FIM Documentation

Jian-Wen Bao (Column physics parameterization)
Stan Benjamin (Isentropic-sigma hybrid coordinate, team coordinator)
Rainer Bleck (Sigma-isentropic hybrid coordinate)
John Brown (Physical parameterizations, model validation)
Jin Lee (Icosahedral SWM developer, team coordinator)
Alexander MacDonald (Icosahedral SWM developer, code design)
Jacques Middlecoff (Massively parallel processing computer specialist)
Ning Wang (Icosahedral grid generation, graphics)

March 15, 2009

1. Introduction

FIM is a new global weather prediction model currently under development in the Global Systems Division of NOAA/ESRL. The acronym FIM indicates that the model uses a flow-following (i.e. quasi-Lagrangian) vertical coordinate, finite-volume numerics, and an icosahedral global grid.

Global models fall into two classes, spectral and grid-point, depending on the spatial representation of model variables on the sphere. Spectral models have gained almost universal acceptance in the last several decades. However, drawbacks of such models in terms of operations count and communications overhead have led, in recent years, to the development of new types of grid-point global models discretized on geodesic grids (Tomita et al. 2001). Among the various geodesic grids, the icosahedral grid stands out in providing near-uniform coverage over the globe while allowing recursive refinement of grid spacing. If configured as a grid consisting of a large number of hexagonal cells with 12 embedded pentagons, the icosahedral grid lends itself particularly well to the finite-volume approach in which conventional finite-difference operators are replaced by numerically approximated line integrals along grid cell boundaries.

Williamson (1968) and Sadourny et al. (1968) were the first to solve the shallow-water equations on icosahedral grids using finite-difference formulations. More recently, Colorado State University modelers (Heikes and Randall 1995; Ringler et al. 2000) developed an icosahedral-hexagonal shallow-water model (SWM) based on finite-volume numerics. The German Weather Service is currently using an icosahedral-hexagonal model for operational global weather prediction (Majewski et al. 2002). A Japanese group (Tomita et al. 2004) has developed a nonhydrostatic general circulation model (GCM) formulated on an icosahedral-hexagonal grid and has carried out high-resolution cloud-resolving GCM simulations.

Recognizing the potential of finite-volume icosahedral models in high-resolution global weather and climate prediction, MacDonald, then director of NOAA's Forecast Systems Laboratory, teamed up with Lee to develop a finite-volume icosahedral SWM which they sucessfully subjected to the suite of SWM tests of Williamson et al. (1992). In light of the positive outcome of this study, MacDonald in mid-2005 formed a group of modelers to develop an operational 3-D icosahedral global model, initially for numerical weather prediction (NWP) but with an eye on climate simulation. The group was specifically charged with adding a third dimension to the SWM in the form of an Arbitrary Lagrangian-Eulerian (ALE) or "flow-following" vertical coordinate, and to set the stage for its use in operational global NWP at 15 km horizontal resolution.

The ALE vertical coordinate, the second novel

feature in FIM aside from the icosahedral grid, is based upon a compound or "hybrid" grid consisting of isentropic layers in the free atmosphere and terrain-following or σ coordinate layers near the ground. It is an improved version of a scheme used successfully in atmospheric and ocean models such as RUC [Rapid Update Cycle: Bleck and Benjamin (1993), Benjamin et al. (2004)] and HYCOM [HYbrid Coordinate Ocean Model: Bleck (2002)]. Use of a vertical coordinate with Lagrangian attributes is meant to reduce nonphysical dispersion during 3-D transport (as well as during lateral mixing) of tracers such as moisture, potential vorticity, and chemical compounds. Physical parameterizations in FIM match those used operationally by the Global Forecast System (GFS) at the National Centers for Environmental Prediction (NCEP). Validation of FIM, a "draft" hydrostatic version of which was completed by the middle of 2008, will be based on extended-range real-data forecasts using GFS initial conditions.

In preparation for converting the SWM to a "stacked SWM" featuring a hybrid vertical coordinate, a set of 2-D numerical experiments including a single-layer ocean with an emerging seamount (turning ocean bottom into dry land) were carried out. Once the 2-D model's capability to accomodate zero-thickness coordinate layers was confirmed, the team proceded to construct and test a purely isentropic multi-layer coordinate model along the lines of Bleck (1984a), leaving implementation of the isentropic-sigma hybrid coordinate until later. The global circulation in the isentropic model, maintained against dissipation by Newtonian relaxation toward a baroclinically unstable basic state [basically, Held and Suarez (1994) forcing translated to an isentropic framework], remained robust for months of simulation time. Hybridization of the vertical coordinate and implementation of GFS model physics were the final steps in model development.

2. Basic Equations

Denote the spatial coordinate system by (x, y, s) where x, y are the common horizontal coordinates and s is an arbitrary but monotonic function of height, subject only to the requirement that bottom and top of the model atmosphere are s surfaces. The physical dimensions of s are arbitrary; in fact, s can be chosen to be a continuous extension of the vertical layer index.

Let $\mathbf{v}$ be the (Cartesian) horizontal velocity vector; ∇_s the 2-D gradient operator at $s = const$; $\Pi = c_p(p/p_0)^{R/c_p}$ the Exner function; $\theta = c_pT/\Pi$ the potential temperature, $M = gz + \Pi\theta$ the Montgomery potential; ζ the vorticity (i.e., the vertical or $\mathbf{k}$ component of the velocity curl vector); $\dot{\theta}$ the net diabatic heating; and $\mathbf{F}$ the sum of frictional forces. The set of dynamic equations solved in FIM can then be formulated as follows [see Kasahara (1974) and Bleck (1978a) for detailed derivations]

Mass conservation:

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial s}\right) + \nabla_s \cdot \left(\mathbf{v}\frac{\partial p}{\partial s}\right) + \frac{\partial}{\partial s}\left(\dot{s}\frac{\partial p}{\partial s}\right) = 0 \qquad (1)$$

Thermal energy conservation:

$$\frac{\partial}{\partial t}\left(\theta\frac{\partial p}{\partial s}\right) + \nabla_s \cdot \left(\mathbf{v}\frac{\partial p}{\partial s}\theta\right) + \frac{\partial}{\partial s}\left(\dot{s}\frac{\partial p}{\partial s}\theta\right) = \dot{\theta}\frac{\partial p}{\partial s} \qquad (2)$$

Momentum conservation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\zeta + f)\mathbf{k} \times \mathbf{v} + \left(\dot{s}\frac{\partial p}{\partial s}\right)\frac{\partial \mathbf{v}}{\partial p} + \nabla_s\left(M + \frac{\mathbf{v}^2}{2}\right) - \Pi\nabla_s\theta = \mathbf{F} \qquad (3)$$

Hydrostatic Equation:

$$\frac{\partial M}{\partial \theta} = \Pi. \qquad (4)$$

Conservation equations for moisture-related variables (gaseous, liquid, frozen water) and chemical compounds such as ozone have the same form as (2). Source and sink terms in these equations – the analogs of $\dot{\theta}\partial p/\partial s$ – are governed by additional equations expressing what is commonly called "model physics".

Lateral mixing terms are presently omitted from the above equations. The dissipative properties of the horizontal numerical discretization scheme (see below) appear to be sufficient to counteract accumulation of energy on the grid scale. However, the effect of turbulent vertical mixing on mass field constituents is included in the right-hand-side of (2), and there

is a commensurate amount of turbulent vertical momentum mixing symbolically represented in (3) by $\mathbf{F}$.

Discretization in the vertical is accomplished in FIM by integrating prognostic variables, as well as their governing equations, over individual layers bounded by s surfaces. FIM actually goes one step further: following the shallow-water paradigm, it views the atmospheric state as one in which most variables are piecewise constant in the vertical with discontinuities across s surfaces.

Introducing the stairstep discretization of θ in the hydrostatic equation (4) implies that M is vertically constant in each coordinate layer. Since $\nabla_s\theta$ is vertically constant by definition as well, the pressure force terms in (3), if formulated with a layer average of Π, do not create vertical shear within a coordinate layer. Thus, there is no need to distinguish $\mathbf{v}$ from its layer average $\overline{\mathbf{v}}$. The layer-averaged momentum equation therefore can be written as

Momentum conservation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\zeta+f)\mathbf{k}\times\mathbf{v} + \frac{1}{\Delta p}\left[\left(\dot{s}\frac{\partial p}{\partial s}\right)_2(\hat{\mathbf{v}}_2-\mathbf{v}) - \left(\dot{s}\frac{\partial p}{\partial s}\right)_1(\hat{\mathbf{v}}_1-\mathbf{v})\right] + \nabla_s\left(M+\frac{\mathbf{v}^2}{2}\right) - \overline{\Pi}\nabla_s\theta = \mathbf{F} \quad (5)$$

Here, indices 1 and 2 denote the upper and lower interface, respectively, of the coordinate layer in question, and $\Delta p = p_2 - p_1$. The vertical advection terms (those involving $\dot{s}$) are arrived at by integrating

$$\frac{1}{\Delta p}\int\left(\dot{s}\frac{\partial p}{\partial s}\right)\frac{\partial \mathbf{v}}{\partial p}dp$$

by parts and involve interface velocity values denoted here by $\hat{\mathbf{v}}$. Since $\mathbf{v}$ is discontinuous at interfaces, the definition of $\hat{\mathbf{v}}$ is to some extent arbitrary and in practice depends on the finite-difference vertical advection scheme.

The layer-integrated mass and thermal energy equations (1), (2) assume the form

Mass conservation:

$$\frac{\partial \Delta p}{\partial t} + \nabla_s\cdot(\mathbf{v}\Delta p) + \left(\dot{s}\frac{\partial p}{\partial s}\right)_2 - \left(\dot{s}\frac{\partial p}{\partial s}\right)_1 = 0 \quad (6)$$

Thermal energy conservation:

$$\frac{\partial(\theta\Delta p)}{\partial t} + \nabla_s\cdot(\theta\mathbf{v}\Delta p) + \left(\dot{s}\frac{\partial p}{\partial s}\hat{\theta}\right)_2 - \left(\dot{s}\frac{\partial p}{\partial s}\hat{\theta}\right)_1 = \overline{\dot{\theta}}\Delta p \quad (7)$$

As before, the caret denotes interface values needed in the vertical transport terms. The method by which they are constructed (upstream, centered,...) determines properties such as monotonicity and diffusiveness of vertical advection in the model.

Equations for other mass field tracers (moisture etc.) have the same form as (7).

Regardless of whether θ represents potential temperature or another mass field tracer, it is generally necessary to retrieve its value (a concentration or mixing ratio) from the product $\theta\Delta p$ representing the "amount" of θ per unit area. This operation becomes ill-conditioned if Δp at the end of the time step is much smaller (by at least one order of magnitude) than at the beginning. In FIM we avoid generating spurious values when dividing Δp into $\theta\Delta p$ by requiring that the new θ value remain within the range spanned by the old θ values at the point in question and its neighbors. Unfortunately, this introduces an element of nonconservation into the transport process for which a number of "engineering" remedies are available, such as distributing the θ amount gained or lost among neighboring grid cells. Since this problem arises only in nearly massless cells, the θ amount in question tends to be small, and redistributing it is generally not worth the effort.

3. The Dynamic Core

3a. *Time differencing*

Given right-hand side values for the generic differential equation $u_t = F(u,x,t)$ at 3 consecutive time levels n–$2, n$–$1, n$, and u at time level n, the 3^{rd} order Adams-Bashforth scheme (Durran 1991) expresses u at time level n+1 by

$$u^{n+1} = u^n + \frac{\Delta t}{12}[23F^n - 16F^{n-1} + 5F^{n-2}] \quad (8)$$

where Δt is the model time step. The Adams-Bashforth scheme is an explicit scheme and requires only one evaluation of F per time step. Previous studies (Lee and MacDonald 2000; MacDonald et al. 2000) have shown that it is an accurate and efficient scheme for numerical weather prediction models.

3b. *The icosahdedral horizontal mesh*

The governing equations in FIM are formulated in a local stereographic projection (Lee et al. 1995) of an icosahedral grid onto a plane surface. The grid is generated from an icosahedron which has 12 vertices and 20 equilateral spherical triangles with 30 shared edges. Each edge is a segment of a great circle on the sphere. The icosahedral grid provides quasi-uniform coverage of the sphere and allows hierarchical refinement of grid spacing by progressively subdividing triangles into smaller ones. The triangular cells are used to construct hexa- or pentagonally shaped Voronoi cells (Heikes and Randall 1995), the pentagons occupying the 12 vertices of the original icosahedron. After G consecutive bisections of triangle edges, the total number of grid cells is $n = 10 \times (2^G)^2 + 2$.

An icosahedron can be constructed from evenly-sized and -shaped hexagons (though some end up creased and amputated), but these attributes are lost when the icosahedron is transformed into a sphere. There are tradeoffs between minimizing geometric distortion and minimizing variations in cell size during the transformation. Tests indicate that numerical accuracy benefits from minimizing distortion, achievable by radially projecting the original icosahedron onto the sphere. However, in the interest of uniformity in mesh size, some distortion of the grid cells is tolerated in FIM.

Many of our early experiments with FIM have been carried out on a grid obtained by five consecutive subdivisions ($G = 5$), resulting in 10,242 grid cells spaced approximately 240 km apart. Nine subdivisions, the highest number we have explored so far, yield approximately 2.6 million cells with an approximate mesh size of 15 km.

3c. *Horizontal Finite-Volume Operators*

The governing equations are solved with the finite-volume approach (van Leer 1977; Lin et al. 1994), which defines model variables as mean quantities over each grid cell (or control volume). The finite-volume approach excels among the plethora of grid-point discretization schemes because of its integral conservation properties (Lin et al. 1997). Even on irregular grids, conservation laws for various quantities are easily built into the finite-difference equations by approximating terms like vorticity, divergence, and gradient by line integrals along the perimeter of each grid cell:

$$\zeta = A^{-1} \oint \mathbf{v} \cdot d\mathbf{s}; \qquad \nabla \cdot \mathbf{v} = A^{-1} \oint \mathbf{v} \cdot d\mathbf{n};$$

$$\frac{\partial p}{\partial x} = A^{-1} \oint p\,\mathbf{j} \cdot d\mathbf{s}; \qquad \frac{\partial p}{\partial y} = -A^{-1} \oint p\,\mathbf{i} \cdot d\mathbf{s}.$$

Here, $\mathbf{s}$, $\mathbf{n}$ are, respectively, the unit vectors along, and normal to, the line circumscribing the grid cell of area A while $\mathbf{i}$, $\mathbf{j}$ are the unit vector in x, y direction, respectively. The hexagonally and pentagonally shaped elements of the icosahedral mesh are well-suited for approximating such line integrals. The numerical properties of the finite-volume approach on this grid are such that the prognostic equations can be solved in a stable manner for extended periods of time without lateral mixing terms.

3d. *Flux Corrected Transport (FCT)*

The term *Layer Model* describes a class of models in which the vertical spacing of layer interfaces is variable in space and time, with interface movement controlled primarily by the convergence and divergence of lateral mass fluxes in each layer. FIM belongs to this class of vertically quasi-Lagrangian models. To assure the numerical integrity of a layer model, mass fluxes must be constructed with strong emphasis on positive-definiteness and monotonicity. The scheme chosen for this purpose in FIM is known as Flux-Corrected Transport (Boris and Book 1973; Zalesak 1979). The multi-step operation is numerically complex and results in a fairly large horizontal stencil, as illustrated in Fig. 1 for the case of an icosahedral grid and space-centered flux expressions of 2^{nd} order accuracy.

The task at hand is to solve the continuity equation in the hexagonal cell marked by an X on the left. This requires evaluation of six mass fluxes, one for each edge of the hexagon. The stencil for evaluating individual fluxes is shown on the right, with an X marking the edge segment which the computed flux value is assigned to. Note that information from at most four grid cells is needed to compute "raw" flux values of 2^{nd} order accuracy. It is the flux limiting

process that is responsible for extending the stencil because flux limiting requires knowledge of all six fluxes into and out of the two cells abutting a given edge segment.

Overlaying the stencil on the right, appropriately rotated, over the six edges of hexagon X yields the stencil shown on the left. Note that Fig. 1 illustrates the overall distance by which information spreads laterally during a single time step. In practice, the solution procedure is broken into several steps, each one requiring a much smaller stencil.

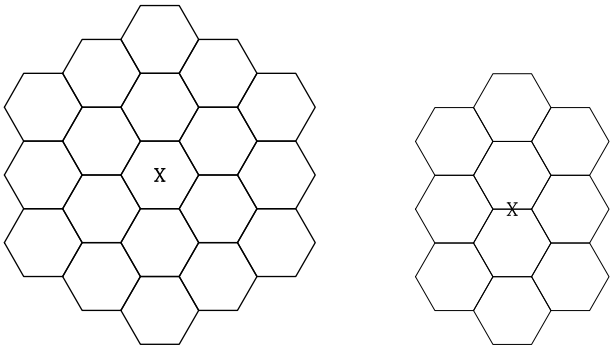

Figure 1: *Left:* stencil of grid cells affecting the outcome of thickness change calculations in the central cell (marked by an X). *Right:* stencil of grid cells affecting the calculation of mass fluxes across the central hexagonal edge segment (marked by an X).

Given a transport equation of the type u_t=$-F_x$, where F represents the flux of the variable u in x direction, the FCT scheme proceeds in two steps. In step 1, the u field is advanced in time using an upstream, forward-in-time ("low-order") scheme for computing F_x that due to its diffusive character is known to maintain positive-definiteness and monotonicity. In step 2, antidiffusive fluxes based on more accurate ("high-order") approximations are added to the diffusive fluxes. In what constitutes the essence of FCT, these antidiffusive fluxes are locally reduced or "limited" just enough to avoid violating the positive-definiteness constraint and creating new extrema in u.

In contrast to the usual practice of forming high-order fluxes from space-centered 4^{th} or 6^{th} order finite-difference expressions, high-order fluxes are constructed in FIM in the spirit of the (at best) 2^{nd} order PLM scheme (van Leer 1974; Colella and Woodward 1984). In one spatial dimension, PLM is monotonicity-preserving. It may be possible, with some effort, to preserve this property even in the present case where each cell exchanges mass with 5 or 6 neighbors. We have not yet explored this possibility, relying instead on FCT-type flux limiting. The spatial average of the transported variable in the slab upstream of a cell edge is computed by assuming that the variable changes linearly between cell center and cell edge. The variable's cell edge value is formed by averaging the two nearest vertex values, which in turn are unweighted averages of the 3 surrounding cell values. (To preserve 2^{nd} order accuracy of this interpolation scheme in the case of slightly distorted hexagons or in the presence of a pentagon, nonuniform weights would have to be introduced. This is planned for the future.)

To adapt the low-order transport scheme and the flux-limiting algorithm, both of which are inherently 2-time level schemes, to the 4-time level scheme (8), we proceed as follows:

1. Low-order fluxes from three consecutive time levels are blended as shown in (8) to generate a low-order solution at time level $n+1$.

2. The flux limiting process is based on "worst-case" tendencies of the transported variable, obtained by selectively bundling high-order incoming and outgoing fluxes. These tendencies are computed in FIM by blending, in the manner of the right-hand-side of (8), the current unclipped high-order flux (F^n) with the "final" (low-order plus clipped antidiffusive) fluxes from the previous two time levels (F^{n-1}, F^{n-2}).

3. The clipped antidiffuse fluxes are combined with the low-order ones to form final fluxes F^n; these are then used in (8) to compute the final value of u^{n+1}.

Lateral transport of θ and other state variables, such as moisture, is handled analogously. The prototype transport term for these variables is $\nabla\cdot(\theta\mathbf{v}\Delta p)$ in (7). Antidiffusive fluxes are limited in this case on the basis of extrema in the transported variable θ itself, not in the product $\theta\Delta p$. This is to say that we enforce monotonicity in the tracer *concentration* field, not the tracer *amount* field.

3e. *Vertical mesh*

As already mentioned, FIM is a *layer* model, meaning that the vertical spacing of grid points is allowed to vary in space and time. Since the prognostic equations resemble the shallow-water equations, layer models are also referred to as stacked shallow-water models.

The hybrid grid in FIM resembles that of RUC (Benjamin et al. 2004; Bleck and Benjamin 1993), but vertical staggering of variables has been changed because the RUC scheme does not rigorously conserve mass field constituents (referred to here as "tracers"). Vertical staggering in FIM replicates the layer treatment in the hybrid-isopycnic ocean model HYCOM (Bleck 2002) where only pressure and geopotential are carried on interfaces while all other variables are defined in layers. Coordinate layers conform to isentropic layers except in locations where the latter intersect the earth's surface. There, layers are locally redefined as terrain-following (σ coordinate) layers. An individual coordinate layer can be isentropic in one geographic region and terrain-following in another.

The hybridization concept employed here and in RUC differs from non-ALE hybrid schemes developed elsewhere (Bleck 1978b; Konor and Arakawa 1997; Pierce et al. 1991; Webster et al. 1999; Zapotocny et al. 1991, 1994) in that it relies on locally mandated adjustment of vertical grid spacing rather than on a fixed formula typically consisting of a weighted average of two or more traditional coordinate choices. The present scheme adds one important element to the original Arbitrary Lagrangian-Eulerian technique (Hirt et al. 1974): it provides a mechanism for keeping coordinate layers aligned with their designated target isentropes wherever possible. The original ALE scheme *(loc.cit.)* only concerned itself with the maintenance of nonzero grid spacing in Lagrangian coordinate simulations.

While the flexibility of coordinate placement in ALE-type schemes is disconcerting to some users because grid point location in model space cannot be expressed in terms of a simple analytic formula, it allows the model designer to maximize the size of the purely isentropic domain. The salient point here is that ALE can set the height above ground of the σ-to-θ coordinate transition in each geographic location separately, i.e., unencumbered by global considerations. This is a major advantage.

FIM manages the vertical grid structure as follows. If a given layer is "on target" (meaning that θ matches the target potential temperature assigned to this layer) and if, in addition, the 2-D shallow-water continuity equation (eq. (6) without the $\dot{s}$ terms)

$$\frac{\partial \Delta p}{\partial t} + \nabla \cdot (\mathbf{v} \Delta p) = 0 \tag{9}$$

yields a layer thickness Δp that does not fall below a predetermined minimum value, the Δp obtained from (9) is accepted. In other words, FIM sets $\dot{s} = 0$ in this case, meaning that it treats interfaces as material. If one of the above conditions is not met, the "grid generator" (see Appendix A) takes over and changes $\partial \Delta p / \partial t$ in a way that either maintains minimum thickness or, if the layer has become separated from its target potential temperature, moves it closer to it. In these situations, the selected Δp tendency is inserted into the full continuity equation (6) which at this point becomes a diagnostic equation for the interlayer mass fluxes $\dot{s} \partial p / \partial s$. The latter are used to vertically advect momentum and other variables.

Hydrostatic models infer the vertical component of motion from the vertically integrated horizontal mass flux divergence. The grid generator in an ALE model divides this material vertical motion into vertical motion of the coordinate surface and vertical air motion relative to it (Bleck 1978b):

$$\begin{pmatrix} \text{vertical} \\ \text{motion} \\ \underline{\textit{of}} \\ \text{interface} \end{pmatrix} + \begin{pmatrix} \text{vertical} \\ \text{motion} \\ \underline{\textit{through}} \\ \text{interface} \end{pmatrix} = \begin{pmatrix} \text{vertically} \\ \text{integrated} \\ \text{horizontal} \\ \text{mass-flux} \\ \text{divergence} \end{pmatrix} . \tag{10}$$

The decision whether to accept the solution of (9) – that is, whether to balance the right-hand side of (10) by only the first term or by some combination of both terms on the left – is made by the grid generator at each gid point and each time step individually. The grid generator in FIM actually performs two tasks: it carries out the "regridding" just described, followed by a vertical "remapping" of all prognostic variables to the modified grid. The remapping is formally equivalent to vertical advection because it is

driven by nonzero values of $\dot{s}$. However, since vertical displacement of atmospheric constituents due to actual air motion is already accounted for in the heaving and slumping of coordinate layers, the evaluation of the $\dot{s}$ terms in the prognostic equations is best viewed as a secondary property redistribution initiated by the migration of coordinate surfaces through resting air. With the atmosphere conceptually remaining at rest, remapping should conserve certain integral properties such as column integrals of momentum, thermal energy, etc.

3f. *Long-time step tracer transport*

The use of the numerically complex FCT scheme makes lateral transport costly in FIM. If the model were to be used to simulate the evolution of O(100) interacting chemical species, execution time would be prohibitive. One approach to reducing the amount of time spent in the FCT routine is to carry out tracer transport *intermittently*, using a longer time step. This is possible because the time step in FIM dynamics is controlled by the speed of gravity and Lamb waves, not by the typically much smaller wind speed which governs transport processes. In other words, advecting tracers using a time step geared toward maintaining numerical stability in gravity wave transmission is not very cost-effective.

Due to the fluctuating height of grid cells, tracer conservation during long-time step transport is not easily achieved in layer models. Since the transport equations in layer models are formulated in flux form, transport with a time step longer than Δt, say, $J\Delta t$ where $J > 1$, must be based on a rigorously time-integrated form of the mass continuity equation (6),

$$\frac{\Delta p^{n+J} - \Delta p^n}{J\Delta t} + \nabla_s \cdot \overline{\mathbf{v}\Delta p}^J + \left(\overline{\dot{s}\frac{\partial p}{\partial s}}^J\right)_2 - \left(\overline{\dot{s}\frac{\partial p}{\partial s}}^J\right)_1 = 0, \qquad (11)$$

where the overbar denotes integration over J time steps. To compute all terms in this equation and to assure that it holds, the dynamically active fields in FIM must have been stepped forward from time level n to $n+J$. At that instant, both the tendency term and the horizontal flux divergence term in (11) can be determined, the latter by summing up the instantaneous fluxes over the past J time steps. The time-integrated vertical flux terms can then be obtained by vertically summing up (11), using $\dot{s} = 0$ at the top or bottom of the column as a starting point.

By combining (11) with the equation $dQ/dt = 0$ expressing conservation of a tracer Q during transport (sources and sinks of Q can be evaluated separately), we arrive at the transport equation

$$\frac{(Q\Delta p)^{n+J} - (Q\Delta p)^n}{J\Delta t} + \nabla_s \cdot (Q\,\overline{\mathbf{v}\Delta p}^J) + \left(\overline{\dot{s}\frac{\partial p}{\partial s}}^J\,\hat{Q}\right)_2 - \left(\overline{\dot{s}\frac{\partial p}{\partial s}}^J\,\hat{Q}\right)_1 = 0. \qquad (12)$$

which can be solved for the tracer amount $Q\Delta p$ at time level $n+J$. The meaning of the caret and the method by which $Q\Delta p$ is converted to mixing ratio Q are discussed in the context of (7).

Eqn. (12) is solved by Flux Corrected Transport. Details are as follows.

1. Vertical Q fluxes are based on the Piecewise Parabolic Method (PPM). To avoid numerical stability problems posed by combinations of large vertical velocities and thin layers, integration of Q over the slab upstream of a given interface may extend over multiple layers.

2. The vertical PPM-based fluxes are used in conjunction with horizontal upstream fluxes to arrive at a low-order solution for Q.

3. High-order horizontal fluxes are of 2nd order accuracy, i.e., involve averages of Q over two neighboring grid cells.

4. The limiters applied to the antidiffusive (high-minus low-order) fluxes to assure monotonicity are based on the maxima and minima of "old" Q values in (a) the cell in question; (b) its lateral neighbors; and (c) the upstream slab(s) above or below the cell.

Comparisons of tracer fields advected over long and short time steps indicate that $J = 10$ works well in general, the single exception encountered so far being late-winter major stratospheric warming events when winds in the strato- and mesosphere can reach speeds exceeding 200 m/s. In those few cases, J had to be reduced to 5 to yield meaningful results.

Note that (6) is solved using the 3-time level Adams-Bashforth time differencing scheme whereas transport in (12) is carried out in forward-in-time mode. To achieve consistency between the thickness tendency term and the horizontal mass flux divergence, mass fluxes from 3 consecutive time levels must therefore be combined in the manner indicated in (8) before they are added to the flux time integral.

While 3-dimensional transport is the dominant process by which tracers are redistributed in the atmosphere, other processes such as subgridscale turbulent mixing cannot be neglected – especially if tracers advected by (12) are to evolve consistently with the primary mass field variables which in FIM are subjected to subgridscale vertical mixing. At the end of each long transport step, the relevant "physics" equations should therefore be solved for each tracer in question, using $J\Delta t$ as time step.

Two issues which arise in this context and are the subject of ongoing work shall be mentioned here:

1. Buoyancy-driven mixing in the isentropic subdomain generally leads to vertical mass fluxes $\dot{s}\partial p/\partial s$ which, if used in (12), redistribute tracers in the column in a manner closely paralleling the mixing of potential temperature. Hence, solving a separate diffusion equation in the isentropic subdomain might lead to excess mixing.

2. If different time steps are used to advect chemical tracers and hydrometeors, conservation issues arise if the two types of tracers interact. The problem is that during the J short time steps, chemical tracers are "locked up" in grid cells of thickness Δp^n whereas Δp and hydrometeor mixing ratio, being part of FIM dynamics, continue to evolve. Short of advecting hydrometeors on long time steps, the issue can be resolved by introducing two hydrometeor populations – population 1 remaining associated with the chemical tracers for the duration of the J short time steps and population 2 allowed to evolve with Δp and other dynamics-related variables. After tracers and population 1 have been subjected to transport via (12) and all progostic variables in FIM are therefore at time level $n+J$, the two hydrometeor populations must then be merged in a way that conserves all relevant properties acquired since time level n.

4. Model Physics

In order to facilitate comparison of FIM to existing operational NWP models, column physics parameterizations are taken directly from the Global Forecast System (GFS) of the NOAA National Centers for Environmental Prediction (NCEP). Treatment of vertical mixing, including the planetary boundary layer, is based on Hong and Pan (1996) and Troen and Mahrt (1986). The effects of short-wave radiation are modeled by a scheme described in Chou (1992), Chou and Lee (1996), Chou et al. (1998), while long-wave radiation is treated by the Rapid Radiative Transfer Model of Mlawer et al. (1997). Penetrative convection is based on Arakawa and Schubert (1974) as modified by Grell (1993). The effect of nonprecipitating shallow clouds is incorporated following Tiedtke (1983). Cloud formation on resolved scales is treated according to Sundqvist et al. (1989) as modified by Zhao and Carr (1997).

4a. *Temperature diagnostics*

Temperature is not a prognostic variable in FIM and hence must be inferred from θ and p when needed. Since model "dynamics" in FIM is formulated in terms of θ while model "physics" is formulated predominantly in terms of T, the $\theta - T$ conversion takes place frequently and in both directions. To avoid numerical degradation during this frequent back-and-forth, we define T, like θ, as a layer variable and solve physics equations in layers, not on interfaces.

There appears to be considerable freedom in how to define a "layer" pressure (or Exner function), needed for relating θ to T, in terms of interface pressure. One particularly compelling and widely used choice [e.g., Sela (1980), eq.10; Arakawa and Lamb (1977), eq. 250] is based on the notion that the column integral for the sum of potential and internal energy should not depend on whether it is written in terms of θ or T. The two forms of the integral are

$$\frac{c_p}{g}\int T dp = \frac{c_p p_0}{g(1+\kappa)}\int \theta d\left(\frac{p}{p_0}\right)^{1+\kappa} \qquad (13)$$

Equality is assured if

$$\frac{\theta}{1+\kappa} d \left(\frac{p}{p_0}\right)^{1+\kappa} = T d \left(\frac{p}{p_0}\right)$$

in each model layer. This condition is met if the Exner function value relating θ to T in a model layer is defined as a finite-difference analog of

$$\frac{c_p}{1+\kappa} \frac{\partial (p/p_0)^{1+\kappa}}{\partial (p/p_0)}.$$

By satisfying (13), the model correctly translates temperature changes resulting from, for example, radiation or cloud physics into available potential energy changes represented in FIM in terms of p and θ.

4b. *Vertical motion diagnostics*

One particular parameter required by the convective cloud parameterization scheme is the updraft velocity $\omega \equiv dp/dt$. It is diagnosed in FIM by expanding the material derivative dp/dt into

$$\frac{dp}{dt} = \left(\frac{\partial p}{\partial t}\right)_s + \mathbf{v} \cdot \nabla_s p + \left(\dot{s} \frac{\partial p}{\partial s}\right) \qquad (14)$$

and evaluating the three terms on the right individually.

By virtue of (1), the sum of the first and third term on the r.h.s. of (14) at a given horizontal location x, y and level $s = S$ is given by the horizontal mass flux divergence at x, y integrated from the top of the atmosphere to level S:

$$\left[\frac{\partial p}{\partial t} + \dot{s} \frac{\partial p}{\partial s}\right]_S = - \int_{top}^{S} \nabla_s \cdot \left(\mathbf{v} \frac{\partial p}{\partial s}\right) ds.$$

According to (6), the finite-difference analog of this expression, valid on the lower interface of layer k (k increasing upward), is

$$\left[\frac{\partial p}{\partial t} + \dot{s} \frac{\partial p}{\partial s}\right]_{k-1/2} = - \sum_{n=k}^{k_{top}} \nabla \cdot (\mathbf{v}_n \Delta p_n). \qquad (15)$$

Completing the calculation of ω on an interface would require evaluating the second term on the r.h.s. of (14) on that interface. This is not straightforward because horizontal air motion is usually different on opposite sides of an interface, yielding two possible values of $(\mathbf{v} \cdot \nabla p)$. This dilemma can be avoided by evaluating (14) as a layer average, not an interface value. Fortuitously, the convective parameterization scheme calls for a layer average.

Generating a layer average of (15) is straightforward owing to the fact that $\mathbf{v}$ does not vary within a layer:

$$\left[\frac{\partial p}{\partial t} + \dot{s} \frac{\partial p}{\partial s}\right]_k = -\frac{1}{2} \nabla_s \cdot (\mathbf{v}_k \Delta p_k) - \sum_{n=k+1}^{k_{top}} \nabla_s \cdot (\mathbf{v}_n \Delta p_n).$$

For the same reason, the layer average of the 2nd term on the r.h.s. of (14) is simply $\mathbf{v}_k \cdot \nabla \overline{p}_k$ where $\overline{p}_k$ is the mid-layer pressure. In summary,

$$\omega_k = \mathbf{v}_k \cdot \nabla \overline{p}_k - \frac{1}{2} \nabla_s \cdot (\mathbf{v}_k \Delta p_k) - \sum_{n=k+1}^{k_{top}} \nabla_s \cdot (\mathbf{v}_n \Delta p_n). \qquad (16)$$

5. Sequence of Operations

Variables are updated during each model time step in the following order.

1. Starting from momentum $\mathbf{v}^n$ and layer thicknes Δp^n at time level n, preliminary values at time level $n+1$ are obtained by solving (5) and (9) under the assumption that all interfaces are material ($\dot{s} = 0$). We refer to the resulting values as $\mathbf{v}_{sw}^{n+1}, \Delta p_{sw}^{n+1}$ where subscript sw stands for *shallow water*, a reminder that the 2-D versions of the respective prognostic equations have been used.

2. Thermal energy $(\theta \Delta p)$ is advanced in time by solving (7), once again with $\dot{s}$ set to zero, and with the r.h.s. set to zero. The outcome of this process is $(\theta \Delta p)_{sw}^{n+1}$.

3. Values of θ_{sw}^{n+1} are obtained by dividing $(\theta \Delta p)_{sw}^{n+1}$ resulting from step 2 by Δp_{sw}^{n+1}. Monotonicity constraints are applied to safeguard against indeterminacies in the limit of zero layer thickness.

4. Diabatic forcing due to radiation, surface fluxes, release of latent heat, etc., are evaluated using the GFS physical parameterization module

based on $\theta_{sw}^{n+1}, \Delta p_{sw}^{n+1}$, and other state variables at time level $n+1$. These calculations yield an updated potential temperature θ_{phy}^{n+1} where *phy* stands for model physics.

5. Fields of $\theta_{phy}^{n+1}, \Delta p_{sw}^{n+1}$ are fed to the grid generator which decides on the magnitude of interface fluxes $\dot{s}\partial p/\partial s$ at each grid point. These flux values are used to evaluate the missing vertical advection terms in (5), (6), (7), yielding final values $\mathbf{v}^{n+1}, \Delta p^{n+1}, \theta^{n+1}$.

Other variables carried by the model to define the physical state of the atmosphere (moisture, hydrometeors,...) are advanced in time like the variable θ. This is to say that transport takes place in flux form analogously to (7); source terms are evaluated as in step 4; and the variables are advected vertically as part of step 5.

Vertical advection of model variables, only necessary in regions where $p_{sw}^{n+1} \neq p^{n+1}$, is implemented as a vertical remapping of the stairstep profiles resulting from the preliminary shallow-water integration. To remain stable in situations where layer thickness approaches zero while $\dot{s}\partial p/\partial s$ remains finite, the remapping algorithms available in FIM are formulated to allow Courant numbers > 1.

6. Model Initialization

Initial conditions for FIM are based on fields provided by NOAA's Global Forecast System. The state of the atmosphere is represented in that system by surface pressure and by layer averages of virtual temperature T_v, humidity, ozone concentration, and horizontal velocity components in 64 hybrid $\sigma - p$ layers on a spherical grid. These fields are processed as follows.

1. Values of geopotential ϕ on $\sigma - p$ interfaces[1] on the spherical grid are obtained by integrating the hydrostatic equation layerwise in the form

$$\partial\phi/\partial\Pi = -\theta_v \qquad (17)$$

where $\theta_v = T_v(p_0/p)^{R/c_p}$ is the potential virtual temperature.

[1] interfaces = coordinate surfaces separating coordinate layers

2. Terrain height, interface pressure, interface geopotential, as well as layer averages of humidity, wind, and ozone concentration are interpolated horizontally to the icosahedral grid.

3. Virtual potential temperature in the original $\sigma - p$ layers on the icosahedral grid is deduced from pressure and geopotential, once again using (17). Deducing θ_v from ϕ minimizes the risk of introducing hydrostatic height errors during horizontal interpolation.

4. In each grid column on the icosahedral grid, the stairstep profile defined in terms of θ_v and Π in the $\sigma - p$ system, to be referred to as $\theta_{in}(\Pi)$, is converted into a new stairstep profile $\theta_{out}(\Pi)$ in which the stair "risers" are the prescribed θ coordinate values. The height values for the "landings", i.e., the horizontal sections in $\theta_{out}(\Pi)$, are the unknowns in this transformation problem.

   The transformation, which we refer to as "restepping" and which is described more fully in Appendix A, can result in the formation of one or more zero-thickness layers at the top and bottom of the column. Those at the bottom will subsequently be inflated and converted into terrain-following σ layers.

   In order to minimize truncation problems while converting one stairstep profile into another, the restepping process is broken into two steps. First, the piecewise constant profile $\theta_{in}(\Pi)$ is converted into a continuous, piecewise *linear* profile using an extension of the integral-conserving method described in Bleck (1984b). The linear segments are then integrated piecewise to form a stairstep profile with "risers" in the appropriate places.

   Use of the Exner function Π as opposed to p as vertical coordinate guarantees that the height $\int\theta d\Pi$ of the input column is preserved during each step of the vertical coordinate transform. Without this constraint, large-amplitude external gravity waves would likely be excited in the model at the beginning of the forecast.

5. The grid generator (see Appendix A) is invoked to inflate zero-thickness layers at the surface that may have been generated while transforming the original $\sigma - p$ layers to isentropic layers. The piecewise linear θ profile in the column

is then integrated over the newly formed FIM-specific σ layers.

6. Moisture, ozone concentration, and velocity components are expanded into piecewise linear profiles and then integrated over the hybrid $\sigma - \theta$ coordinate layers resulting from the previous step.

7. Fields of Montgomery potential are obtained by integrating the hydrostatic equation (4). Integration starts with M defined in the lowest layer as $M_1 = \Pi_{sfc}\theta_1 + \phi_{sfc}$ where θ_1 is the lowest coordinate value and sfc stands for values at ground level. (Layer 1 may be massless but must be in contact with the ground for this formula to hold.)

7. Outlook

This document is likely to evolve as the team gains experience with FIM and the dust settles on implementation details that are too much in flux at the time of this writing to warrant documentation.

8. Appendix A: The Vertical Grid Generator

8a. *Background*

The first-generation HYCOM grid generator, whose design principles are described in some detail in Bleck (2002), has been modified and tuned over the years to address grid degeneracies that came to light as the range of applications of HYCOM grew. This tuning has added branches to the decision tree in the original algorithm, creating a situation where the underlying logic is no longer transparent to the user. Complexity in the grid generator discourages experimentation and adaptation of HYCOM/FIM to special modeling needs, and hence should be avoided.

The algorithm described below represents an attempt to get "back to basics" when moving layer interfaces for the sake of maximizing the part of the atmosphere represented by isentropic layers while at the same time subjecting the layers to minimum-thickness constraints. In the first-generation grid generator, each grid point is inspected and adjusted recursively in light of its distance to grid points above and below, using a variety of semi-empirical criteria. The algorithm proposed here is more straightforward in that it begins by transforming a given hybrid stairstep potential temperature (θ) profile into a purely isentropic one, i.e., into a stairstep profile whose θ levels are prescribed beforehand. Depending on the stratification and θ range in the original profile, this process can produce massless (zero thickness) layers at the top and bottom of the column. Massless layers that occur at the ground are subsequently inflated to a prescribed minimum thickness.

As outlined in Bleck (2002), differences between the resulting hybridized layer interface pressures and those of the input profile imply mass exchange among layers. Model variables such as humidity and momentum must then be exchanged between layers as well. Any one of the standard conservative advection schemes can be used for this task.

8b. *Transformation of non-isentropic stairstep θ profiles to isentropic coordinates*

The following is an adaptation of the ocean-oriented scheme described in Appendix D of Bleck (2002).

Let $\mathcal{Z}$ be a monotonic function of p decreasing with height, and let $\theta_{in}(\mathcal{Z})$ be a piecewise constant ("stairstep") vertical profile of θ. Both the step width $\Delta\theta_{in}$ and the step height $\Delta\mathcal{Z}$ can vary from step to step. Our task is to transform $\theta_{in}(\mathcal{Z})$ into another stairstep profile differing from the original one in that the location of the "risers" on the θ axis is prescribed. Ideally, the transformation should be accomplished without perturbing the potential/internal energy of the column. Another quantity worth preserving is the geopotential height of the column, because a transformation that changes the column height is likely to set off external gravity waves.

Let θ_k ($k = 1, ..., n$, $\theta_{k+1} > \theta_k$, k increasing upward) mark the points on the θ axis where we want the new risers to be placed. We require that the θ_k

values span the θ range of the input profile,

$$\theta_1 \le \theta_{in}(\mathcal{Z}) \le \theta_n \text{ for all } \mathcal{Z}, \tag{18}$$

and that the input profile be monotonic. Denoting the pressure[2] of the lower and upper interface bounding the k-th layer by $\mathcal{Z}_{k-1/2}$ and $\mathcal{Z}_{k+1/2}$, respectively, the condition we wish to satisfy can then be stated as

$$\sum_{k=1}^{n} \theta_k(\mathcal{Z}_{k-1/2} - \mathcal{Z}_{k+1/2}) = \int_{\mathcal{Z}_{n+1/2}}^{\mathcal{Z}_{1/2}} \theta_{in}\, d\mathcal{Z}. \tag{19}$$

The interface pressures are the unknowns in the problem.

Integration by parts (on the left this amounts to reordering the terms under the summation sign) allows us to rewrite (19) as

$$\theta_1\mathcal{Z}_{1/2} - \theta_n\mathcal{Z}_{n+1/2} + \sum_{k=1}^{n-1} \mathcal{Z}_{k+1/2}(\theta_{k+1} - \theta_k)$$
$$= [\theta_{in}\mathcal{Z}]_{\mathcal{Z}_{n+1/2}}^{\mathcal{Z}_{1/2}} + \int_{\theta_{in}(\mathcal{Z}_{1/2})}^{\theta_{in}(\mathcal{Z}_{n+1/2})} \mathcal{Z}_{in}(\theta)\, d\theta$$

where $\mathcal{Z}_{in}(\theta)$ is the inverse of $\theta_{in}(\mathcal{Z})$.

In situations where the θ range of the input profile does not span the entire range $\theta_1 ... \theta_n$, we can, without altering the physical appearance of the input profile, lower $\theta_{in}(\mathcal{Z}_{1/2})$ to θ_1 and/or raise $\theta_{in}(\mathcal{Z}_{n+1/2})$ to θ_n. With these modifications, the above expression reduces to

$$\sum_{k=1}^{n-1} \mathcal{Z}_{k+1/2}(\theta_{k+1} - \theta_k) = \int_{\theta_1}^{\theta_n} \mathcal{Z}_{in}(\theta)\, d\theta. \tag{20}$$

Our strategy is to satisfy (20) by breaking the integral into pieces taken over intervals (θ_k, θ_{k+1}) and conserving each integral individually. This immediately leads to

$$\mathcal{Z}_{k+1/2} = \frac{1}{\theta_{k+1} - \theta_k} \int_{\theta_k}^{\theta_{k+1}} \mathcal{Z}_{in}(\theta)\, d\theta \tag{21}$$

$(k = 1, ..., n-1)$. If condition (18) is violated, evaluation of (21) is postponed until the offending input layer is brought into compliance by "diluting" it with mass from adjacent layers. Persistent heating at the model top, for example, is thereby transformed into a gradual thickening of the uppermost coordinate layer.

[2] While $\mathcal{Z}$ is a function of pressure, we will refer to it as pressure for short.

8c. *Enforcement of layer thickness constraints*

Suppose the prescribed potential temperature values $\theta_1, \theta_2, ...$ in the output profile cover a wide enough range to yield $\theta_k < \theta_{in}$ for some $k > 1$. In this case, (21) will yield $\mathcal{Z}_{1/2} = \mathcal{Z}_{3/2} = ... = \mathcal{Z}_{k+1/2}$, i.e., layers $1, ..., k$ in the transformed profile will be massless. Likewise, if $\theta_{in} < \theta_k$ for some $k < n$, layers $k+1, ..., n$ will be rendered massless ($\mathcal{Z}_{k+1/2} = ... = \mathcal{Z}_{n+1/2}$).

The strategy in FIM is to accept massless layers aloft, but to always inflate massless layers at the bottom of the grid column. Layer inflation rules can be as simple as specifying a constant minimum thickness $\Delta\mathcal{Z}_0$. In this case the set of isentropic interface values $\mathcal{Z}_{k+1/2}$ obtained from (21), to be identified here as $\hat{\mathcal{Z}}_{k+1/2}$ to distinguish them from the final "hybridized" values, are recursively subjected to the constraint

$$\mathcal{Z}_{k+1/2} = \min(\hat{\mathcal{Z}}_{k+1/2}, \mathcal{Z}_{k-1/2} - \Delta\mathcal{Z}_0) \tag{22}$$

$(k = 1, 2, ...)$. Note that $\Delta\mathcal{Z}_0$ can easily be made layer-dependent or scaled by terrain height. (Scaling by mixed layer depth, on the other hand, is problematic as this will spawn large vertical displacements of layer interfaces during day-night mixed layer transitions. Such displacements cause excessive interlayer mass exchange whose suppression is one of the original motivations for isentropic modeling.)

It is advisable to smooth out large lateral variations in layer thickness that typically occur where a hybridized coordinate layer transitions from the fixed-depth to the isentropic subdomain. These variations are created when, for a given k, the 2^{nd} argument in the *minimum* function of (22) is chosen in one grid column, while the 1^{st} argument is chosen in a neighboring column. One way to smooth out the transition, short of exchanging information among neighboring grid columns, is to increase layer thickness in situations where the two arguments are of similar magnitude. This is the purpose of the "cushion" function originally introduced into hybrid-coordinate ocean modeling by Bleck and Boudra (1981) and later adapted for atmospheric use by Bleck and Benjamin (1993). Use of the cushion function entails replacing (22) by

$$\mathcal{Z}_{k+1/2} = \min(\hat{\mathcal{Z}}_{k+1/2}, \mathcal{Z}_{k-1/2}$$
$$-cushn[\hat{\mathcal{Z}}_{k-1/2} - \mathcal{Z}_{k+1/2}, \Delta\mathcal{Z}_0]).$$

In the two extreme cases where $\hat{\mathcal{Z}}_{k-1/2} - \mathcal{Z}_{k+1/2}$ is either large negative or large positive compared to $\Delta\mathcal{Z}_0$, the cushion function is designed to replicate the functionality of (22). In other words, $cushn(a,b)$ returns a if $a >> b$, and it returns b if $-a >> b$ $(b > 0)$. In between the two extremes, $cushn$ varies smoothly, returning values as high as $2\max(a,b)$. In many cases, this widens a layer if its potential temperature is close to target, thereby softening the lateral interface height contrast between locations where the underlying layer is isentropic and where it is not.

If more effective interface smoothing in the σ-θ transition region is deemed nececessary, a sideways-looking smoothing algorithm may be required.

At the time of this writing, the minimum thickness value $\Delta\mathcal{Z}_0$ is set as follows.

1. A default value $\Delta\mathcal{Z}_0(k)$ is specified for each layer k. Typical values (stated here in pressure units for easier reference) are 3 hPa in the bottom layer, gradually increasing to 10 or 15 hPa in layers above.

2. In an attempt to mimic the vertical spacing of conventional σ coordinate layers, all $\Delta\mathcal{Z}_0(k)$ are multiplied by the factor $(p_{srf} - p_{top})/(1000\,hPa - p_{top})$ where p_{srf} is the surface pressure and p_{top} is the pressure level (400 hPa or smaller) where coordinate surfaces in a conventional σ coordinate model cease to be terrain-following.

3. Starting in the lowest layer and moving up the column, $\mathcal{Z}_{k+1/2}$ is compared against the lesser of (23) and $\mathcal{Z}_{1/2} - \sum_{n=1}^{k} \Delta\mathcal{Z}_0(n)$. If it exceeds the minimum of these two values, it is replaced by that minimum. This is done recursively, i.e., altered interface values affect the inflation test in layers above.

4. The lowest layer *not* in need of inflation is labeled $k_{\sigma\theta}$; it marks the transition from the σ to the θ subdomain.

5. The upper interface of layer $k_{\sigma\theta}$ stays fixed by definition, but very thin isentropic layers qualifying for inflation based on (23) can occur higher up in the atmosphere. To keep these from unnecessarily being inflated, the value $\Delta\mathcal{Z}_0(k)$ is reduced in layers $k_{\sigma\theta}+1, ..., k_{\sigma\theta}+4$ by the factors 0.4, 0.2, 0.1, and 0.05, respectively. The factor 0.05 is also used in layers $k > k_{\sigma\theta}+4$.

8d. *Vertical advection*

The "regridding" process described in the preceding sections must be followed by a "remapping" process in which model variables are advected vertically in response to changes in interface pressure. Borrowing from HYCOM, vertical advection of momentum as well as tracers such as humidity, liquid water content, etc., is currently handled by either the piecewise linear or the piecewise parabolic method (PLM, PPM) (van Leer 1974; Colella and Woodward 1984). All variables mentioned are remapped in p space to conserve their mass-weighted column integral.

Potential temperature is a special case. The regridding process described earlier yields a new θ distribution that may be viewed as resulting from upstream or donor cell advection in $\mathcal{Z}$ space. To suppress the numerical diffusivity implied by this low-order scheme, FIM actually discards the θ field resulting from the regridding exercise and replaces it by a field advected by the same higher-order scheme that is used for the other prognostic variables.

There is a price to be paid for inferring the amount of mass transferred between layers from a piecewise constant θ distribution, as is done in (19)–(21), and subsequently using a higher-order scheme to remap θ. Neither will the slab of air arriving in a layer have the potential temperature needed to precisely restore that layer to target, nor will the transfer leave θ in the donor layer unchanged. However, we find that repetitive use of the restoring algorithm allows layers to reach their target relatively quickly.

One advantage of using a higher-order advection scheme for θ is that the variable $\mathcal{Z}$ in (19) – (23) no longer needs to be chosen with an eye on the conservation properties of the regridding scheme. Any variable monotonic in pressure, including p itself, is acceptable. What matters now is the vertical coordinate used during *remapping* of θ. FIM allows the use of either $(p/p_0)^{\kappa}$ or $(p/p_0)^{1+\kappa}$ (where $\kappa = R/c_p$). The rationale for providing these two options is given in the following section.

8e. *Conservation alternatives*

It follows from (17) that the height of an air column can be preserved during vertical regridding/remapping by setting $\mathcal{Z} = \Pi$. Unfortunately, this choice of $\mathcal{Z}$ does not allow us to satisfy another important constraint: conservation of column-integrated internal energy $I = \int c_v T dp$ and column-integrated potential energy $P = \int g\rho z\, dz$. (In an ideal gas, internal and potential energy are proportional to one another, so conservation of one entails conservation of the other.) The incompatibility of column height conservation with internal/potential energy conservation becomes clear if one writes P and I in terms of θ and p and compares the resulting expressions

$$I = \frac{c_v p_0}{g(1+k)} \int \theta\, d \left(\frac{p}{p_0}\right)^{1+k} \tag{23}$$

$$P = \frac{R\, p_0}{g(1+k)} \int \theta\, d \left(\frac{p}{p_0}\right)^{1+k} \tag{24}$$

($k = R/c_p$) with the formula for column height, $\int \theta d\Pi$, in which θ is integrated over a variable proportional to $(p/p_0)^\kappa$. It is easy to see now that conservation of I and P can be achieved during remapping of θ by using $(p/p_0)^{1+k}$ as vertical coordinate, but that this can only be done at the price of violating the height preservation constraint. The relative importance of height versus internal/potential energy conservation is hard to assess without practical tests.

8f. *Miscellaneous refinements*

Discretization of a continuous profile $\theta(p)$ in term of a stairstep profile is not unique, because stairsteps can be broken into smaller steps or combined into bigger ones without violating any continuity or conservation principle. This ambiguity can lead to computational modes in the vertical layer structure, leading to the gradual disappearance of, say, odd-numbered layers accompanied by a thickening of even-numbered ones. Initial experiments with FIM indeed revealed a propensity for amplifying this mode. To suppress it, a special algorithm has been added to the grid generator.

The algorithm is still in a state of development. Its present version scans each grid column for sequences of 5 Δp values, numbered $\Delta p_1, ..., \Delta p_5$, that satisfy the following three conditions:

$$\begin{aligned} \Delta p_1 &< \Delta p_2 \\ \Delta p_5 &< \Delta p_4 \\ \Delta p_3 &< \min(\Delta p_2, \Delta p_4). \end{aligned}$$

If all three conditions are met, layer 3 is inflated by drawing mass from both layers 2 and 4 such that (a) the column integral of θ is conserved and (b) $\Delta p_3^{new} = \min(\Delta p_2^{new}, \Delta p_4^{new})$. Requirement (a) leads to the constraint

$$\frac{\Delta p_2 - \Delta p_2^{new}}{\Delta p_4 - \Delta p_4^{new}} = \frac{\theta_4 - \theta_3}{\theta_3 - \theta_2}$$

which may put a limit on the mass transfer stipulated by (b). The resulting interface displacements are added to those associated with the primary regridding process.

Suppression of the layer thickness computational mode improves the performance of the GFS column physics parameterization scheme which has been found to be sensitive to large variations in layer thickness.

9. Appendix B: Turbulent Vertical Mixing

9a. *Background*

The following is a simplified version of a numerical scheme developed by McDougall and Dewar (1998) for carrying out vertical mixing in fluid models whose vertical coordinate is a function of the diffused variable(s). They deal with the specific problem of mixing temperature and salinity in ocean models whose vertical coordinate is potential density (a function of both temperature and salinity), constrained to remain constant in each coordinate layer during mixing.

Here we deal with the much simpler problem of solving the diffusion equation in an atmospheric column where there is only one diffused variable (potential temperature θ) doing double duty as vertical coordinate. The only variable capable of capturing the effects of thermal diffusion in this case is the thickness of coordinate layers.

9b. *The mixing scheme*

The equations expressing conservation of mass and heat in a column, basically 1-D versions of the equations listed in the beginning, are

$$\frac{\partial}{\partial t}\left(\frac{\partial z}{\partial s}\right)_s + \frac{\partial}{\partial s}\left(\dot{s}\frac{\partial z}{\partial s}\right) = 0. \qquad (25)$$

$$\left(\frac{\partial \theta}{\partial t}\right)_s + \left(\dot{s}\frac{\partial z}{\partial s}\right)\frac{\partial \theta}{\partial z} = -\frac{\partial F_\theta}{\partial z} \qquad (26)$$

The turbulent heat flux $F_\theta = \overline{w'\theta'}$ is usually parameterized as $F_\theta = -K\partial\theta/\partial z$ where θ is the resolved-scale potential temperature and K is a thermal diffusivity coefficient.

The flux form of (26), obtained by combining (25) and (26), is

$$\frac{\partial}{\partial t}\left(\theta\frac{\partial z}{\partial s}\right) + \frac{\partial}{\partial s}\left(\dot{s}\frac{\partial z}{\partial s}\theta\right) = -\frac{\partial F_\theta}{\partial s}. \qquad (27)$$

The task at hand is to discretize the above equations for use in a model framework where stratification is represented by a piecewise constant, stairstep θ profile. The discretization will be done by formally integrating the equations over individual stairsteps.

If θ is to remain constant in each layer during the mixing process, F_θ must be vertically constant in each layer. If this were not the case, integrating (26) over an individual layer would yield a nonzero right-hand side. Of the two terms on the left, the second one integrates to zero since $\partial\theta/\partial z = 0$ inside the layer. (The vertical mass flux $(\dot{s}\partial z/\partial s)$ remains finite.) Hence, a nonzero r.h.s. implies a nonzero tendency term $\partial\theta/\partial t$ which clashes with the stated requirement.

We conclude: for diffusion to leave a mark on the profile under the constraint $F_\theta = const$ in individual layers, F_θ must be allowed to vary from layer to layer. The implied infinite heat flux divergence at layer interfaces is consistent with the notion that air crossing an interface undergoes an instantaneous change in θ.

A simple centered finite-difference expression for the heat flux in layer n is

$$F_\theta^n = \frac{K^n}{2}\,\frac{\theta^{n+1} - \theta^{n-1}}{z^{n+1/2} - z^{n-1/2}} \qquad (28)$$

where fractional superscripts indicate quantities defined on interfaces.

The central task is to determine the mass flux across layer interfaces, $(\dot{s}\partial z/\partial s)$. For this we integrate (27) over an s interval representing an infinitesimal slab bracketing a layer interface. Since the tendency term drops out as ∂z approaches zero and the mass flux $(\dot{s}\partial z/\partial s)$ is continuous in the vertical, we obtain in the limit of zero slab thickness

$$\left(\dot{s}\frac{\partial z}{\partial s}\right)^{n+1/2} = \frac{F_\theta^{n+1} - F_\theta^n}{\theta^{n+1} - \theta^n}. \qquad (29)$$

Expressions (28) and (29) encompass the sought-after solution to the problem of diffusing heat in a stairstep θ profile while maintaining θ in individual layers. Note that, in the absence of externally imposed heat fluxes, the column integral $\int\theta dz$ is conserved regardless of the physical and numerical approximations made in evaluating the heat flux (28).

The heat flux as approximated by (28) becomes infinite in massless layers. To avoid division by zero, the denominator in (28) must therefore be bounded away from zero. The parameter representing minimum layer thickness, together with K and the time step used in solving (27), can be tuned to concentrate the effect of vertical diffusion almost entirely on very thin layers. We use the scheme in this mode as an alternative to the grid generator to avoid generating zero-thickness layers in the isentropic subdomain that may result from strongly layer-dependent diabatic forcing. The advantage of the present scheme over the grid generator is that it does not produce local deviations from target θ. Suitable parameter values are: 1 m^2 for the product of time step and mixing coefficient, and 2×10^{-3}m for the minimum thickness.

Ideally, vertical mixing should conserve the total heat content of the column, $\int c_v T dp$. From (23) we note that in order to conserve total heat, the variable z in (25) – (29) must be replaced by a variable proportional to p^{1+k}. No other changes are required in the solution procedure, except that the mixing coefficient K in (28) must be rendered dimensionally compatible with the new vertical coordinate.

In a similar vein, preservation of the total height $\int\theta d\Pi$ of the column during mixing can be achieved by using a variable proportional to p^k in place of z. Note that height preservation is incompatible with

heat conservation.

References

Arakawa, A. and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Meth. Comput. Phys.*, **17**, Academic Press, New York, 173–265.

Arakawa, A. and W. H. Schubert, 1974: Interaction of a cumulus ensemble with the large-scale environment, part 1. *J. Atmos. Sci.*, **31**, 674–704.

Benjamin, S., G. Grell, J. Brown, T. Smirnova, and R. Bleck, 2004: Mesoscale weather prediction with the ruc hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494.

Bleck, R., 1978a: Finite difference equations in generalized vertical coordinates. part i: Total energy conservation. *Beitr. Phys. Atm.*, **51**, 360–372.

— 1978b: On the use of hybrid vertical coordinates in numerical weather prediction models. *Mon. Wea. Rev.*, **106**, 1233–1244.

— 1984a: An isentropic coordinate model suitable for lee cyclogenesis simulation. *Riv. Meteorol. Aeronaut.*, **44**, 189–194.

— 1984b: Vertical coordinate transformation of vertically discretized atmospheric fields. *Mon. Wea. Rev.*, **112**, 2537–2541.

— 2002: An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. *Ocean Modelling*, **4**, 55–88.

Bleck, R. and S. Benjamin, 1993: Regional weather prediction with a model combining terrain-following and isentropic coordinates. part 1: model description. *Mon. Wea. Rev.*, **121**, 1770–1785.

Bleck, R. and D. B. Boudra, 1981: Initial testing of a numerical ocean circulation model using a hybrid (quasi-isopycnic) vertical coordinate. *J. Phys. Oceanogr.*, **11**, 755–770.

Boris, J. P. and D. L. Book, 1973: Flux-corrected transport. i. shasta, a fluid transport algorithm that works. *J. Comput. Phys.*, **11**, 38–69.

Chou, M.-D., 1992: A solar radiation model for use in climate studies. *J. Atmos. Sci.*, **49**, 762–772.

Chou, M.-D. and K. T. Lee, 1996: Parameterizations for the absorption of solar radiation by water vapor and ozone. *J. Atmos. Sci.*, **53**, 1204–1208.

Chou, M.-D., M. J. Suarez, C. H. Ho, M. M. H. Yan, and K. T. Lee, 1998: Parameterizations for cloud overlapping and shortwave single scattering properties for use in general circulation and cloud ensemble models. *J. Climate*, **11**, 202–214.

Colella, P. and P. Woodward, 1984: The piecewise parabolic method (ppm) for gas-dynamical simulations. *J. Comput. Phys.*, **54**, 174–201.

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

Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.*, **121**, 764–787.

Heikes, R. H. and D. A. Randall, 1995: Numerical integration of the shallow-water equations on a twisted icosahedral grid. i. basic design and results of tests. *Mon. Wea. Rev.*, **123**, 1862–...

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.*, **75**, 1825–1830.

Hirt, C. W., A. A. Amsden, and J. L. Cook, 1974: An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *J. Comput. Phys.*, **14**, 227–253.

Hong, S.-Y. and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339.

Kasahara, A., 1974: Various vertical coordinate systems used for numerical weather prediction. *Mon. Wea. Rev.*, **102**, 509–522.

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

Lee, J.-L., G. L. Browning, and Y.-F. Xie, 1995: Estimating divergence and vorticity from the wind profiler network hourly wind measurements. *Tellus*, **47A**, 892–910.

Lee, J.-L. and A. E. MacDonald, 2000: Qnh: Mesoscale bounded derivative initialization and winter storm test over complex terrain. *Mon. Wea. Rev.*, **128**, 1037–1051.

Lin, S. J., , and R. B. Rood, 1997: An explicit flux-form semi-lagrangian shallow-water model on the sphere. *Quart. J. Roy. Meteor. Soc.*, **123**, 2477–2498.

Lin, S. J., W. Chao, Y. Sud, and G. Walker, 1994: A class of the van leer-type transport schemes and its application to the moisture transport in a general circulation model. *Mon. Wea. Rev.*, **122**, 1575–1593.

MacDonald, A. E., J. L. Lee, and S. Sun, 2000: Qnh: Design and test of a quasi-nonhydrostatic model for mesoscale weather prediction. *Mon. Wea. Rev.*, **128**, 1016–1036.

Majewski, D., D. LierMann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, and W. Wergen, 2002: The operational global icosahedral-hexagonal gridpoint model gme: Description and high-resolution tests. *Mon. Wea. Rev.*, **130**, 319–338.

McDougall, T. J. and W. K. Dewar, 1998: Vertical mixing and cabbeling in layered models. *J.Phys.Oceanogr.*, **28**, 1458–1480.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682.

Pierce, R., D. Johnson, F. Reames, T. Zapotocny, and B. Wolf, 1991: Numerical investigations with a hybrid isentropic-sigma model, part i: Normal mode characteristics. *J. Atmos. Sci.*, **48**, 2005–2024.

Ringler, T. D., R. P. Heikes, and D. A. Randall, 2000: Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores. *Mon. Wea. Rev.*, **128**, 2471–2490.

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.

Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. *Mon. Wea. Rev.*, **108**, 1279–1292.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud studies with mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1757.

Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ecmwf model. *ECMWF Workshop on Convection in Large-Scale Models*, 297–316.

Tomita, H., M. Satoh, and K. Goto, 2004: A new dynamical framework of global nonhydrostatic model using the icosahedral grid. *Fluid Dyn. Res.*, **34**, 357–400.

Tomita, H., M. Tsugawa, M. Satoh, and K. Goto, 2001: Shallow water model on a modified icosahedral geodesic grid by using spring dynamics. *J. Comput. Phys.*, **174**, 579–613.

Troen, I. and L. Mahrt, 1986: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Bound.-Layer Meteor.*, **37**, 129–148.

van Leer, B., 1974: Towards the ultimate conservative difference scheme. ii. monotonicity and conservation combined in a second order scheme. *J. Comput. Phys.*, **14**, 361–370.

— 1977: Towards the ultimate conservative difference scheme. iii. upstream-centered finite-difference schemes for ideal compressible flow. *J. Comput. Phys.*, **23**, 263–275.

Webster, S., J. Thuburn, B. J. Hoskins, and M. J. Rodwell, 1999: Further development of a hybrid-isentropic GCM. *Quart. J. Roy. Meteor. Soc.*, **125**, 2305–2331.

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

Williamson, D., 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**, 221–224.

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

Zapotocny, T., D. Johnson, and F. Reames, 1994: Development and initial test of the University of Wisconsin global isentropic-sigma model. *Mon. Wea. Rev.*, **122**, 2160–2178.

Zapotocny, T., D. Johnson, F. Reames, R. Pierce, and B. Wolf, 1991: Numerical investigations with a hybrid isentropic-sigma model, part ii: The inclusion of moist processes. *J. Atmos. Sci.*, **48**, 2025–2043.

Zhao, Q. Y. and F. H. Carr, 1997: A prognostic cloud scheme for operational nwp models. *Mon. Wea. Rev.*, **125**, 1931–1953.

Primarily responsible for content: R. Bleck