A PV-Based Shallow-Water Model on a Hexagonal–Icosahedral Grid

JOHN THUBURN

Centre for Global Atmospheric Modelling, Department of Meteorology, University of Reading, Whiteknights, Reading, United Kingdom

(Manuscript received 11 July 1996, in final form 19 December 1996)

ABSTRACT

A new global shallow-water model has been developed. It uses a hexagonal–icosahedral grid, potential vorticity as a prognostic variable, and a conservative, shape-preserving scheme for advection of mass, potential vorticity, and tracers. A semi-implicit time scheme is used so that the maximum time step for stable integrations is limited by the advection speed rather than the gravity wave phase speed. This combination of numerical methods avoids some of the major problems of more traditional numerical methods, such as pole problems, and spurious oscillations and negatives in advected quantities. Sample results from a standard set of test cases are presented to illustrate the model's performance. In a pure advection test case the model's advection scheme shows good isotropy and phase-speed properties, but it is a little diffusive. In the remaining test cases the model's overall accuracy is comparable to that of other gridpoint models for which results are available. Two sources of error are noted. One is the dissipation inherent in the advection scheme, which is estimated to be significantly stronger than the dissipation usually imposed in climate models of comparable resolution. The other is the grid structure, which leads to conspicuous symmetry errors in test cases where the true solution is symmetrical. The symmetry errors appear to arise because the hexagonal grid boxes are not perfectly regular but are somewhat distorted, particularly in certain regions of the grid, leading to larger truncation errors in the advection scheme in those regions.

1. Introduction

No completely satisfactory numerical method has yet been developed for solving the equations of atmospheric dynamics and constituent transport in spherical geometry. Some of the major problems associated with methods currently in widespread use include Gibbs ripples (or ''spectral ringing'') in spectral models, problems near the poles in gridpoint models using regular latitude–longitude grids, spurious oscillations and unphysical negatives in fields of moisture and chemicals caused by some advection schemes, and lack of conservation with semi-Lagrangian advection schemes. Many methods have been proposed that address one or more of these problems. For example, the use of alternative grids [see Rasch (1994) and Rančić et al. (1996) for recent examples] or the use of semi-Lagrangian advection schemes (e.g., Staniforth and Côté 1991) can reduce pole problems; many advection schemes are available that do not produce spurious oscillations, that is, they are ''shape preserving'' (e.g., Williamson and Rasch 1989; Smolarkiewicz and Rasch 1991; Lin et al. 1994); and conservative semi-Lagrangian schemes have recently been developed (e.g., Leslie and Purser 1995; Lin and Rood 1996; Leonard et al. 1996). Nevertheless, it is worthwhile attempting to develop new numerical models that avoid as many as possible of the problems mentioned above. This paper describes a shallow-water model based on a novel combination of numerical methods. It was designed from the outset to avoid as far as possible the problems mentioned above.

The shallow-water equations describe many of the physical phenomena that are described by the full three-dimensional primitive equations, including slow, balanced vortical motions and fast gravity wave motions. Therefore, developing a model to integrate the shallow-water equations on a sphere presents many of the main challenges of developing a full three-dimensional primitive equation model, including the problem of how to deal with the inevitable singularities that arise (e.g., at the poles) in any particular coordinate system. The shallow-water equations are a useful test bed for any proposed numerical method before moving on to the full complexity of the three-dimensional primitive equations.

The shallow-water model described here uses the following combination of special features: a hexagonal–icosahedral grid; potential vorticity (PV) as a prognostic variable; and a newly developed conservative, shape-preserving, genuinely multidimensional scheme for the advection of mass, PV, and tracers.

Variations on the hexagonal–icosahedral grid have been proposed for modeling atmospheric flows by Wil-

Corresponding author address: Dr. John Thuburn, CGAM, Department of Meteorology, University of Reading, 2 Earley Gate, Whiteknights, Reading RG6 6BB, United Kingdom.
E-mail: swsthubn@met.rdg.ac.uk

liamson (1968), Sadourny et al. (1968), Cullen (1974), and, more recently, Masuda and Ohnishi (1986) and Heikes and Randall (1995a,b). Such grids give nearly uniform coverage of the sphere. In contrast, regular latitude–longitude grids have an accumulation of grid points near the poles. Depending on the time-stepping scheme for the gravity wave terms and on the advection scheme, this can impose excessive restrictions on the allowable time step unless some form of spatial filtering is used. Also, such a distribution of grid boxes is wasteful because a disproportionate amount of computational effort is spent in the polar regions. In addition, the variation in gridbox size could potentially lead to complications in a full global circulation model if parameterization schemes require different tuning for different resolutions. The hexagonal–icosahedral grid avoids all of these problems.

The shallow-water equations in terms of the fluid depth h^*, the PV Q, and the divergence δ, in the absence of friction and mass sources are

$$h^*_t + \nabla\cdot(\mathbf{v}h^*) = 0, \tag{1}$$

$$(h^*Q)_t + \nabla\cdot(\mathbf{v}h^*Q) = 0, \tag{2}$$

$$\delta_t = -\nabla\cdot\left[h^*Q\mathbf{k}\times\mathbf{v} + \nabla\left(gh + \frac{\mathbf{v}^2}{2}\right)\right], \tag{3}$$

where $\mathbf{v}$ is the velocity, $h = h^* + h_0$ is the surface elevation, h_0 is the orography, g is the acceleration due to gravity, and $\mathbf{k}$ is the unit vertical vector. Subscript t indicates a time derivative. In the shallow-water system Q is defined by

$$Q = \frac{f+\zeta}{h^*}, \tag{4}$$

where $\zeta = \mathbf{k}\cdot\nabla\times\mathbf{v}$ is the relative vorticity. The mixing ratio q of any inert tracer satisfies an equation like (2):

$$(h^*q)_t + \nabla\cdot(\mathbf{v}h^*q) = 0. \tag{5}$$

This system of equations is closed by expressing the velocity in terms of the vorticity and divergence via a streamfunction ψ and a velocity potential χ:

$$\nabla^2\psi = \zeta = h^*Q - f, \tag{6}$$

$$\nabla^2\chi = \delta, \tag{7}$$

$$\mathbf{v}_{\rm rot} = \mathbf{k}\times\nabla\psi, \tag{8}$$

$$\mathbf{v}_{\rm div} = \nabla\chi, \tag{9}$$

$$\mathbf{v} = \mathbf{v}_{\rm rot} + \mathbf{v}_{\rm div}. \tag{10}$$

Potential vorticity has several remarkable properties that make it a powerful diagnostic tool for the study of atmospheric dynamics (e.g., Hoskins et al. 1985; Haynes and McIntyre 1990; McIntyre 1992). (i) It is conserved following fluid parcels to the extent that friction and diabatic effects are negligible, so that it labels air parcels. (ii) Under suitable balance assumptions it can be "inverted" to give the wind and mass fields—information about balanced motions is embodied in the PV distribution. (iii) It satisfies a conservation law of the form

$$(h^*Q)_t + \nabla\cdot\mathbf{F} = 0 \tag{11}$$

for some flux $\mathbf{F}$ even when friction and diabatic effects are not negligible—for three-dimensional flow governed by the primitive equations a similar conservation law holds with h^* replaced by $\sigma = -g^{-1}(dp/d\theta)$ and ∇ replaced by ∇_θ, the gradient operator along isentropes. (iv) For three-dimensional flow its behavior can be interpreted as if isentropes were impermeable to a notional substance whose mixing ratio is Q, that is, the cross-isentropic component of $\mathbf{F}$ vanishes. Some of these properties suggest that it might be advantageous to use PV as a prognostic variable in numerical models. For example, it should make PV-based diagnostics easier to derive and interpret. It might even lead to more accurate or efficient models, for example, by providing a cleaner separation between the balanced and unbalanced components of the flow, facilitating the use of the most appropriate numerical methods for each component. However, despite these potential advantages, the use of PV as a prognostic variable has so far been rather limited, notable exceptions being the shallow-water models described by Temperton and Staniforth (1987) and Bates et al. (1995).

Note, incidentally, that (1)–(3) are written in coordinate-independent form. Directly discretizing this form of the equations, without the introduction of coordinates, sidesteps the problem of coordinate singularities that inevitably occur in any global coordinate system on the sphere. Furthermore, the prognostic variables are scalars, not the components of vectors. The velocity is a vector, but when it is used in the model's calculations it is expressed as components normal and tangential to a local box edge. Again this avoids introducing any singularities.

Equations (1), (2), and (5) are conservation laws that can be integrated using an advection scheme [although the mass continuity equation must be treated slightly differently from the other two—compare (19) with (20) below]. An advection scheme is required that is conservative and shape preserving. Because of the grid geometry, the scheme must have these properties for arbitrary two-dimensional gridbox shapes. The method described by Thuburn (1996a) is used to construct such a scheme.

One consequence of using this advection scheme is that no explicit scale-selective dissipation (with its concomitant tunable parameters) is necessary to obtain stable, noise-free integrations of the model. [More generally, explicit scale-selective dissipation can often be omitted when using upwind-biased, forward-in-time advection schemes or other dissipative schemes (e.g., Smolarkiewicz and Margolin 1993; Walsteijn 1994).] It has been argued (Thuburn 1995) that for two-dimen-

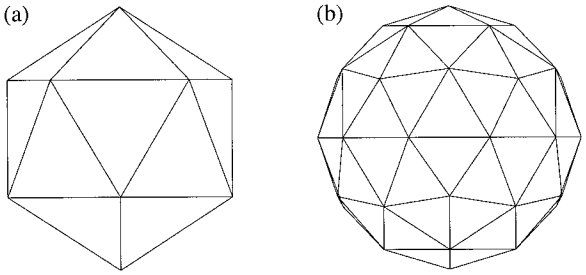

FIG. 1. (a) Regular icosahedron. (b) Grid obtained by subdividing each face of the icosahedron into four triangles and then projecting the new vertices onto the surface of the sphere.

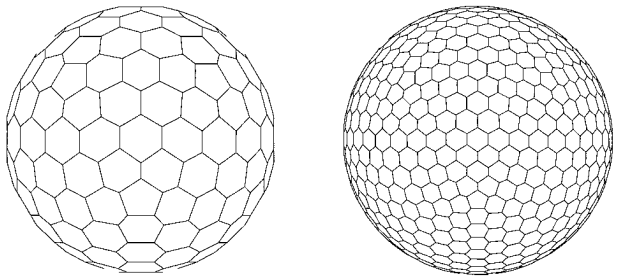

FIG. 2. Model grids 3 and 4 in the hierarchy.

sional and quasi-two-dimensional flow, where the cascade to small scales is an advective process and where the strain rate tends to be dominated by the larger-scale (and hence resolved) eddies, the cascade to unresolved scales can be adequately represented by a shape-preserving advection scheme alone, without the need for any additional parameterization.

The numerical schemes used in the model are described in detail in section 2. The model has been systematically tested on the suite of seven test cases proposed by Williamson et al. (1992). Complete results of those tests are documented by Thuburn (1996b) for comparison with the results for other models that have also been applied to the test cases (e.g., Heikes 1993; Jakob et al. 1993; Malcolm 1994). A few key results are presented in section 3 to illustrate the strengths and weaknesses of the model described here. Some issues that arise in extending the numerical schemes to the three-dimensional primitive equations are discussed in section 4. The results are summarized in section 5.

2. Description of the model

The model's hexagonal–icosahedral grid is described in section 2a. The integration of (1)–(3) involves finding $\mathbf{v}$ from h^*, Q, and δ (section 2b), advecting h^* and Q (section 2c), and evaluating terms on the right-hand side of the divergence equation (section 2d). The time stepping scheme is described in section 2e.

a. The model grid

The grid used here is very similar to the simple twisted icosahedral grid described by Heikes and Randall (1995a) but without the twist. The sphere is covered by pentagonal and hexagonal grid boxes; the number of hexagons depends on the resolution, while there are always exactly 12 pentagons. There is a hierarchy of possible grids of different resolutions, each member of the hierarchy having approximately four times as many grid boxes as the previous member. The first member of the hierarchy is the regular dodecahedron.

The starting point for constructing the grid is the regular icosahedron (Fig. 1a). Given a grid of triangles, a new finer grid of triangles is generated by placing new vertices at the midpoints of the existing edges and then projecting these new vertices onto the surface of the sphere (Fig. 1b). This process may be repeated indefinitely to obtain higher-resolution grids.

The grid actually used by the model is the Voronoi grid associated with the triangular grid just described (Augenbaum and Peskin 1985) (Fig. 2). Each vertex of the triangular grid corresponds to a face, that is, a grid box, of the Voronoi grid; each face of the triangular grid corresponds to a vertex of the Voronoi grid; each edge of the triangular grid corresponds to an edge of the Voronoi grid. The edges of the Voronoi grid are the perpendicular bisectors of the edges of the triangular grid. Note that the hexagonal grid boxes vary somewhat in their exact shape and size. The most distorted hexagons are those immediately adjacent to the pentagons. The pentagons, however, are perfectly regular, and the grid has a fivefold symmetry relative to any one of the 12 pentagons.

All the information about the grid needed by the model, such as the areas of the grid boxes and the lengths of the edges, is computed and tabulated before running the model. Some of the properties of the grid as functions of resolution are shown in Table 1.

Since this work was begun, an important advance has been made by Heikes and Randall (1995b) in the design

TABLE 1. Properties of the model grid for different resolutions. Here N_{grid} indicates which grid in the hierarchy; N_{box} is the number of grid boxes; N_{eq} is the number of grid boxes around the equator; A_{ave} is the average grid box area; A_{min} is the minimum gridbox area; A_{max} is the maximum grid box area; d_{ave} is the average distance between the centers of neighborings grid boxes; d_{min} is the minimum distance between the centers of neighboring grid boxes; d_{max} is the maximum distance between the centers of neighboring grid boxes; and T_{equiv} is the approximate equivalent spectral truncation based on the number of degrees of freedom.

N_{grid}	N_{box}	N_{eq}	A_{ave} (km^2)	$\frac{A_{\text{min}}}{A_{\text{max}}}$	d_{ave} (km)	$\frac{d_{\text{min}}}{d_{\text{max}}}$	T_{equiv}
1	12	—	4.25×10^7	1.0	7053	1.0	T2–T3
2	42	10	1.215×10^7	0.884	3765	0.881	T6
3	162	20	3.149×10^6	0.842	1914	0.848	T12
4	642	40	7.945×10^5	0.764	961	0.839	T24
5	2562	80	1.991×10^5	0.742	481	0.837	T50
6	10 242	160	4.980×10^4	0.736	241	0.837	T100
7	40 962	320	1.245×10^4	0.735	120	0.837	T201

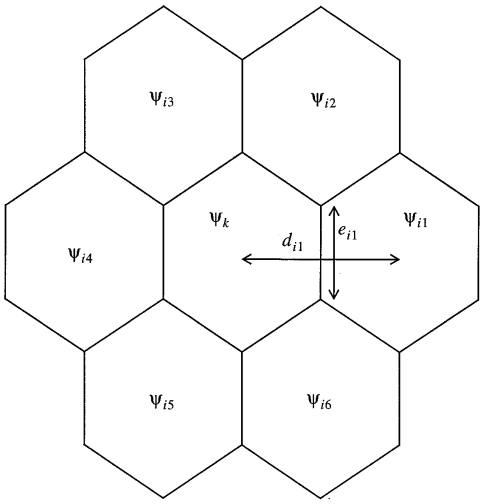

FIG. 3. Schematic showing the quantities used to calculate the finite-difference Laplacian (12). The index i in (12) runs over $i1$, $i2$, . . . , $i6$.

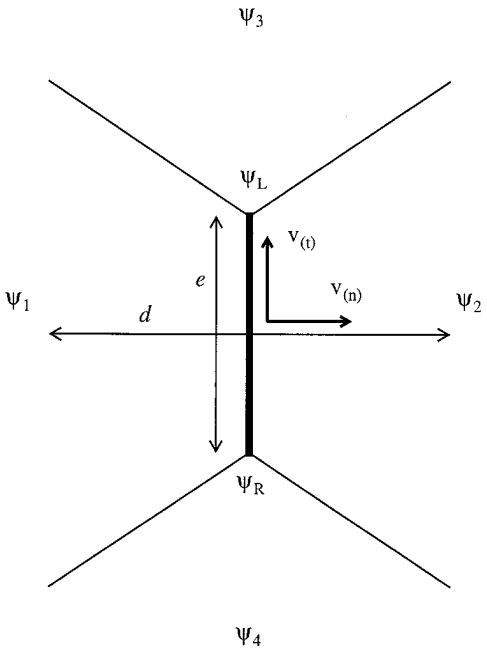

FIG. 4. Schematic showing the ψ values used to calculate the normal and tangential components of the rotational wind at the box edge indicated by the heavy line [(13) and (14)]. Here ψ_1, . . . , ψ_4 are values at the box centers; ψ_L and ψ_R are values at vertices.

of the grid. They noted that with the simple grid described above (or with their twisted version of it), the error obtained in evaluating a Laplacian using the natural finite-difference expression given in section 2b below does not converge to zero uniformly as resolution is increased. The problem is worst near the most distorted hexagons. Heikes and Randall describe an improved grid-generation algorithm that gives a hierarchy of grids on which the error in evaluating the Laplacian does appear to converge to zero uniformly as resolution is increased. Similar improvements to the grid would be highly desirable in future versions of the model described here.

b. Relationship between the winds and the vorticity and divergence

In the continuous equations, the vorticity and divergence are related to the streamfunction and velocity potential by (6) and (7). A discrete approximation to the Laplacian operator is given by

$$\frac{1}{A_k}\sum_i \frac{e_i}{d_i}(\psi_i - \psi_k) = \zeta_k, \tag{12}$$

(see Fig. 3), where the index i runs over the five or six edges and boxes surrounding box k. The term e_i is the length of the ith edge, d_i is the distance between the center of box k and the center of box i, and A_k is the area of box k. This approximation would be second-order accurate on a regular hexagonal grid. However, as pointed out by Heikes and Randall (1995b), on the simple hexagonal–icosahedral grid it is locally less than first-order accurate. Given a global vorticity field ζ, the elliptic equation (12) is solved for ψ using a multigrid relaxation method. The details are given in appendix A. The velocity potential χ is found from δ in a similar way.

In the continuous equations, ψ and χ are related to the rotational and divergent components of the wind by (8) and (9). The discrete form of (8) is (see Fig. 4)

$$v_{\mathrm{rot}(n)} = \frac{\psi_R - \psi_L}{e}, \tag{13}$$

$$v_{\mathrm{rot}(t)} = \frac{\psi_2 - \psi_1}{d}, \tag{14}$$

where $v_{\mathrm{rot}(n)}$ and $v_{\mathrm{rot}(t)}$ are the normal and tangential components of the rotational wind at the face in question (positive means in the direction of the arrows). Similarly,

$$v_{\mathrm{div}(n)} = \frac{\chi_2 - \chi_1}{d}, \tag{15}$$

$$v_{\mathrm{div}(t)} = \frac{\chi_L - \chi_R}{e}, \tag{16}$$

for the divergent wind. Note that this discretization is consistent in the sense that the divergence of the rotational wind and the vorticity of the divergent wind are both identically zero, that is, $\Sigma_i v_{\mathrm{rot}(n)i} e_i = 0$ and $\Sigma_i v_{\mathrm{div}(t)i} e_i = 0$, where i runs over the edges of a grid box and the normal velocity component is taken to be positive outward and the tangential component is taken to be positive anticlockwise in these sums.

The elliptic solver gives the streamfunction and velocity potential only in the grid boxes, ψ_1, . . . , ψ_4, not at the vertices, ψ_L, ψ_R. There are various ways to define ψ_L, ψ_R in terms of ψ_1, . . . , ψ_4. Masuda and Ohnishi (1986) and Heikes and Randall (1995a) use the formula

$$\psi_R = \frac{\psi_1 + \psi_2 + \psi_4}{3}, \tag{17}$$

etc., because this conserves energy and potential enstrophy for nondivergent flow with their schemes. However, the distorted gridbox shapes significantly reduce the accuracy of this formula, and it gives noticeable wind errors in some simple tests (e.g., specifying a streamfunction proportional to the sine of latitude and computing the wind). The model described here uses instead a value given by fitting a two-dimensional linear function through the three nearest grid boxes:

$$\psi_R = \alpha_1\psi_1 + \alpha_2\psi_2 + \alpha_4\psi_4. \tag{18}$$

The coefficients α_1, α_2, . . . , vary from vertex to vertex. They are computed and tabulated along with all the other information about the grid. The schemes described in this paper do not formally conserve energy or potential enstrophy (but see section 3b), even with the choice (17), so the conservation properties are not affected by using (18) instead of (17).

c. The advection scheme

The advection scheme is similar to that described by Thuburn (1995). It is conservative and uses a genuinely multidimensional extension of Leonard's (1991) ''universal limiter'' to ensure that there is no spurious amplification of extrema and that there is minimal distortion of advected profiles.

Discrete forms of (1) and (2) are

$$h_k^{*m+1} = h_k^{*m} + \sum_{\text{in}} c_{(n)i}\hat{h}_i^* - \sum_{\text{out}} c_{(n)j}\hat{h}_j^*, \tag{19}$$

$$Q_k^{m+1} = \frac{Q_k^m + \sum_{\text{in}} \tilde{c}_{(n)i}\hat{Q}_i + \sum_{\text{out}} \tilde{c}_{(n)j}\hat{Q}_j}{1 + \sum_{\text{in}} \tilde{c}_{(n)i} - \sum_{\text{out}} \tilde{c}_{(n)j}}. \tag{20}$$

Here h_k^{*m} is the (area-weighted average) fluid depth in box k at time step m, Q_k^m is the (mass-weighted average) PV in box k at time step m, Σ_{in} means a sum over the inflow edges of box k and Σ_{out} means a sum over the outflow edges, and quantities with a circumflex are values at the box edges to be defined below. The value $c_{(n)i}$ is a Courant number for the flow normal to edge i,

$$c_{(n)i} = \frac{|v_{(n)i}|e_i\Delta t}{A_k}, \tag{21}$$

and $\tilde{c}_{(n)i}$ is a modified Courant number,

$$\tilde{c}_{(n)i} = \frac{c_{(n)i}\hat{h}_i^*}{h_k^{*m}}. \tag{22}$$

Note that each edge has two Courant numbers, one with respect to the box on each side.

Equations (19) and (20) ensure that the scheme conserves the total mass of fluid $\Sigma_k A_k h_k^*$ and the total vorticity $\Sigma_k A_k h_k^* Q_k$, irrespective of how $\hat{h}^*$ and $\hat{Q}$ are defined. The values of $\hat{h}^*$ and $\hat{Q}$ can be chosen to make the scheme accurate and shape preserving. First preliminary values of $\hat{h}^*$ and $\hat{Q}$ are chosen using a basic advection scheme, then these values are modified if necessary to ensure shape preservation.

The detailed formulas for computing the preliminary values of $\hat{h}^*$ and $\hat{Q}$ are given in appendix B. The concept behind this scheme is similar to that behind the ''UTOPIA'' (uniformly third-order polynomial interpolation algorithm) scheme (Leonard et al. 1993; see also Rasch 1994). A two-dimensional quadratic function is fitted through a set of boxes in an upstream-biased neighborhood of the edge in question, then the amount of advected substance swept across that edge in one time step is estimated by integrating the quadratic over the appropriate area. Note that the formula (B1) differs slightly from that used by Thuburn (1995) through the inclusion of the second term on the right-hand side. In that paper, the quadratic was fitted so that, on a regular hexagonal grid, it gave the values *a*, *b*, *c*, etc., (see Fig. 18) at the gridbox centers. In this paper the quadratic is fitted so that, on a regular hexagonal grid, it gives the values *a*, *b*, *c*, etc., as gridbox averages. Tests have shown that this refinement reduces the already small phase errors of the scheme to virtually zero (see section 3a). However, it has very little effect on the dissipation inherent in the scheme. Also, the interpretation of *a*, *b*, *c*, etc., as box average values is required for the idea of conservation to make sense, and the new scheme is more consistent with this. The formula (B2) for the outflow edge of a pentagonal box, however, is based on fitting the quadratic to give *a*, *b*, *d*, etc., as box center values rather than box average values simply because the algebra required to fit the quadratic to box average values is even more difficult.

On a plane grid of regular hexagons with an advecting wind that is constant in space and time, this advection scheme would be third-order accurate in space and time. When used on the hexagonal–icosahedral grid, which does not consist solely of regular hexagons, or when the advecting wind varies in space or time, the scheme is formally no more than first-order accurate in space and time. However, in practice it is found to be much more accurate than a simple ''donor cell'' first-order upwind scheme (see section 3a).

After the preliminary edge values have been calculated using the basic scheme, they are adjusted, if necessary, to ensure that extrema are not amplified and no spurious extrema are generated. This flux-limiting step is exactly as described by Thuburn (1995, 1996a) and will not be described here.

Note that the algorithm for choosing $\hat{h}^*$ is exactly the same as the algorithm for choosing $\hat{Q}$, except that $c_{(n)}$, $c_{(t)}$ are used instead of $\tilde{c}_{(n)}$, $\tilde{c}_{(t)}$. But note too that the update formulas are different [compare (19) with (20)].

d. The divergence equation

This section describes the spatial discretization of the terms in the divergence equation. The second term on

the right-hand side of (3) can be written as $-\nabla^2 E$, where $E = gh + \mathbf{v}^2/2$. In grid box k,

$$E_k = gh_k + \frac{(\mathbf{v}^2)_k}{2}. \tag{23}$$

The value of $\mathbf{v}^2$ in grid box k is obtained by taking a simple mean of $\mathbf{v}^2$ around the edges of the box, $\mathbf{v}$ at the edges of the box being given by (13)–(16). The Laplacian is then evaluated by a formula analogous to (12).

The first term on the right-hand side of (3) is evaluated for box k as

$$\frac{1}{A_k}\sum_i \zeta_{ai} v_{(t)i} e_i, \tag{24}$$

where i runs over the five or six edges of box k. Here $v_{(t)i}$ is the tangential component of the velocity at edge i, taken to be positive when anticlockwise about box k. The absolute vorticity at edge i, ζ_{ai}, is given by the simple mean of the absolute vorticity $\zeta_a = h^*Q$ in the boxes either side of face i.

e. Time-stepping scheme

The basic idea behind the time-stepping scheme is to use an explicit forward time step for the advection of h^* and Q and a semi-implicit step to advance δ and modify h^*. In this way, the permissible time step for stable integrations is limited by the advection speed rather than the gravity wave phase speed, and this allows time steps about three to five times longer than would be possible with a fully explicit scheme under typical conditions. The simplest such scheme is

$$\frac{h^{*m+1} - h^{*m}}{\Delta t} = -\nabla\cdot(\mathbf{v}^m h^{*m}) + \beta h_{\text{ref}}(\delta^m - \delta^{m+1}) \tag{25}$$

$$\frac{\delta^{m+1} - \delta^m}{\Delta t} = -\nabla\cdot\left\{\zeta_a^m \mathbf{k}\times\mathbf{v}^m + \nabla\left[(1-\beta)gh^{*m} + \beta g h^{*m+1} + gh_0 + \frac{(\mathbf{v}^m)^2}{2}\right]\right\} \tag{26}$$

$$\frac{(h^*Q)^{m+1} - (h^*Q)^m}{\Delta t} = -\nabla\cdot[(\mathbf{v}h^*)^{(\text{SI})} Q^m]. \tag{27}$$

The coefficient β controls the semi-implicit correction: $\beta = 0$ would give an explicit scheme, while $\beta = 1$ would give a backward Euler semi-implicit correction. In all the test cases described here $\beta = 0.5$. The semi-implicit correction corresponds to the mass continuity equation linearized about the reference depth h_{ref}. For stability, the gravity wave speed associated with the semi-implicit scheme should be greater than the fastest gravity wave speed in the model, including the nonlinear effect of advection:

$$(gh_{\text{ref}})^{1/2} > |\mathbf{v}| + (gh^*)^{1/2}. \tag{28}$$

However, if h_{ref} is excessively large, then the scheme becomes less accurate. Therefore, h_{ref} should be chosen as small as possible while allowing a comfortable safety margin in (28). The value $(\mathbf{v}h^*)^{(\text{SI})}$ in (27) is the mass flux after allowing for the semi-implicit correction:

$$(\mathbf{v}h^*)^{(\text{SI})} = \mathbf{v}^m h^{*m} + \beta h_{\text{ref}}(\mathbf{v}_{\text{div}}^{m+1} - \mathbf{v}_{\text{div}}^m). \tag{29}$$

It is essential to use this mass flux for the advection of PV and tracers, for consistency with the mass continuity equation, to ensure conservation. However, small errors in conservation still do occur via the semi-implicit correction because the finite number of iterations taken by the elliptic solver mean that $\nabla\cdot(\mathbf{v}_{\text{div}}^{m+1})$ does not exactly equal δ^{m+1}.

Unfortunately, the scheme (25), (26), (27) allows two kinds of instability. Both of them are really just manifestations of the well-known instability of the forward Euler time scheme. The first kind of instability is essentially the Rossby wave instability described by Bates et al. (1995) in which parcel displacements grow exponentially, at least while the disturbance is linear, even though the PV values themselves remain bounded. It occurs because $\mathbf{v}_{\text{rot}}$ at time level m is used to advect PV. The second kind of instability is a Rossby–gravity-mode instability. It occurs if either the advection of vorticity by the divergent wind or the vorticity and rotational wind terms in the divergence equation are evaluated at time level m. For both kinds of instability, the fastest-growing modes are global in scale and so are not affected by spatial resolution. To ameliorate these instabilities, the scheme (25), (26), (27) is modified to replace certain terms at time level m by estimates of their values at the intermediate time level $m + 1/2$:

$$\frac{h^{*m+1} - h^{*m}}{\Delta t} = -\nabla\cdot(\mathbf{v}^{(+)} h^{*m}) + \beta h_{\text{ref}}(\delta^m - \delta^{m+1}) \tag{30}$$

$$\frac{\delta^{m+1} - \delta^m}{\Delta t} = -\nabla\cdot\left\{\zeta_a^{(+)}\mathbf{k}\times(\mathbf{v}_{\text{rot}}^{(+)} + (1-\beta)\mathbf{v}_{\text{div}}^m + \beta\mathbf{v}_{\text{div}}^{m+1}) + \nabla\left[(1-\beta)gh^{*m} + \beta g h^{*m+1} + gh_0 + \frac{\mathbf{v}^{(+)2}}{2}\right]\right\} \tag{31}$$

$$\frac{(h^*Q)^{m+1} - (h^*Q)^m}{\Delta t} = -\nabla\cdot[(\mathbf{v}h^*)^{(+\mathrm{SI})}Q^m]. \quad (32)$$

The quantity $\zeta_a^{(+)}$ is a cheap estimate of the absolute vorticity at step $m + 1/2$. It is obtained by advecting the absolute vorticity for one-half of a time step using the wind $\mathbf{v}_{\mathrm{rot}}^m + \mathbf{v}_{\mathrm{div}}^m$ and using a cheap (second-order centered difference) advection scheme. The value $\mathbf{v}_{\mathrm{rot}}^{(+)}$ is the rotational wind corresponding to $\zeta_a^{(+)}$ and $\mathbf{v}^{(+)} = \mathbf{v}_{\mathrm{rot}}^{(+)} + \mathbf{v}_{\mathrm{div}}^m$. The value $(\mathbf{v}h^*)^{(+\mathrm{SI})}$ is the mass flux after allowing for the semi-implicit correction:

$$(\mathbf{v}h^*)^{(+\mathrm{SI})} = \mathbf{v}^{(+)}h^{*m} + \beta h_{\mathrm{ref}}(\mathbf{v}_{\mathrm{div}}^{m+1} - \mathbf{v}_{\mathrm{div}}^m). \quad (33)$$

The use of $\zeta_a^{(+)}$, $\mathbf{v}^{(+)}$, etc., in place of ζ_a^m, $\mathbf{v}^m$, etc., makes the relevant parts of the time scheme resemble a second-order Runge–Kutta scheme. When applied to the ordinary differential equation $\dot{F} = i\omega F$, the second-order Runge–Kutta scheme gives a solution that amplifies in magnitude by a factor of $1 + (\omega\Delta t)^4/4$ per step. Thus, the scheme is still unstable, but only weakly. For example, taking $\omega = 2\pi\ \mathrm{day}^{-1}$ corresponding to the highest frequency planetary Rossby mode for nondivergent flow on a sphere and taking a time step of one hour gives an *e*-folding time of 71 days for the instability. This is long enough for the instability to be negligible in virtually all practical applications. The *e*-folding time increases rapidly like Δt^{-3} for even shorter time steps. In a trial integration on grid 5 with $\Delta t = 1800$ s, starting from the initial conditions of section 3c below, the model ran for 200 days with no sign of instability. In the discussion below this time scheme will be referred to as the ''operational'' time scheme.

In all of the tests presented in section 3 below, an inferior time scheme was used in which $\zeta_a^{(+)}$ is obtained in two stages. In the first stage the absolute vorticity is advected for one-half of a time step using only the rotational component of the wind, $\mathbf{v}_{\mathrm{rot}}^m$, to give ζ_a'. In the second stage the resulting vorticity is modified to make some allowance for vortex stretching by the divergent flow over one-half of a time step:

$$\zeta_a^{(+)} = \frac{\zeta_a'(h^{*m} - \delta^m h_{\mathrm{ref}}\Delta t/2)}{h^{*m}}. \quad (34)$$

The modification (34) allows for vortex stretching by the divergent flow but not for advection along vorticity gradients by the divergent flow. As a result, the Rossby–gravity-mode instability is still significant with this scheme; in a trial integration on grid 5 with $\Delta t = 1800$ s, starting from the initial conditions of section 3c, the unstable mode had an *e*-folding time of about 15 days and caused the model to ''blow up'' after 72 days. However, in the shorter test integrations presented in section 3, this instability was not a conspicuous problem.

Equations (30) and (31) both involve terms on the right-hand side at step $m + 1$. These equations can be combined to give an equation of the form

$$\nu^2\nabla^2\delta^{m+1} - \delta^{m+1} = R + \beta\Delta t\nabla\cdot(\zeta_a^{(+)}\mathbf{k} \times \mathbf{v}_{\mathrm{div}}^{m+1}), \quad (35)$$

where $\nu^2 = \beta^2\Delta t^2 g h_{\mathrm{ref}}$ and R is a known function of values at step m and intermediate values [superscript $(+)$]. Equation (35) can be solved using a multigrid iterative method. The presence of the $\mathbf{v}_{\mathrm{div}}^{m+1}$ term in (35) makes the solution of the equation rather less straightforward than the solution of Poisson's equation; see appendix A for details.

In total, five elliptic problems must be solved at each time step to find $\mathbf{v}_{\mathrm{rot}}^m$, to find $\mathbf{v}_{\mathrm{div}}^m$, to find $\mathbf{v}_{\mathrm{rot}}^{(+)}$, and to make the semi-implicit correction, which involves solving for δ^{m+1} and $\mathbf{v}_{\mathrm{div}}^{m+1}$ together. In principle this could be reduced to four elliptic problems per step by using $\mathbf{v}_{\mathrm{div}}^{m+1}$ found during one step as $\mathbf{v}_{\mathrm{div}}^m$ for the next step. This has not been done in the current implementation of the model because it would require extra fields to be saved to ensure that numerically identical results are obtained independently of whether or not the model is stopped and restarted.

In summary, the sequence of operations involved in one time step is as follows. The rotational and divergent wind $\mathbf{v}_{\mathrm{rot}}^m$, and $\mathbf{v}_{\mathrm{div}}^m$ are diagnosed. The intermediate values $\zeta_a^{(+)}$, $\mathbf{v}_{\mathrm{rot}}^{(+)}$, and $\mathbf{v}^{(+)}$ are calculated. Then the first term on the right-hand side of (30) is evaluated using the advection scheme described in section 2c. Then (35) is solved for δ^{m+1}. Substitution of δ^{m+1} in (30) then gives h^{*m+1}. Finally, the mass flux (33) is computed, and the PV and any tracers are updated using the advection scheme described in section 2c.

3. Results of some test cases

The suite of test cases proposed by Williamson et al. (1992) has been carried out using the model described in this paper. Complete results of these test cases, including all the diagnostics requested by Williamson et al., are presented in the report by Thuburn (1996b). This report is available from the author, or in postscript format on the World Wide Web at http://www.met.rdg.ac.uk/ugamp/publications/jt. Some selected results are presented here to illustrate the strengths and weaknesses of the model. The test cases themselves are described in detail by Williamson et al. (1992); only very brief descriptions are given here.

Some of the test cases do not have analytic solutions so errors must be estimated by comparing the computed solution with a reference solution from a high-resolution (T213) spectral model. These reference solutions, truncated to T63, were obtained in netCDF format from the anonymous ftp account sage.cgd.ucar.edu, along with software to interpolate them to any desired grid.

The time step used in the test cases described below is 3600 s on grid 5 and 1800 s on grid 6, except in the Rossby–Haurwitz wave test case where, because of the stronger winds, a time step of 900 s is used on grid 6.

An indication of the CPU cost of the model when integrating the full shallow-water equations is given in

TABLE 2. CPU cost as a function of resolution on one processor of a Cray YMP8 computer. Here N_{grid} indicates which grid in the hierarchy; Δt is the time step.

N_{grid}	Δt (s)	CPU (seconds per model day)
4	3600	1.55
5	3600	5.33
6	1800	34.26

Table 2. In the current implementation all relevant inner loops ''vectorize'' on machines with pipeline architecture, giving moderately good performance (e.g., about 73 Mflops on one processor of a Cray YMP8). No attempt has been made to optimize the code for use on a distributed memory massively parallel machine. The CPU cost is dominated by the solution of Poisson's equation and the elliptic problem that arises in taking a semi-implicit time step, which together account for about 70% of the cost. The CPU cost could be reduced, possibly with some loss of accuracy, by taking fewer iterations in the multigrid solver (appendix A). Despite the complexity of the advection scheme, it accounts for only about 15% of the CPU cost.

a. Advection of a cosine bell

This test does not deal with the full shallow-water equations but tests only the advection scheme. The time stepping of the PV and divergence is switched off and the depth is advected by a specified wind field. The wind field is a solid-body rotation that makes one complete revolution in 12 days. The initial depth field is a compact cosine bell of height 1000 m and radius $a/3$, where a is the earth's radius.

Figure 5 shows the computed solutions, true solutions, and errors for two different model resolutions when the axis of the specified wind field passes through the poles (i.e., through two antipodal pentagonal boxes). Generally the advection scheme performs very well. There are no undershoots or overshoots. The phase error is extremely small, and there is little distortion of the advected profile, only a slight elongation in the direction of the flow on grid 5. On grid 5 there is significant erosion of the maximum to about one-half of its original height. On grid 6 the maximum is eroded by about 150 m. Most of this erosion is due to the basic quasi-third-order upwind advection scheme rather than the flux limiter. Additional tests show that the errors are hardly sensitive at all to the orientation of the axis of the flow relative to the grid; in particular, there are no pole problems. Other diagnostics confirm that the scheme is conservative and maintains the minimum value of zero. They also show the errors growing steadily with time; there are no particular places where the errors are suddenly amplified or where there are large oscillations in the errors. The advection scheme used here compares well with other Eulerian advection schemes for which this test case has been documented (e.g., Jakob et al. 1993; Malcolm 1994; Rasch 1994; Heikes and Randall 1995a).

b. Zonal flow over an isolated mountain

The initial state for this test case is solid-body zonal flow, making one complete revolution in 12 days, with a balanced depth field, flowing over a conical hill of height 2000 m and radius 20° centered at (30°N, 90°W). The subsequent evolution of the flow has no analytical solution, so it is compared with a reference solution from a high-resolution spectral model.

Figure 6 shows the surface elevation field at days 5, 10, and 15 of an integration on grid 6, and Fig. 7 shows differences from the reference solution. The elevation errors are generally quite small (compared, for example, with the equator-to-pole elevation difference), with the maximum error less than 30 m on day 15. The large-scale flow features forced by the mountain are well resolved at this resolution and even at lower resolution. A circular pattern of small-scale noise is visible in the error maps at the location of the mountain. This arises because the reference solution is stored in the form of spectral coefficients at T63 truncation and has small oscillations near the edge of the mountain. It does not signify any problem with the computed solution. (Compare Fig. 6 here with Figs. 5.1 and 5.2 of Jakob et al. 1993.)

The continuous equations have a number of global invariants, including

mass

$$\frac{I[h^*(t)] - I[h^*(0)]}{I[h^*(0)]};$$

energy

$$\frac{I[E(t)] - I[E(0)]}{I[E(0)]}, \quad \text{where } E = \frac{1}{2}h^*\mathbf{v}^2 + \frac{1}{2}g(h^2 - h_0^2);$$

and

potential enstrophy

$$\frac{I[P(t)] - I[P(0)]}{I[P(0)]}, \quad \text{where } P = \frac{1}{2}h^*Q^2.$$

Here $I[x]$ is shorthand for the mean of x over the surface of the sphere, and the three invariants have been normalized. Discrete approximations to these quantities are obtained by replacing the global integral by a discrete summation over grid boxes. But note that the discrete energy and potential enstrophy are only the resolved contributions to the total; they neglect any implied contributions due to subgrid-scale variations in h^*, $\mathbf{v}$, and Q. The discrete mass, on the other hand, does equal the total because the value of h^* in a grid box is interpreted as a gridbox average.

Figure 8 shows the normalized discrete mass, energy,

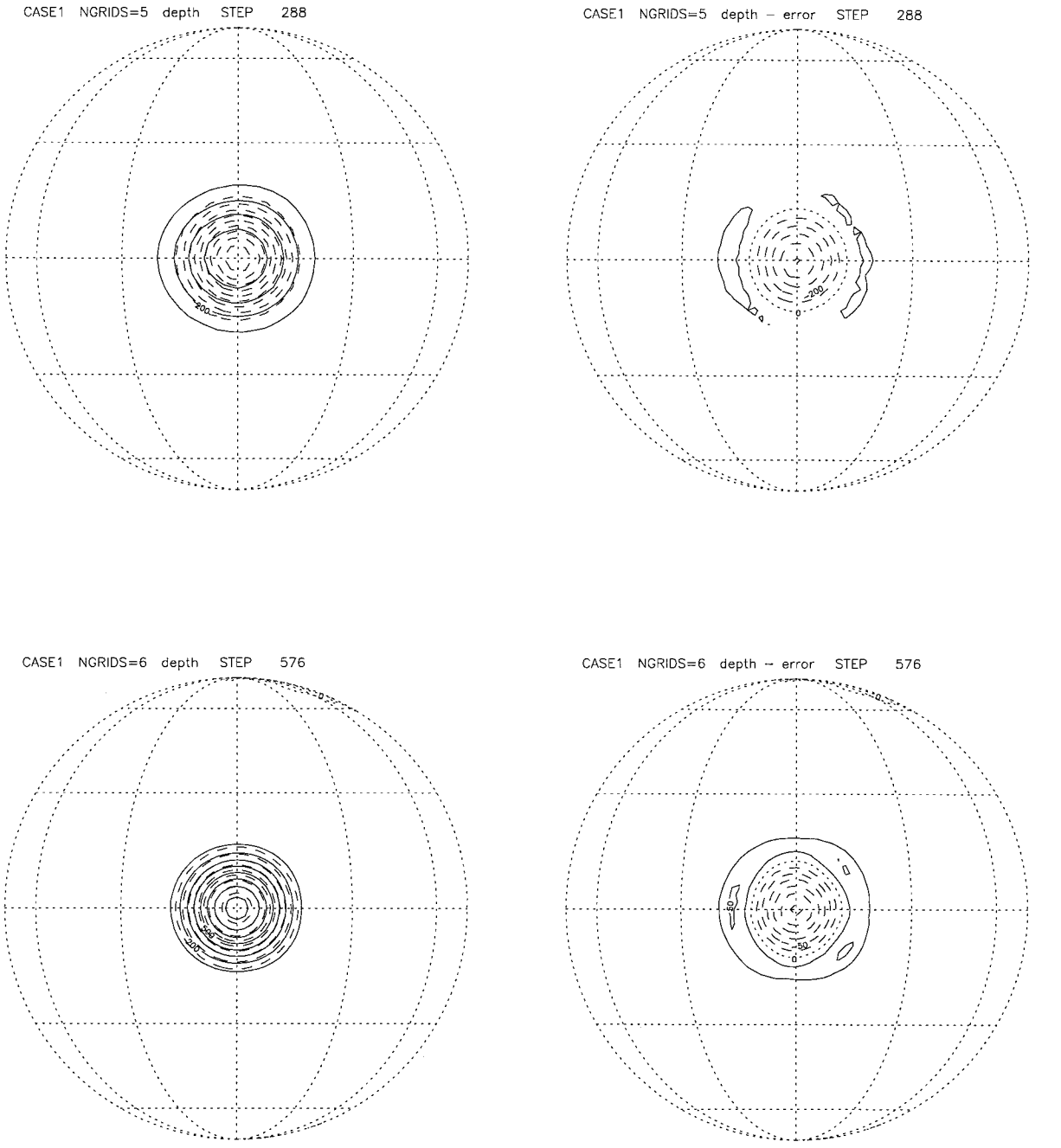

FIG. 5. Results of the advection test case. Upper left: grid 5, computed solution (solid contours) and exact solution (dashed contours), contour interval 100 m. Upper right: grid 5, computed solution minus exact solution, contour interval 100 m. Lower left: grid 6, computed solution (solid contours) and exact solution (dashed contours), contour interval 100 m. Lower right: grid 6, computed solution minus exact solution, contour interval 25 m.

and potential enstrophy as functions of time for this test case. The steps in the curves indicate the precision of the diagnostics as written out by the model. The mass is accurately conserved. Although the schemes do not formally conserve the resolved energy, it is very nearly constant. (At lower resolution there is a small decrease in the resolved energy with time.) There is some decrease in the resolved potential enstrophy with time. Note that the true solution will involve a cascade of potential enstrophy to small scales, some of which will be unresolved at any given resolution. But it is a far from trivial task to calculate what the correct resolved potential enstrophy should be at any given resolution, even with the aid of a high-resolution reference solution.

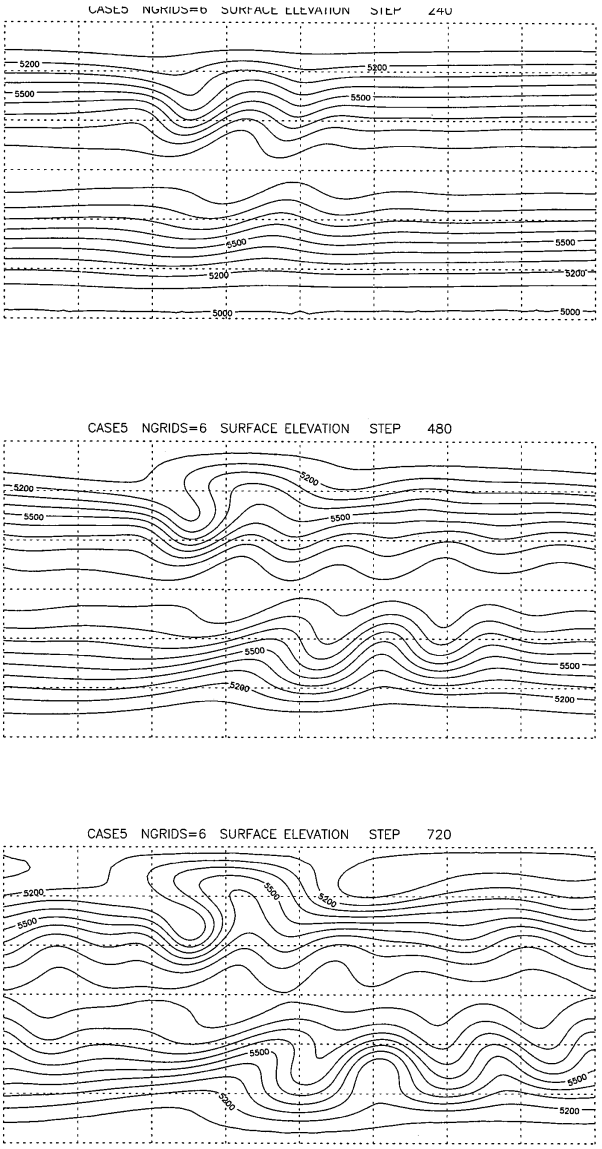

FIG. 6. Longitude–latitude maps of surface elevation for the flow over a mountain test case, contour interval 100 m. Top: day 5. Middle: day 10. Bottom: day 15.

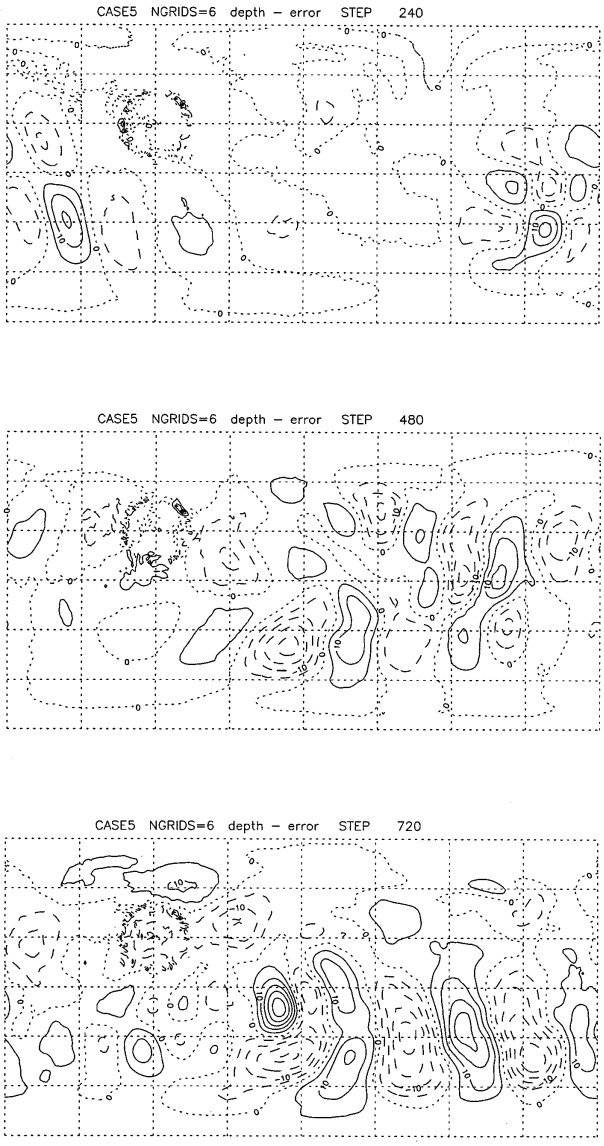

FIG. 7. As in Fig. 6 except that the field shown is the computed solution minus the reference solution, contour interval 5 m.

c. Rossby–Haurwitz wave

This test case is a zonal wavenumber 4 Rossby–Haurwitz wave. The initial state is an exact steadily propagating solution of the nondivergent shallow-water equations. Although it is not an exact solution of the full shallow-water equations, its use by numerous authors has made it a de facto standard test case. Reference integrations with a high-resolution spectral model show that the solution should remain stable and propagate steadily, maintaining its wavenumber 4 structure with only slight vacillations in shape. However, as discussed by Bates et al. (1995), the true solution to this test case is not known, and the details of the vacillations in high-resolution reference solutions depend on the dissipation used in the calculation. This caveat should be borne in mind when interpreting ''error'' maps calculated using such reference solutions.

Figures 9 and 10 show the depth field and its difference from the reference solution at days 1, 7, and 14 of a run on grid 6. The phase speed of the wave is well captured throughout the run. The wave's structure is well simulated at day 1, and even at day 7, but it becomes significantly disrupted by day 14. In particular, the longitudinal and cross-equatorial symmetry of the wave structure has become disrupted. Some symmetry remains, however; the fields in the two hemispheres are related by a reflection across the equator followed by a rotation of 180° in longitude.

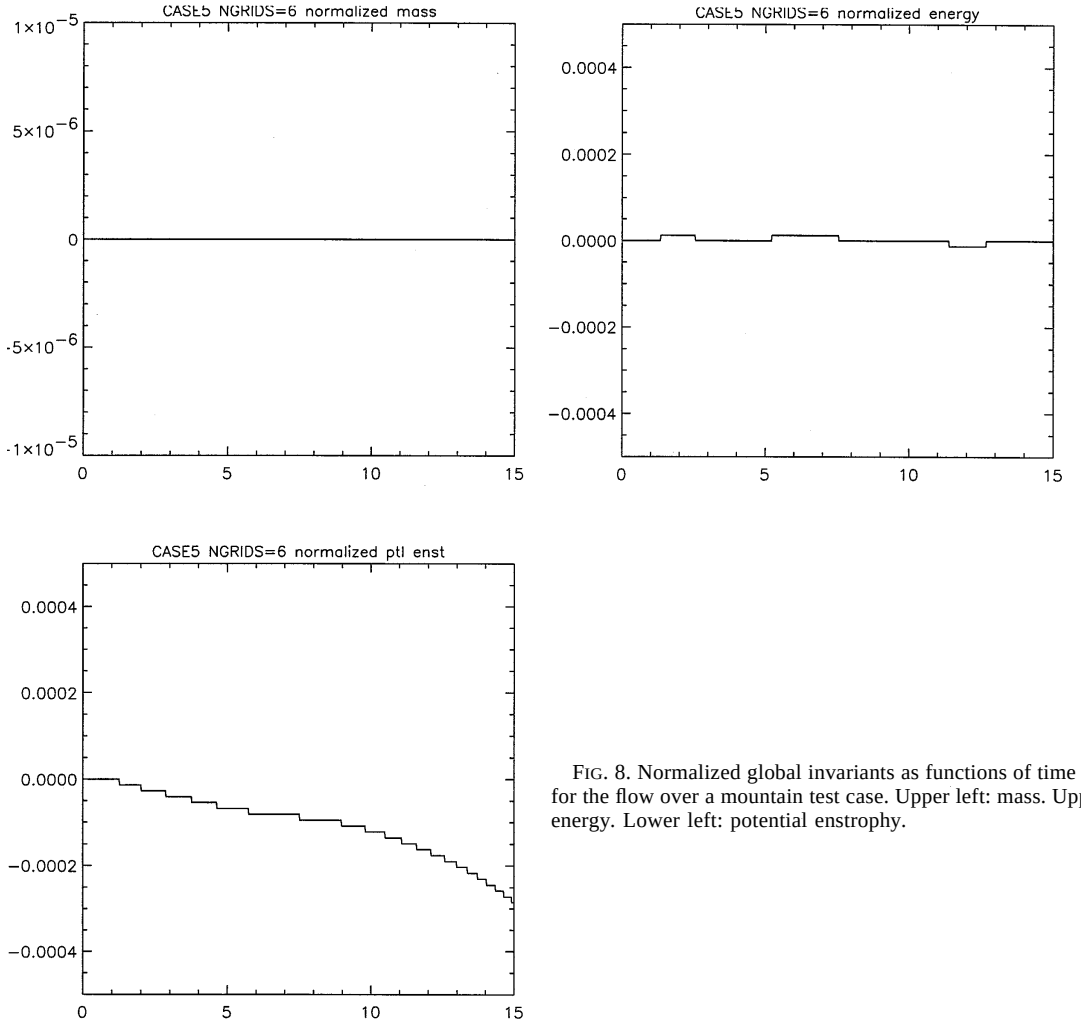

FIG. 8. Normalized global invariants as functions of time (in days) for the flow over a mountain test case. Upper left: mass. Upper right: energy. Lower left: potential enstrophy.

These symmetry errors are obviously related to the structure of the model grid. (Repeating the test case for other grid orientations confirms this.) For this very reason, Heikes and Randall (1995a) chose to use a twisted version of the hexagonal–icosahedral grid, so that a solution initially symmetric about the equator would at least remain symmetric about the equator. The zonal wavenumber 1 pattern at high latitudes in the depth error at day 14 (Fig. 10) arises through an interaction between the wavenumber 4 flow pattern and the fivefold symmetry of the grid. (Repeating the test case using the operational time scheme of section 2e gave virtually identical results, showing that the wavenumber 1 pattern is not associated with an instability of the time scheme.) The distortion of the grid boxes degrades the accuracy of at least two parts of the model's algorithm: the calculation of the Laplacian and its inverse (section 2a), and the advection scheme (section 2c). Repeating the test case with a shorter time step of 300 s (maximum Courant number about 0.17) revealed an interesting relationship between the symmetry errors and the time step. The depth field at day 7 and day 14 (Fig. 11) is much more symmetrical than in the case with a 900-s time step (maximum Courant number about 0.5). This suggests that the symmetry errors may arise through an inhomogeneous pattern of time truncation errors related to the grid structure. This probably occurs through the advection scheme, where both the space *and* time truncation are degraded from quasi-third order to first order by the distortion of the grid boxes.

Figure 12 shows the initial PV for this test case along with the PV at day 7 from the run with the 900-s time step and and PV at day 7 derived from the reference solution depth and vorticity. The most conspicuous difference between the model and the reference solution is that the reference solution PV is rather noisy. (This may be because the reference integration itself is noisy, or it may arise partly through the truncation of the reference solution to T63 and its interpolation to the hexagonal–icosahedral grid.) Apart from this, the behavior

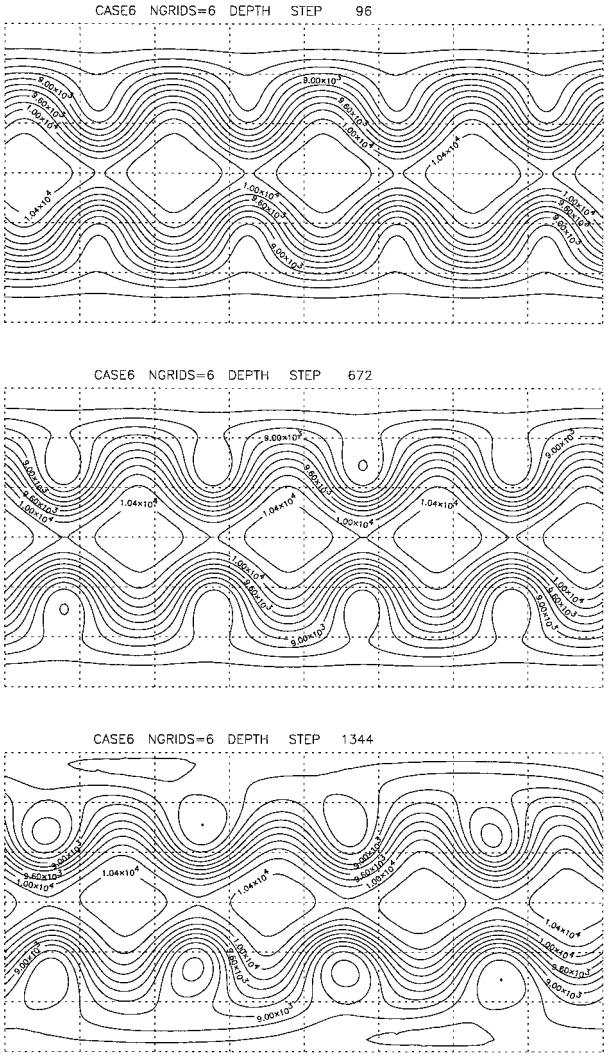

FIG. 9. Longitude–latitude maps of depth for the Rossby–Haurwitz wave test case, contour interval 200 m. Top: day 1. Middle: day 7. Bottom: day 14.

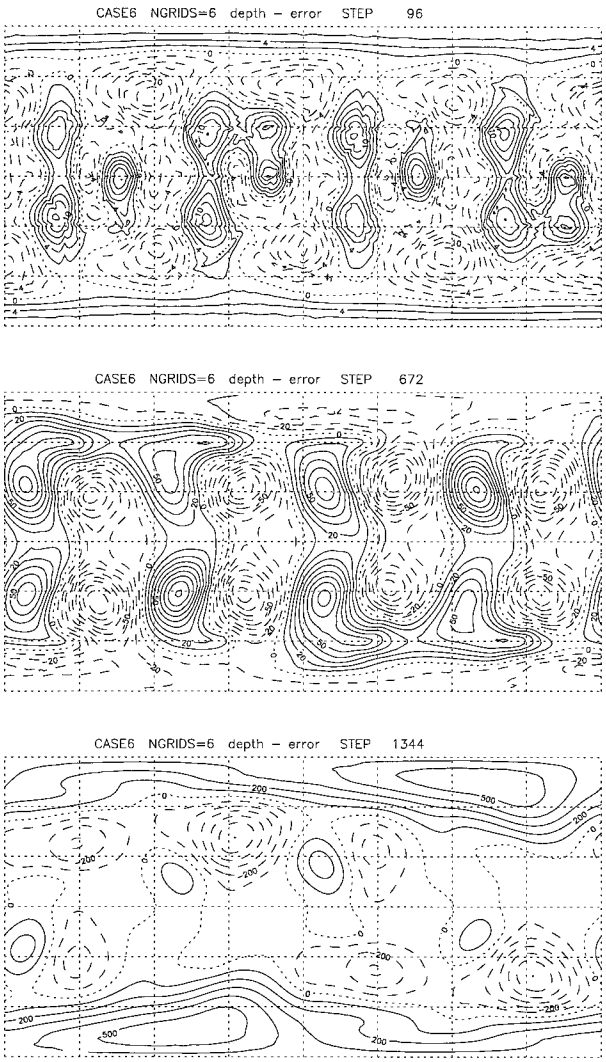

FIG. 10. Longitude–latitude maps of computed depth minus reference solution depth for the Rossby–Haurwitz wave test case. Top: day 1, contour interval 2 m. Middle: day 7, contour interval 10 m. Bottom: day 14, contour interval 100 m.

in the model is very similar to that in the reference solution. In particular, note the cutting off of the cyclonic centers and the vestiges around 60°–70° latitude of PV streamers that initially connected the cyclonic centers to the polar regions. These vestigial streamers indicate a tendency for a cyclonic wrapping up. However, after day 7 the model solution shows an increasing tendency for cyclonic wrapping up, whereas the reference solution shows a shift in the vestigial PV streamer toward the eastern side of the cyclonic center. In view of this and the discussion by Bates et al. (1995), a less noisy reference solution for PV and, ideally, some consensus among reference solutions from different models are clearly required in order to allow a more detailed assessment of model accuracy on this test case.

Repeating the test case using a variance-conserving centered difference scheme for advection of PV instead of the scheme of section 2c gave height errors at day 14 about one-half of the magnitude of those in the original integration. (The symmetry errors were also reduced, but the PV field was noisy.) Also, both the integration with the 900-s time step and the integration with the 300-s time step using the advection scheme of section 2c show positive depth errors at high latitudes and negative depth errors at low latitudes at day 14, signifying a weakening of the mean westerlies. These results suggest that the dissipative nature of the advection scheme of section 2c may be contributing to the error (see, e.g., Smolarkiewicz and Margolin 1993). An estimate for the strength and scale selectivity of the dissipation associated with an advection scheme very

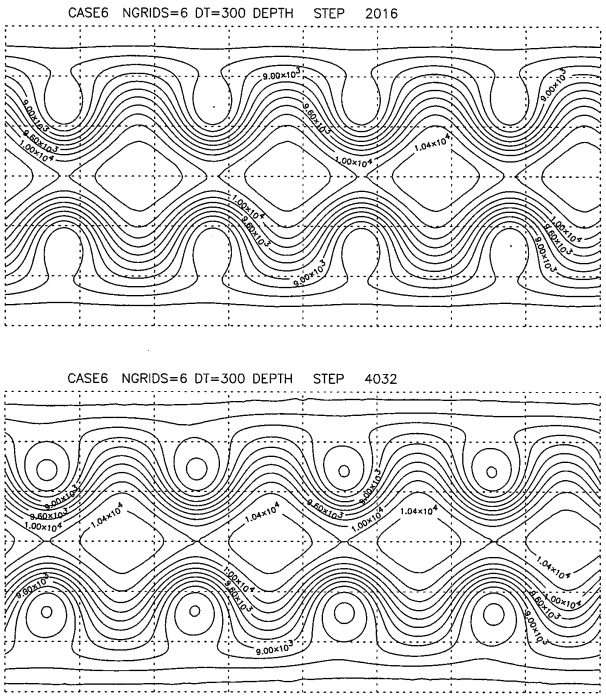

FIG. 11. Longitude–latitude maps of depth for the Rossby–Haurwitz wave test case with 300-s time step, contour interval 200 m. Above: day 7. Below: day 14.

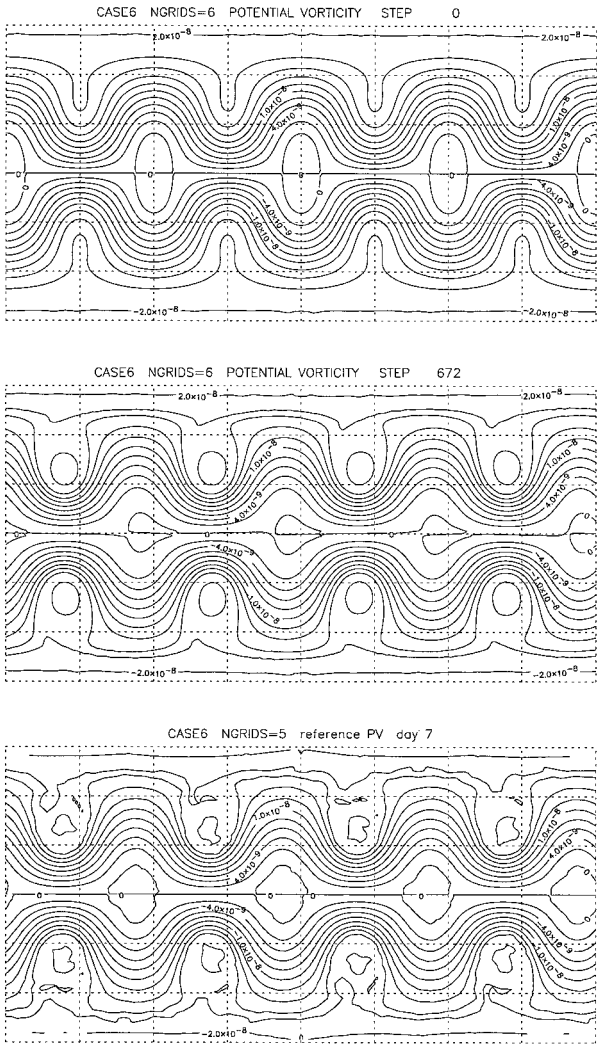

FIG. 12. Longitude–latitude maps of PV for the Rossby–Haurwitz wave test case, contour interval 2×10^{-9} m^{-1} s^{-1}. Top: day 0. Middle: day 7. Bottom: day 7 for the reference solution.

similar to that of section 2c was given by Thuburn (1995) as $(0.01–0.06)\alpha^3|\mathbf{v}|\nabla^4$, where α is the gridbox size. For the Rossby–Haurwitz wave test case on grid 6, $\alpha \approx 240\,000$ m and $|\mathbf{v}|$ peaks at about 100 m s^{-1}, implying dissipation strengths up to about $8 \times 10^{16}\nabla^4$. This is comparable to the scale-selective dissipation typically used in spectral climate models at T42 resolution, whereas the number of degrees of freedom on grid 6 is comparable to a spectral resolution of about T100 (Table 1), suggesting that the dissipation in this case may be stronger than optimal. Certainly this dissipation is far stronger than that in the model used to calculate the reference solution. For example, the dissipation used by Jakob et al. (1993) for the Rossby–Haurwitz wave test at T42 was about 16 times weaker than this, while the dissipation used in calculating the T213 reference solution was a factor of 10^4 weaker. However, to put things in perspective, the dissipation associated with the advection scheme of section 2c is far weaker than that associated with a first-order upwind scheme (Thuburn 1995).

d. Analyzed 500-mb height and wind field initial conditions

The final test case consists of initial conditions based on observed atmospheric 500-hPa heights and winds giving flow conditions typical of the real atmosphere. Of course, the evolution of the shallow-water equations from this state cannot be expected to match the evolution of the real atmosphere. The computed solution is compared with a reference solution obtained using a high-resolution spectral model. Results shown are for a run on grid 6.

The initial state is based on data for 0000 UTC 21 December 1978 (Fig. 13). The depth fields and depth errors for day 1 and day 5 are shown in Figs. 14 and 15. The model captures most of the features of the evolution of the reference solution, including the formation of a blocklike anticyclone near the Greenwich meridian. In particular, it has no problem coping with the strong flow across the pole. However, it does fail to capture a cyclonic feature to the south of the blocklike anticyclone that appears in the reference solution.

The PV field at day 1 (Fig. 16) shows smooth, coherent cyclones, anticyclones, and connecting streamers

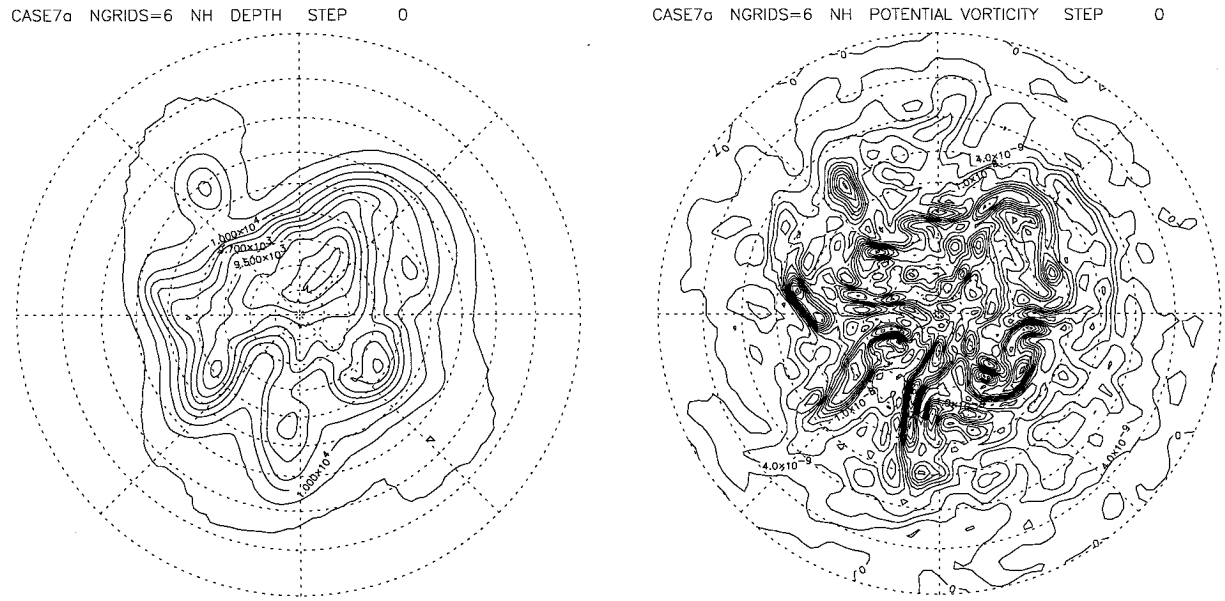

FIG. 13. Northern Hemisphere initial conditions for the test case using analyzed height field initial data. Left: depth, contour interval 100 m. Right: PV, contour interval 2×10^{-9} m^{-1} s^{-1}. The Greenwich meridian is toward the bottom of the figure.

typical of those we expect to see in a turbulent two-dimensional fluid. The day 1 PV is certainly much less noisy than the initial PV (Fig. 13) or the day 1 PV calculated from the reference solution. Comparison of the PV with the depth error at day 1 is quite revealing. The depth error consists of a number of compact regions of positive error. These regions coincide with the centers of cyclonic PV. Apparently the PV in the cyclonic centers is smoothed by the advection scheme, weakening the lows and allowing them to fill in. Subsequently the errors cascade upscale. By day 5 the model PV field is much smoother than at day 1 (Fig. 16). The PV calculated from the reference solution at day 5 also has fewer large-scale strong PV features than at day 1 but is rather noisy. Thus, it is difficult to judge to what extent the model PV at day 5 is excessively smoothed and to what extent the reference solution is insufficiently smoothed. The reference solution does have an intense,

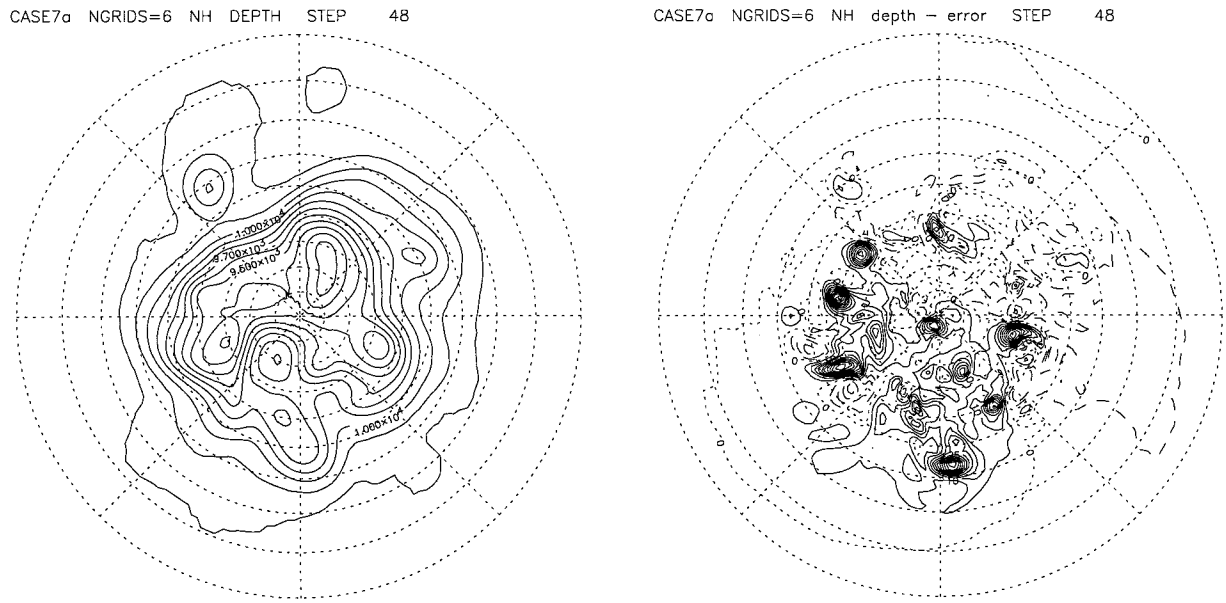

FIG. 14. Northern Hemisphere fields at day 1 for the test case using analyzed height field initial data. Left: depth, contour interval 100 m. Right: computed depth minus reference solution depth, contour interval 5 m.

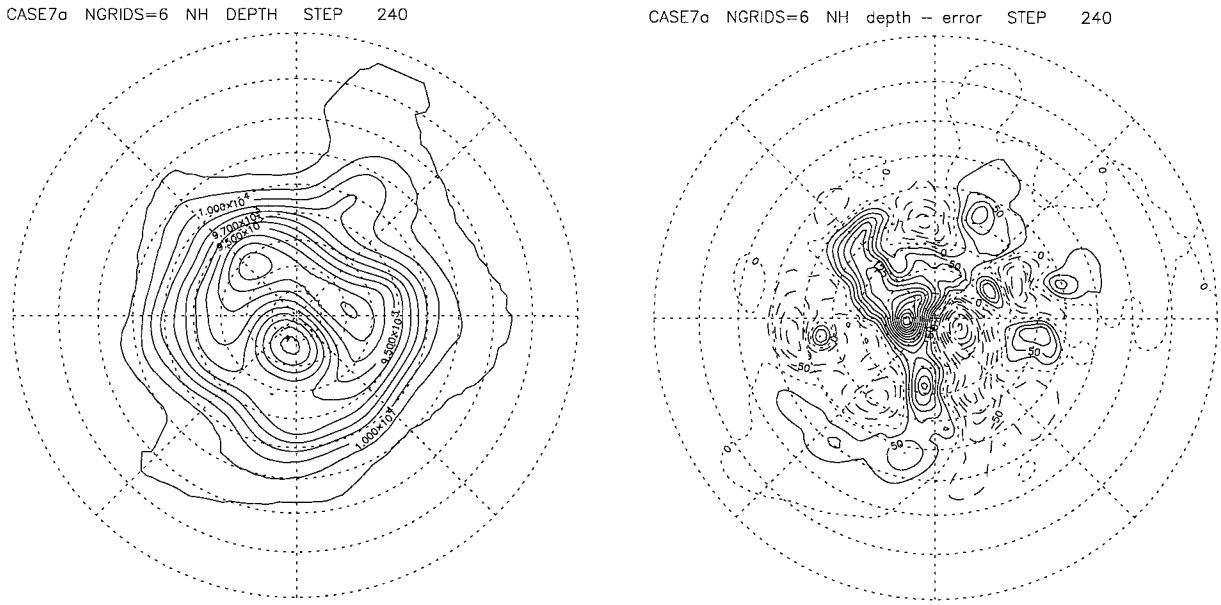

FIG. 15. Northern Hemisphere fields at day 5 for the test case using analyzed height field initial data. Left: depth, contour interval 100 m. Right: computed depth minus reference solution depth, contour interval 25 m.

small-scale cyclonic PV feature at about 60°N, 20°W, which is severely damped in the model and that corresponds to the cyclonic feature that the model failed to capture in the height field mentioned above.

4. Extension to three dimensions

Some basic issues must be addressed in order for the schemes used here for the shallow-water equations to be extended to the three-dimensional primitive equations. One issue concerns the use of PV as a prognostic variable. In most coordinate systems, including terrain-following coordinate systems, recovery of the rotational wind from the PV distribution requires the solution of a three-dimensional elliptic equation, which would be expensive. In isentropic coordinates, on the other hand, only a two-dimensional elliptic equation need be solved at each model level. However, isentropic

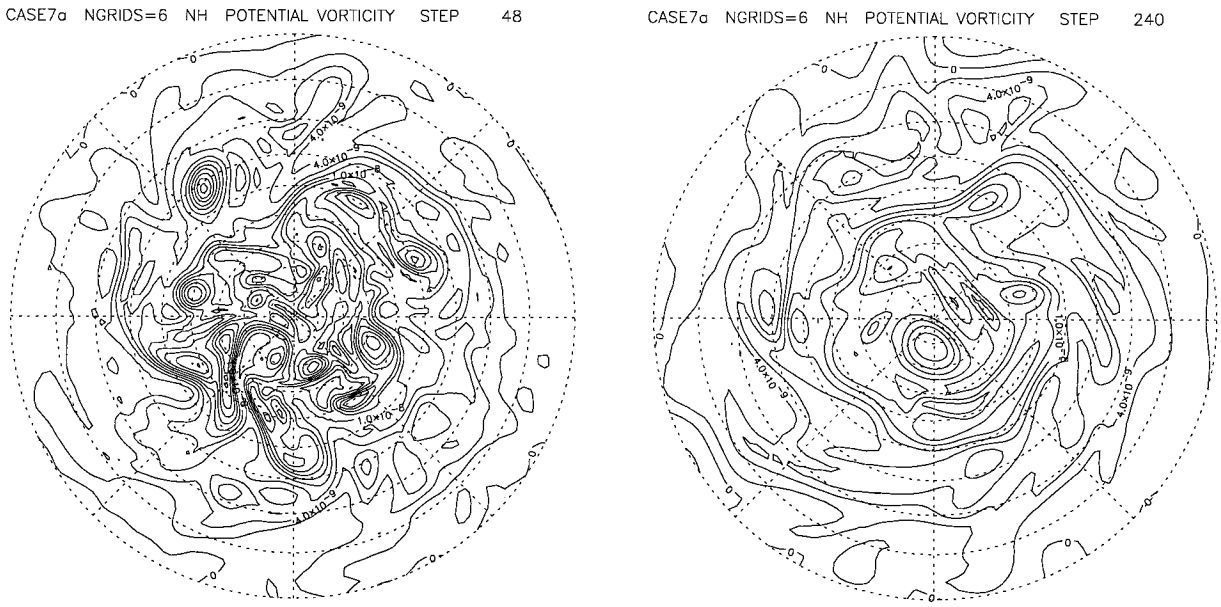

FIG. 16. Northern Hemisphere PV for the test case using analyzed height field initial data, contour interval 2×10^{-9} m^{-1} s^{-1}. Left: day 1. Right: day 5.

coordinates introduce problems in treating a solid lower boundary. (Note that the recovery of wind and temperature fields from PV under some balance assumption using ''PV invertibility'' requires the solution of a three-dimensional elliptic equation in *any* coordinate system, but that is a distinct problem from the one involved in integrating the full primitive equations without any assumption of balance.)

Another issue concerns the use of a semi-implicit time scheme. In three dimensions a semi-implicit time scheme would require the solution of a three-dimensional elliptic equation at each step. Also, a semi-implicit scheme could cause problems in some coordinate systems, for example, in isentropic coordinates, when there is a strong gradient of layer thickness or layer thicknesses close to zero, for example, if massless layers (e.g., Hsu and Arakawa 1990) are used at the lower boundary. In particular, a semi-implicit scheme might undermine the shape-preserving property for advection of mass. One way to sidestep these problems is to use a split-explicit time scheme (e.g., Cullen and Davies 1991), in which several short explicit ''adjustment'' steps are taken for mass and divergence per single long advection step for PV and tracers.

A model that integrates the three-dimensional primitive equations in isentropic coordinates using extensions of the methods described here has been built by A. Gregory (1996, personal communication). The model is being used to study stratospheric dynamics, for which an isentropic lower boundary at $\theta = 414$ K is adequate, thereby avoiding the problems with a solid lower boundary mentioned above. Also, it uses a split-explicit time step, thereby avoiding the potential problems with a semi-implicit time scheme mentioned above.

5. Summary

A shallow-water model has been presented based on a novel combination of numerical methods, including a hexagonal–icosahedral grid, PV as a prognotic variable, and a conservative, shape-preserving scheme for the advection of mass, PV, and tracers. This combination of schemes avoids Gibbs ripples, pole problems, and spurious oscillations in advected quantities and ensures conservation of advected quantities. The overall accuracy of the model has been assessed on the suite of test cases proposed by Williamson et al. (1992). The model showed no stability problems in any of the test cases when reasonable choices were made for the time step and for the reference depth in the semi-implicit scheme. Nor did the model show any tendency to create small-scale noise in either space or time. On the advection test the scheme performs as well as any other Eulerian scheme for which results of this test have been documented (e.g., Heikes 1993; Jakob et al. 1993; Malcolm 1994; Rasch 1994). On the other Williamson et al. test cases the spectral method performs significantly better than any other scheme yet documented (Jakob et al. 1993). (The limitations of the spectral method, such as Gibbs ripples when tendencies from parameterized processes are calculated on a quadratic Gaussian grid, do not show up in these purely dynamical tests.) On these tests the model presented here is similar in accuracy to the other nonspectral models whose results are available (e.g., Heikes 1993; Malcolm 1994).

Two particular sources of error in the model have been identified. The first is the distortion of the hexagonal grid boxes, particularly near the pentagonal grid boxes, giving rise to error maps that reflect the symmetry of the grid in some test cases. The results of section 3c show that the symmetry errors are sensitive to the size of the time step, suggesting that they are dominated not by the calculation of the Laplacian but, rather, by the advection scheme whose space *and* time truncation errors depend on the distortion of the grid boxes. Heikes and Randall (1995b) have proposed an improvement to the way the grid is constructed that minimizes the box-distortion-dependent errors in the calculation of the Laplacian. It remains to be seen whether this improvement, or something similar, can reduce the box-distortion-dependent errors in the advection scheme too.

The second source of error identified in the model is the dissipative nature of the advection scheme. Although the dissipation is far weaker than that of a first-order upwind scheme, the estimate of dissipation given in section 3c, the reduced height errors with variance-conserving advection of PV in the test case of section 3c, and the weakening of the lows in the test case of section 3d are all evidence that dissipation is a source of error. The dissipation associated with upwind biased forward-in-time advection schemes does tend to decrease as the order of accuracy of the schemes increases (e.g., Leonard 1991). However, it is not clear if there is a lower limit (greater than zero) to the strength of dissipation of such a scheme or what that limit might be, particularly if we insist that the scheme be shape preserving. Also, it is a far from trivial exercise to extend the advection scheme used here beyond its current order of accuracy on the hexagonal–icosahedral grid. Whether such dissipation errors are important or not in practice will depend on the proposed application of the model and on the timescales of other dissipative processes, such as radiation (or its shallow-water analog, depth relaxation), included in the model.

Acknowledgments. I am grateful to Piotr Smolarkiewicz and two anonymous reviewers, whose comments led to improvements of this paper, and to Andrew Gregory for many interesting discussions on time schemes and on extensions of the methods presented here. I appreciate the help of John Truesdale in getting hold of the reference solutions for the test cases described in section 3. This work was carried out under the U.K. Universities Global Atmospheric Modelling Programme funded by the U.K. Natural Environment Research Council.

APPENDIX A

Multigrid Elliptic Solver

This appendix summarizes the multigrid method used to solve Poisson's equation and the elliptic equation that arises in taking a semi-implicit time step. The hierarchy of grids for different model resolutions, described in section 2a, is also the natural hierarchy of grids to use in a multigrid solver. Figure A1 illustrates schematically the relationship between grid boxes on one grid and grid boxes on the next coarser grid.

An outline of the algorithm for solving Poisson's equation (12) is as follows.

- The quantity ζ_k on the right-hand side of the equation is projected onto successively coarser grids by area-weighted averaging.
- A solution is calculated by iterated relaxation on the coarsest grid.
- This solution is interpolated to the next finer grid; some of the box centers on the finer grid coincide with box centers on the coarser grid and so they simply take the corresponding value of the solution. The remaining box centers take interpolated values of the solution. A linear interpolation between two neighboring boxes is used.
- A few relaxation sweeps are performed in which only the interpolated box values are relaxed.
- A solution is calculated on this grid by iterated relaxation.
- The process of going to finer grids is repeated until the solution is found at the full model resolution.

The area-weighted averaging of the right-hand side to go from the finest grid to coarser grids is done as follows. Let $Z_k^{(i)}$ be the vorticity in box k multiplied by the area of box k on grid i:

$$Z_k^{(i)} = \zeta_k^{(i)} A_k^{(i)}. \tag{A1}$$

On the finest grid, $i = N$, $Z_k^{(N)}$ is calculated directly using (A1). On the second finest grid, $i = N - 1$, the values of $Z_k^{(N-1)}$ are given by

$$Z_k^{(N-1)} = Z_k^{(N)} + \sum_j Z_j^{(N)}/2, \tag{A2}$$

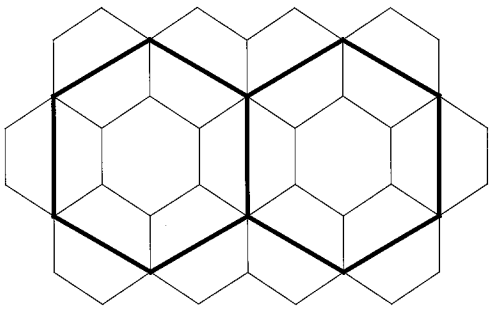

FIG. A1. Illustration of the relationship between grid boxes on one grid (thin lines) and on the next coarser grid (heavy lines).

where j runs over the neighbors of box k on grid N. The grid indexing is chosen so that boxes on grids $i - 1$ and i whose centers coincide have the same index, for example, index k in (A2). Note that each box on grid N either contributes all of its Z to a single box on grid $N - 1$ or contributes one-half of its Z to each of two boxes on grid $N - 1$. Thus, the global integral $\Sigma_k Z_k^{(i)}$ is the same on both grids. Furthermore, the integral over any more local region is approximately the same on both grids. This is important because it means that the implied circulation around the region is approximately the same on both grids.

For other grids in the hierarchy, $Z^{(i)}$ is calculated in terms of $Z^{(i+2)}$:

$$Z_k^{(i)} = \sum_j Z_j^{(i+2)} + \sum_l Z_l^{(i+2)}\Big/2, \tag{A3}$$

where j runs over the boxes on grid $i + 2$ that are contained completely within box k on grid i, while l runs over the boxes on grid $i + 2$ that overlap the edge of box k on grid i. Note that each box on grid $i + 2$ either contributes all of its Z to a single box on grid i or contributes one-half of its Z to each of two boxes on grid i. Thus, the global integral $\Sigma_k Z_k^{(i)}$ is the same on all grids and the integral over any more local region is approximately the same on all grids. This procedure of going directly from grid $i + 2$ to grid i introduces less smoothing of the right-hand side of Poisson's equation than going from grid $i + 2$ to grid $i + 1$ and then to grid i using a formula like (A2), and it significantly improves the performance of the multigrid method.

Rearranging (12) gives

$$\frac{\sum_i (e_i/d_i)\psi_i - Z_k}{\sum_i e_i/d_i} - \psi_k = 0. \tag{A4}$$

After n iterations this equation will not be satisfied exactly, but there will be some residual, r_k^n, say:

$$\frac{\sum_i (e_i/d_i)\psi_i^n - Z_k}{\sum_i e_i/d_i} - \psi_k^n = r_k^n. \tag{A5}$$

A simple Jacobi iteration would define ψ_k^{n+1} so that

$$\frac{\sum_i (e_i/d_i)\psi_i^n - Z_k}{\sum_i e_i/d_i} - \psi_k^{n+1} = 0, \tag{A6}$$

that is,

$$\psi_k^{n+1} = \psi_k^n + r_k^n. \tag{A7}$$

In practice convergence can be accelerated by "over-relaxation,"

$$\psi_k^{n+1} = \psi_k^n + \gamma r_k^n, \tag{A8}$$

where γ is some number slightly greater than 1. Here $\gamma = 1.2$ is used for all the tests described in this paper.

TABLE A1. Number of iterations taken by the multigrid solver on each grid.

Grid number	1	2	3	4	5	6
Number of iterations	30	30	30	25	20	15

A Gauss–Seidel iteration, in which updated values ψ_i^{n+1} are used in the calculation of r_k^n as soon as they become available, formally converges faster. However, the resulting code does not ''vectorize'' on machines with a pipeline architecture, whereas the code for the Jacobi iteration does. So, in practice, the Jacobi iteration is more efficient.

The interpolation of a partial solution to a finer grid is straightforward. If the center of a grid box on grid $i + 1$ coincides with the center of a grid box on grid i, then the solution is simply copied into the box on the finer grid. Each other grid box on grid $i + 1$ lies between two grid boxes on grid i, so the value in the box on grid $i + 1$ is set to the mean of the values in the two boxes on grid i. This linear interpolation from grid i to grid $i + 1$ causes a weak damping of features in the solution close to the grid scale of grid i, and extra iterations on grid $i + 1$ are required to restore them to their full amplitude. An efficient way to do this is to perform a few sweeps in which only the interpolated values (not the copied values) are relaxed. In all the tests presented here three such partial sweeps are performed on each grid. [An alternative approach is to use a quadratic interpolation (Heikes and Randall 1995a).]

The algorithm for solving (35) is more complicated because the second term on the right-hand side is also at time step $m + 1$. Discretizing the left-hand side of (35) and rearranging leads to an equation for the residual after n iterations:

$$\frac{\nu^2 \sum_i (e_i/d_i)(\delta_i^{m+1})^n - B_k^n}{\nu^2 \sum_i (e_i/d_i) + A_k} - (\delta_k^{m+1})^n = r_k^n, \quad \text{(A9)}$$

where B_k^n equals A_k times an approximation after n iterations to the right-hand side of (35) in box k. The estimate of δ_k^{m+1} after $n + 1$ iterations $(\delta_k^{m+1})^{n+1}$ is calculated by overrelaxation using an equation like (A8). At the same time $(\chi^{m+1})^{n+1}$ is calculated from $(\delta^{m+1})^{n+1}$ using equations analogous to (A4)–(A8). This enables a new approximation $(\mathbf{v}_{\text{div}}^{m+1})^{n+1}$ to be calculated, and hence B_k^{n+1}. Thus, the two elliptic problems for δ^{m+1} and χ^{m+1} are solved together. The algorithm is as follows.

- The quantity R in (35) is projected onto successively coarser grids by area-weighted averaging. The intermediate value of the absolute vorticity, $\zeta_a^{(+)}$, is also projected onto successively coarser grids.
- Solutions for both δ^{m+1} and χ^{m+1} are calculated by iterated relaxation on the coarsest grid. Term $\mathbf{v}_{\text{div}}^{m+1}$ is regularly recalculated from the latest estimate of χ^{m+1}, allowing an improved estimate of the right-hand side of (35) to be calculated. However, this calculation is expensive and so it is only carried out every N_{skip} iterations. Here $N_{\text{skip}} = 4$ for all the test results presented.
- These solutions are interpolated to the next finer grid; some of the box centers on the finer grid coincide with box centers on the coarser grid and therefore take the corresponding value of the solution. The remaining box centers take interpolated values of the solution.
- A few relaxation sweeps are performed in which only the interpolated values are relaxed.
- Solutions are calculated on this grid by iterated relaxation.
- The process of going to finer grids is repeated until the solutions are found at the full model resolution.

The number of iterations taken on each grid in the hierarchy is given in Table A1. Tests in which the number of iterations on each grid was doubled showed no discernible difference in the results.

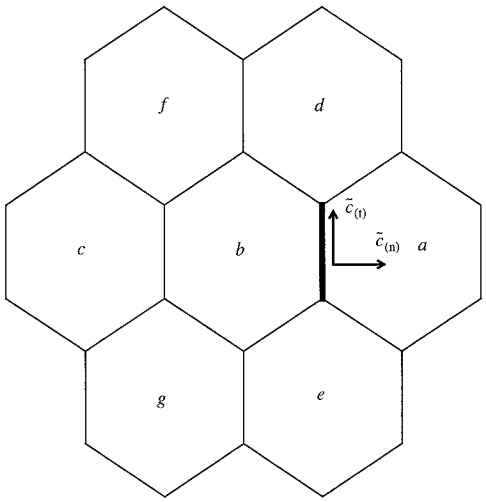

FIG. B1. Schematic showing the values used to calculate the advective flux at an outflow edge of a hexagonal box, indicated by the heavy line.

APPENDIX B

Preliminary Box Edge Values for the Advection Scheme

This appendix gives the formulas used to compute the preliminary box edge value of PV, $\hat{Q}^{(p)}$, from the PV values in nearby grid boxes, *a*, *b*, *c*, etc. When the box immediately upstream of the edge in question is a hexagon, the interpolation formula used is (see Fig. B1):

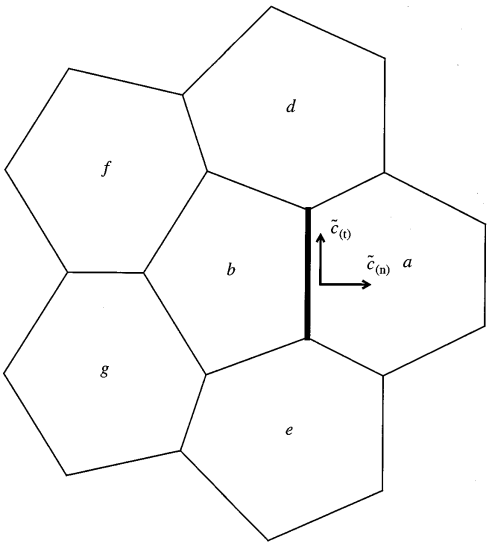

FIG. B2. Schematic showing the values used to calculate the advective flux at an outflow edge of a pentagonal box, indicated by the heavy line.

$$\hat{Q}^{(p)} = \frac{1}{8}(3a + 6b - c) - \frac{5}{108}(c + d + e - 3b) - \frac{3}{4}\tilde{c}_{(n)}(a - b) + \frac{3}{8}\tilde{c}^2_{(n)}(a - 2b + c) + \frac{\sqrt{3}}{4}\tilde{c}_{(t)}(e - d) + \frac{1}{8}\tilde{c}^2_{(t)}[4(e + d) - (3a + 6b - c)] - \frac{\sqrt{3}}{4}\tilde{c}_{(n)}\tilde{c}_{(t)}[(e - d) - (g - f)]. \quad \text{(B1)}$$

Here $\tilde{c}_{(t)}$ is a tangential Courant number defined in an analogous way to the normal Courant number $\tilde{c}_{(n)}$, except that it may be positive or negative (positive means to the left of the normal component of the flow).

When the box immediately upstream of the edge in question is a pentagon, the following formula is used (see Fig. 19):

$$\hat{Q}^{(p)} = T_1 - \frac{5}{8}T_2\tilde{c}_{(n)} + \frac{25}{48}T_3\tilde{c}^2_{(n)} - \frac{5}{8}T_4\tilde{c}_{(t)} + \frac{25}{48}T_5\tilde{c}^2_{(t)} + \frac{25}{48}T_6\tilde{c}_{(n)}\tilde{c}_{(t)}, \quad \text{(B2)}$$

where

$$T_1 = \frac{1}{2}(a + b) - \frac{1}{4}T_3, \quad \text{(B3)}$$

$$T_2 = a - b, \quad \text{(B4)}$$

$$T_3 = K_1\left[\frac{1}{2}(d + e) - b - (a - b)c_2\right] + K_2\left[\frac{1}{2}(f + g) - b + (a - b)c_1\right], \quad \text{(B5)}$$

$$T_4 = \frac{1}{2}K_5(d - e) + \frac{1}{2}K_6(f - g), \quad \text{(B6)}$$

$$T_5 = K_3\left[\frac{1}{2}(d + e) - b - (a - b)c_2\right] + K_4\left[\frac{1}{2}(f + g) - b + (a - b)c_1\right], \quad \text{(B7)}$$

$$T_6 = \frac{1}{2}K_7(d - e) + \frac{1}{2}K_8(f - g), \quad \text{(B8)}$$

and the coefficients c_1, c_2, and K_i ($i = 1, \ldots, 8$) are given by

$$s_1 = \sin\left(\frac{\pi}{5}\right), \quad \text{(B9)}$$

$$s_2 = \sin\left(\frac{2\pi}{5}\right), \quad \text{(B10)}$$

$$c_1 = \cos\left(\frac{\pi}{5}\right), \quad \text{(B11)}$$

$$c_2 = \cos\left(\frac{2\pi}{5}\right), \quad \text{(B12)}$$

$$w_1 = s_1^2(c_2^2 - c_2), \quad \text{(B13)}$$

$$w_2 = s_2^2(c_1^2 + c_1), \quad \text{(B14)}$$

$$K_1 = \frac{s_1^2}{w_1 - w_2}, \quad \text{(B15)}$$

$$K_2 = -\frac{s_2^2}{w_1 - w_2}, \quad \text{(B16)}$$

$$K_3 = -\frac{(c_1^2 + c_1)}{w_1 - w_2}, \quad \text{(B17)}$$

$$K_4 = \frac{c_2^2 - c_2}{w_1 - w_2}, \quad \text{(B18)}$$

$$K_5 = \frac{c_1 + 1/2}{s_2(c_1 + c_2)}, \quad \text{(B19)}$$

$$K_6 = \frac{c_2 - 1/2}{s_1(c_1 + c_2)}, \quad \text{(B20)}$$

$$K_7 = \frac{1}{s_2(c_1 + c_2)}, \quad \text{(B21)}$$

$$K_8 = -\frac{1}{s_1(c_1 + c_2)}. \quad \text{(B22)}$$

[Note that the convention used here for the sign of the tangential Courant number is the opposite of that used in Thuburn (1996b). Also note that in that report the forth and sixth terms on the right-hand side of the equa-

tion corresponding to (B2) here have sign errors, although the correct model code was used in all the test cases presented there.]

REFERENCES

Augenbaum, J. M., and C. S. Peskin, 1985: On the construction of the Voronoi mesh on a sphere. *J. Comput. Phys.,* **14,** 177–192.

Bates, J. R., Y. Li, A. Brandt, S. F. McCormick, and J. Ruge, 1995: A global shallow-water numerical model based on the semi-Lagrangian advection of potential vorticity. *Quart. J. Roy. Meteor. Soc.,* **121,** 1981–2005.

Cullen, M. J. P., 1974: Integrations of the primitive equations on a sphere using the finite element method. *Quart. J. Roy. Meteor. Soc.,* **100,** 555–562.

——, and T. Davies, 1991: A conservative split-explicit integration scheme with fourth-order horizontal advection. *Quart. J. Roy. Meteor. Soc.,* **117,** 993–1002.

Haynes, P. H., and M. E. McIntyre, 1990: On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.,* **47,** 2021–2031.

Heikes, R., 1993: The shallow water equations on a spherical geodesic grid. Colorado State University, Department of Atmospheric Science Paper 524, 175 pp. [Available from Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80521.]

——, 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.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

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

Jakob, R., J. J. Hack, and D. L. Williamson, 1993: Solutions to the shallow water test set using the spectral transform method. NCAR Tech. Note NCAR/TN-388+STR, 82 pp. [Available from National Center for Atmospheric 0Research, Boulder, CO 80303.]

Leonard, B. P., 1991: The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Eng.,* **88,** 17–74.

——, M. K. MacVean, and A. P. Lock, 1993: Positivity preserving schemes for multidimensional advection. NASA Tech. Memo. 106055 ICOMP-93-05, Institute for Computational Mechanics in Propulsion, Lewis Research Center, Cleveland, OH, 62 pp.

——, A. P. Lock, and M. K. MacVean, 1996: Conservative unrestricted-time-step multidimensional constancy-preserving advection schemes. *Mon. Wea. Rev.,* **124,** 2588–2606.

Leslie, L. M., and R. J. Purser, 1995: Three-dimensional mass-conserving semi-Lagrangian scheme employing forward trajectories. *Mon. Wea. Rev.,* **123,** 2551–2556.

Lin, S.-J., and R. B. Rood, 1996: Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.,* **124,** 2046–2070.

——, W. C. Chao, Y. C. Sud, and G. K. Walker, 1994: A class of van Leer–type transport schemes and its application to moisture transport in a general circulation model. *Mon. Wea. Rev.,* **122,** 1575–1593.

Malcolm, A. J., 1994: Evaluation of a numerical scheme on a test set of problems for the shallow water equations. U.K. Meteorological Office, Forecasting Research Division Tech. Rep. 122, 39 pp. [Available from U.K. Meteorological Office, London Road, Bracknell, RG12 252, United Kingdom.]

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. *Proc. WMO/IUGG Symp. on Short- and Medium-Range Numerical Weather Prediction,* Tokyo, Japan, Japan Meteorological Society, 317–326.

McIntyre, M. E., 1992: Atmospheric dynamics: Some fundamentals with observational implications. *Proc. Int. School Physics ''Enrico Fermi,''* 313–386.

Rančić, M., J. R. Purser, and F. Mesinger, 1996: A global shallow water model using an expanded spherical cube: Gnomonic versus conformal coordinates. *Quart. J. Roy. Meteor. Soc.,* **122,** 959–982.

Rasch, P. J., 1994: Conservative shape-preserving two-dimensional transport on a spherical reduced grid. *Mon. Wea. Rev.,* **122,** 1337–1350.

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

Smolarkiewicz, P. K., and P. J. Rasch, 1991: Monotone advection on the sphere: An Eulerian versus semi-Lagrangian approach. *J. Atmos. Sci.,* **48,** 793–810.

——, and L. G. Margolin, 1993: On forward-in-time differencing for fluids: Extension to a curvilinear framework. *Mon. Wea. Rev.,* **121,** 1847–1859.

Staniforth, A., and J. Côté, 1991: Semi-Lagrangian integration schemes for atmospheric models—A review. *Mon. Wea. Rev.,* **119,** 2206–2223.

Temperton, C., and A. Staniforth, 1987: An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. *Quart. J. Roy. Meteor. Soc.,* **113,** 1025–1039.

Thuburn, J., 1995: Dissipation and cascades to small scales in numerical models using a shape-preserving advection scheme. *Mon. Wea. Rev.,* **123,** 1888–1903.

——, 1996a: Multidimensional flux-limited advection schemes. *J. Comput. Phys.,* **123,** 74–83.

——, 1996b: A PV-based shallow water model on a hexagonal-icosahedral grid. UGAMP Tech. Rep. 40, 93 pp. [Available from Centre for Global Atmospheric Modelling, Department of Meteorology, University of Reading, Reading RG6 6AH, United Kingdom.]

Walsteijn, F. H., 1994: Robust methods for 2D turbulence. *J. Comput. Phys.,* **114,** 129–145.

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

——, and P. J. Rasch, 1989: Two-dimensional semi-Lagrangian transport with shape-preserving interpolation. *Mon. Wea. Rev.,* **117,** 102–129.

——, J. B. Drake, J. J. Hack, R. Jakob, and P. N. Swarztrauber, 1992: A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.,* **102,** 211–224.