Geostrophic Adjustment and the Finite-Difference Shallow-Water Equations

DAVID A. RANDALL

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

13 September 1993 and 24 November 1993

ABSTRACT

Numerical simulation of geostrophic adjustment in shallow water is discussed for the case of an unstaggered grid for vorticity, divergence, and mass. The dispersion equation is shown to be very well behaved and superior to that obtained with the Arakawa grids A–E.

1. Introduction

Winninghoff (1968) and Arakawa and Lamb (1977, hereafter AL) discussed the extent to which finite-difference approximations to the shallow-water equations can simulate the process of geostrophic adjustment, in which the dispersion of inertia–gravity waves leads to the establishment of a geostrophic balance, as the energy density of the inertia–gravity waves decreases with time due to their dispersive phase speeds and nonzero group velocity. These authors considered the momentum and mass conservation equations and defined five different staggered grids for the velocity components and mass.

The purpose of this paper is to point out that if the momentum equations are abandoned in favor of the vorticity and divergence equations, a simple unstaggered grid (vorticity, divergence, and mass all defined at the same points) gives very satisfactory approximations to the group and phase speeds of the continuous linearized shallow-water equations.

2. Equations, variables, and grids

Arakawa and Lamb considered the shallow-water equations linearized about a resting basic state in the following form:

$$\frac{\partial u}{\partial t} - fv + g\frac{\partial h}{\partial x} = 0 \tag{1}$$

$$\frac{\partial v}{\partial t} + fu + g\frac{\partial h}{\partial y} = 0 \tag{2}$$

$$\frac{\partial h}{\partial t} + H\delta = 0. \tag{3}$$

Here H is the constant depth of the "water" in the basic state, δ is the divergence, and all other symbols have their conventional meanings. From (1)–(3), we can derive an equivalent set in terms of vorticity ζ and divergence:

$$\frac{\partial \delta}{\partial t} - f\zeta + g\left(\frac{\partial^2}{\partial x^2}h + \frac{\partial^2}{\partial y^2}h\right) = 0, \tag{4}$$

$$\frac{\partial \zeta}{\partial t} + f\delta = 0, \tag{5}$$

$$\frac{\partial h}{\partial t} + H\delta = 0. \tag{6}$$

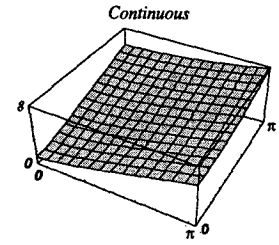
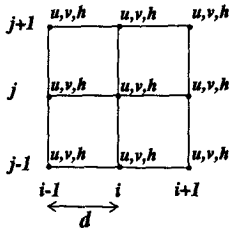
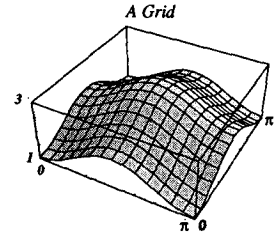
Of course, (6) is identical to (3). We can eliminate the vorticity and mass in (4) by using (5) and (6), respectively. Then by assuming wave solutions, we obtain the dispersion relation:

$$\left(\frac{\sigma}{f}\right)^2 = 1 + \lambda^2(k^2 + l^2). \tag{7}$$

Here σ is the frequency, $\lambda \equiv (gH)^{1/2}/f$ is the radius of deformation, and k and l are the wavenumbers in the x and y directions, respectively. The frequency and phase speed increase monotonically with wavenumber and are nonzero for all wavenumbers. The group speed is also nonzero for all wavenumbers. As discussed by AL, these characteristics of (7) are important for the geostrophic adjustment process.

In their discussion of various numerical representations of (1)–(3), AL defined five grids denoted by "A" through "E," as shown in Fig. 1. Arakawa and Lamb also gave the simplest centered finite-difference approximations to (1)–(3) for each of the five grids; these equations will not be repeated here. The two-dimensional dispersion equations for the various schemes were derived but not published by AL; they are included in Fig. 1. The table also gives a plot of the non-

Corresponding author address: Dr. David A. Randall, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523.

Representation	Dispersion Equation	Plot of Dispersion Equation
Continuous	$\left(\frac{\sigma}{f}\right)^2 = 1 + \left(\frac{\lambda}{d}\right)^2 [(kd)^2 + (ld)^2]$	
A 	$\left(\frac{\sigma}{f}\right)^2 = 1 + \left(\frac{\lambda}{d}\right)^2 [\sin^2(kd) + \sin^2(ld)]$	

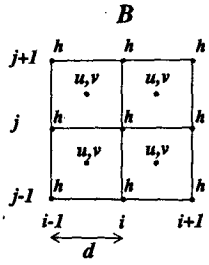
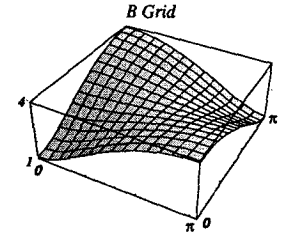
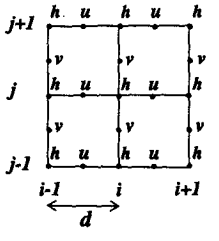
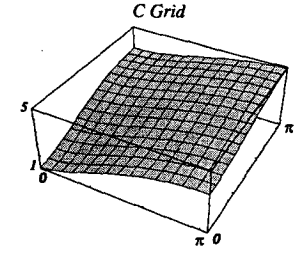
Representation	Dispersion Equation	Plot of Dispersion Equation
B 	$\left(\frac{\sigma}{f}\right)^2 = 1 + 2\left(\frac{\lambda}{d}\right)^2 [1 - \cos(kd)\cos(ld)]$	
C 	$\left(\frac{\sigma}{f}\right)^2 =$ $\frac{1}{4}[1 + \cos(kd) + \cos(ld) + \cos(kd)\cos(ld)]$ $+ 4\left(\frac{\lambda}{d}\right)^2 \left[\sin^2\left(\frac{kd}{2}\right) + \sin^2\left(\frac{ld}{2}\right)\right]$	

FIG. 1. Summary of representation (continuous or finite-difference grid), dispersion equation, and numerical results with $\lambda/d = 2$ for various grids. The horizontal coordinates in the plots are kd and ld, respectively, except for the E grid, for which kd^* and ld^* are used. The vertical coordinate is the normalized frequency, σ/f. For the E grid, the results are meaningful only in the triangular region for which $kd^* + ld^* \leq 2\pi$.

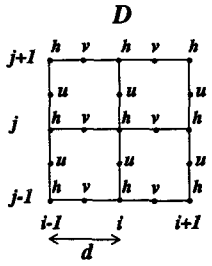
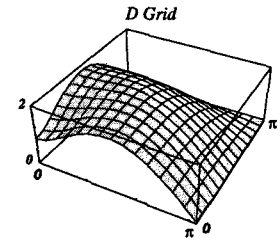
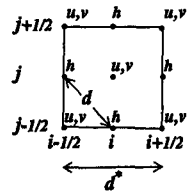
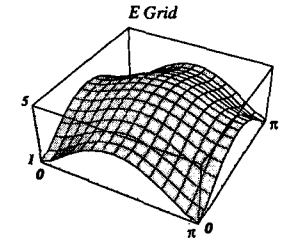
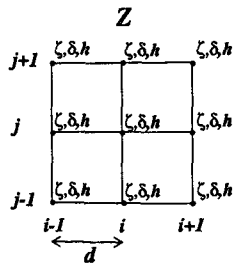
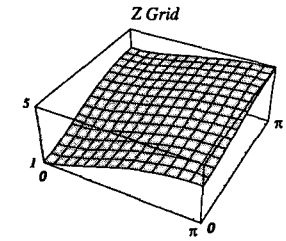
Representation	Dispersion Equation	Plot of Dispersion Equation
D 	$\left(\frac{\sigma}{f}\right)^2 = \frac{1}{4}[1+\cos(kd)+\cos(ld)+\cos(kd)\cos(ld)] + \left(\frac{\lambda}{d}\right)^2\left[\cos^2\left(\frac{kd}{2}\right)\sin^2(ld)+\cos^2\left(\frac{ld}{2}\right)\sin^2(kd)\right]$	
E 	$\left(\frac{\sigma}{f}\right)^2 = 1+4\left(\frac{\lambda}{d^*}\right)^2\left[\sin^2\left(\frac{kd^*}{2}\right)+\sin^2\left(\frac{ld^*}{2}\right)\right]$	
Z 	$\left(\frac{\sigma}{f}\right)^2 = 1+4\left(\frac{\lambda}{d}\right)^2\left[\sin^2\left(\frac{kd}{2}\right)+\sin^2\left(\frac{ld}{2}\right)\right]$	

FIG. 1. (*Continued*)

dimensional frequency (σ/f) as a function of kd and ld for the special case $\lambda/d = 2$. Here d is the grid size, assumed to be the same in the x and y directions. The significance of this particular choice of λ/d is discussed herein. The plots show how the nondimensional frequency varies out to $kd = \pi$ and $ld = \pi$; these wavenumbers correspond to the shortest waves that can be represented on the grid.

From a naive perspective, the A grid is arguably the simplest, since it is unstaggered. For example, the Coriolis terms of the momentum equations are easily evaluated, since u and v are defined at the same points. Approximation of the spatial derivatives in (1)–(3), however, inevitably involves averaging on the A grid. To illustrate this important point, consider the simplest centered approximation to $\partial h/\partial x$ at a u point on the A grid. We must first obtain a value of h at $i + 1/2$ by averaging from i and $i + 1$, and a second value at $i - 1/2$ by averaging from i and $i - 1$. We can then subtract these two average values of i and divide by Δx to obtain the desired approximation to $\partial h/\partial x$. Similarly, averaging is needed to define the mass convergence/divergence at h points.

The averaging described above inevitably "hides" noise at the smallest represented scales (e.g., a checkerboard pattern in h). Such dynamically "invisible" noise cannot participate in the dynamics of the model—for example, by propagating and dispersing as in the process of geostrophic adjustment. A plot of the dispersion equation for the A grid, as shown in Fig. 1, indicates a maximum of the frequency (group speed equal to zero) for some combinations k and l. As a

result, solutions on the A grid are extremely noisy in practice and must be smoothed—for example, through filtering (e.g., Kalnay-Rivas et al. 1977). Because of this well-known problem, the A grid is hardly used today.

The simple perspective of the preceding discussion suggests that it is desirable to avoid averaging in the design of a finite-difference scheme.

Next, consider the B grid. As on the A grid, the Coriolis terms are easily evaluated, without averaging, since u and v are defined at the same points. On the other hand, the pressure-gradient terms must be averaged, again as on the A grid. However, there is an important difference. On the A grid, the averaging used to approximate the x component of the pressure gradient force, $\partial h/\partial x$, is averaging *in the x direction.* On the B grid, the corresponding averages are *in the y direction.* On the B grid, an oscillation in the x direction, on the smallest represented scale, is not averaged out in the computation of $\partial h/\partial x$; it can, therefore, participate in the model's dynamics and so is subject to geostrophic adjustment. A similar conclusion holds for the convergence–divergence terms of the continuity equation. For example, the averaging in the y direction does no harm for solutions that are uniform in the y direction. Nevertheless, it does do some harm, as is apparent in the plot of the B-grid dispersion equation, as shown in Fig. 1. The frequency does not increase monotonically with total wavenumber; for certain combinations of k and l, the group speed is zero. Arakawa and Lamb concluded that the B grid gives a fairly good simulation of geostrophic adjustment but with some tendency to small-scale noise.

Now consider the C grid. The pressure gradient terms are easily evaluated, without averaging, because h is defined east and west of u points, and north and south of v points. Similarly, the mass convergence–divergence terms of the continuity equation can be evaluated without averaging the winds. On the other hand, averaging *is* needed to obtain the Coriolis terms, since u and v are defined at different points. For very small-scale inertia–gravity waves, the Coriolis terms are negligible; we essentially have pure gravity waves. This suggests that the C grid will perform well if the horizontal resolution of the model is high enough so that the smallest waves that can be represented on the grid are insensitive to the Coriolis force. More precisely, AL argued that the C grid does well when the grid size is small compared to λ, the radius of deformation. A plot of the dispersion equation, given in Table 1, shows that the frequency increases monotonically with wavenumber, as in the exact solution, although not as rapidly. Recall, however, that this plot is for the special case $\lambda/d = 2$. We will return to this point later.

Next, we turn to the D grid. Inspection of the stencil shown in Table 1 shows that the D grid allows a simple evaluation of the geostrophic wind. In view of the importance of geostrophic balance for large-scale motions, this may appear to be an attractive property. It is also apparent, however, that considerable averaging is needed in the pressure gradient force, mass convergence/divergence, and even in the Coriolis terms. As a result, the dispersion equation for the D grid is very badly behaved, giving zero phase speed for the shortest represented waves, and also giving a zero group speed for some modes.

Finally, the E grid is shown in Fig. 1. The E grid is essentially a rotated B grid and is sometimes called the "B/E grid" (Janjic 1984). The grid spacing for the E grid is chosen to be $d^* \equiv \sqrt{2}d$, so that the "density" of h points is the same as in the other four grids. The E grid at first seems perfect; no averaging is needed for the Coriolis terms, the pressure-gradient terms, or the mass convergence–divergence terms. Nevertheless, there is a problem, which becomes apparent if we consider a solution that is uniform in one of the grid directions—say, the y direction. In that case, we effectively have a one-dimensional problem. In one dimension, the E grid "collapses" to the A grid, with a reduced grid spacing d $(=d^*/\sqrt{2})$. For such one-dimensional motions, the E grid has all the problems of the A grid. These problems are apparent in the plot of the dispersion equation, given in Fig. 1. (For the E grid, the nondimensional frequency is plotted as a function of kd^* and ld^*, out to a value of 2π; this corresponds to the shortest "one-dimensional" mode.) The group speed is zero for some combinations of k and l.

3. The Z grid

Now recall the conclusion of AL, described earlier, that the C grid gives a good simulation of geostrophic adjustment provided that $\lambda/d > 1$. Large-scale modelers are never happy to choose d so that λ/d can be less than one. Nevertheless, in practice, modes for which $\lambda/d \ll 1$ can be unavoidable, at least for some situations. For example, Hansen et al. (1983) described a low-resolution atmospheric GCM, which they called Model II, designed for very long climate simulations in which low resolution was a necessity. Model II used a grid size of 8° latitude × 10° longitude; this means that the grid size was larger than the radius of deformation for many of the physically important modes that could be represented on the grid. As shown by AL, such modes cannot be well simulated using the C grid. Having experienced these problems with the C grid, Hansen et al. (1983) chose the B grid for Model II.

Ocean models must contend with small radii of deformation, so that very fine grids are needed to ensure that $\lambda/d > 1$, even for external modes. For this reason, ocean models tend to use the B grid (e.g., Semtner and Chervin 1992).

In addition, three-dimensional models of the atmosphere and ocean generate internal modes. With vertical structures typical of current general circulation

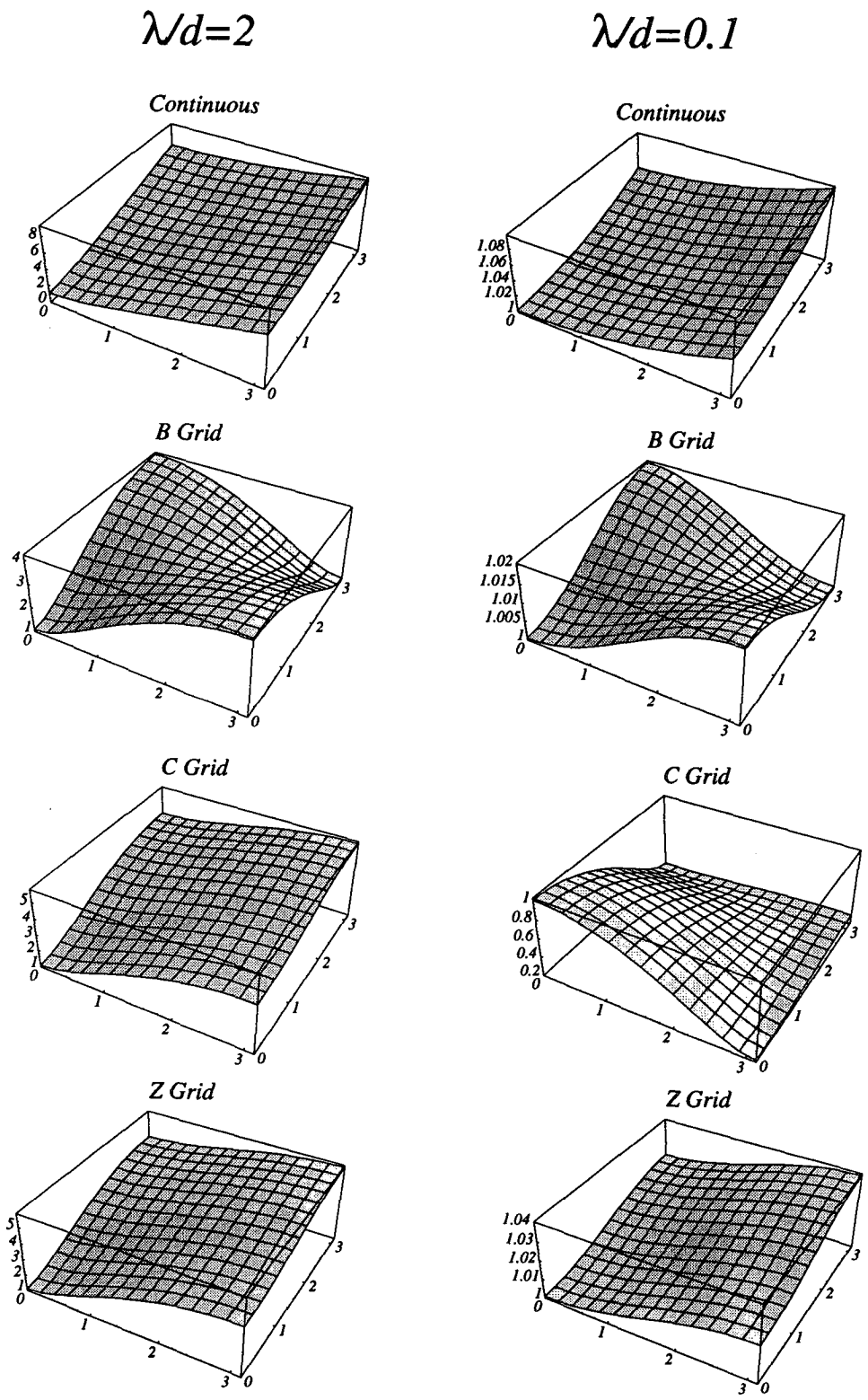


FIG. 2. Dispersion relations for the continuous shallow-water equations and for finite-difference approximations based on the B, C, and Z grids. The horizontal coordinates in the plots are kd and ld, respectively. The vertical coordinate is the normalized frequency, σ/f. The left column shows results for $\lambda/d = 2$, and the right for $\lambda/d = 0.1$.

models, the highest internal modes can have radii of deformation on the order of 50 km or less. The same model may have a horizontal grid spacing on the order of 500 km, so that λ/d can be on the order of 0.1. Figure 2 demonstrates that the C grid behaves very badly for $\lambda/d = 0.1$. The phase speed actually decreases monotonically as the wavenumber increases and becomes very small for the shortest waves that can be represented on the grid. Janjic and Mesinger (1989) have emphasized that, as a result, models that use the C grid have difficulty in representing the geostrophic adjustment of high internal modes. In contrast, the dispersion relation for the B grid is qualitatively insensitive to the value of λ/d. The B grid has moderate problems for $\lambda/d = 2$, but these problems do not become significantly worse for $\lambda/d = 0.1$.

In summary, the C grid does well with deep, external modes but has serious problems with high internal modes, whereas the B grid has moderate problems with all modes.

Now consider an unstaggered grid for the integration of (4)–(6). This grid is also illustrated in Fig. 1; we call it the Z grid.[1] Inspection shows that with the Z grid the components of the divergent part of the wind "want" to be staggered as in the C grid, while the components of the rotational part of the wind "want" to be staggered as in the D grid. This means that the Z grid does not correspond to any of the grids A–E.

No averaging is required with the Z grid. The only spatial differential operator appearing in (4)–(6) is the Laplacian ∇^2, which is applied to h in the divergence equation. With the usual centered finite-difference stencils, the finite-difference approximation to $\nabla^2 h$ is defined at the same point as h itself. An unstaggered grid is thus a natural choice for the numerical integration of (4)–(6).

Figure 2 shows that the dispersion relation for the Z grid is very close to that of the C grid for $\lambda/d = 2$, but is drastically different for $\lambda/d = 0.1$. Whereas the C grid behaves very badly for $\lambda/d = 0.1$, the dispersion relation obtained with the Z grid is qualitatively insensitive to the value of λ/d; it resembles the dispersion relation for the continuous equations, in that the phase speed increases monotonically with wavenumber and the group speed is nonzero for all wavenumbers. Since the Z grid is unstaggered, collapsing it to one dimension has no effect.

4. Conclusions

Our results demonstrate that geostrophic adjustment in shallow water is well simulated on an unstaggered grid when the vorticity and divergence equations are used and that in fact such a scheme simulates geostrophic adjustment better than any of the staggered schemes for the primitive equations. The good behavior of such schemes was recognized by Staniforth and Mitchell (1977) and Williams (1981); see also Neta and Williams (1989), who wrote that "The best finite-element schemes use either staggered elements in the primitive form of the equations, or unstaggered elements in the vorticity–divergence form."

The vorticity and divergence equations are routinely used in global spectral models but are rarely used in global finite-difference models. The reason seems to be that it is necessary to solve elliptic equations to obtain the winds from the vorticity and divergence—for example, to evaluate the advection terms of the nonlinear primitive equations. Such solution procedures can be computationally expensive in finite-difference models (but are not expensive in spectral models), although it may be appropriate to reexamine this point in the light of modern algorithms for solving linear systems (e.g., multigrid methods).

As spectral modelers have recognized, there are some advantages to integrating the vorticity and divergence equations. Conservation principles for potential vorticity and potential enstrophy are easily implemented. Gravity waves can readily be controlled by damping the divergence. More fundamentally, the divergence and vorticity are true scalars, unlike the horizontal velocity components in any particular coordinate system.

Our results encourage us to consider evaluation of a global finite-difference primitive equation model that is based on the vorticity and divergence equations rather than on the momentum equations, and that uses the Z grid. Of course, such models are nonlinear, so that well-behaved discrete approximations for the nonlinear terms of the governing equations are crucial for success. The present paper has discussed only some of the linear aspects of the discretization problem.

Acknowledgments. Support for this research has been provided by the U.S. Department of Energy's CHAMMP Program under Grant DE-FG02-91ER-61218 to Colorado State University. Thanks to Profs. A. Arakawa of the University of California and W. Schubert of Colorado State University for helpful comments.

REFERENCES

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

Hansen, J., G. Russell, D. Rind, P. Stone, A. Lacis, S. Lebedeff, R. Ruedy, and L. Travis, 1983: Efficient three-dimensional global models for climate studies: Models I and II. *Mon. Wea. Rev.*, **111,** 609–662.

Janjic, Z. I., 1984: Nonlinear advection schemes and energy cascade on semi-staggered grids. *Mon. Wea. Rev.*, **112,** 1234–1245.

——, and F. Mesinger, 1989: Response to small-scale forcing on two staggered grids used in finite-difference models of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **115,** 1167–1176.

[1] Modesty precludes us from calling it the R grid.

Kalnay-Rivas, E., A. Bayliss, and J. Storch, 1977: The 4th order GISS model of the global atmosphere. *Contrib. Atmos. Phys.*, **50,** 306–311.

Neta, B., and R. T. Williams, 1989: Rossby wave frequencies and group velocities for finite element and finite difference approximations to the vorticity–divergence and primitive forms of the shallow water equations. *Mon. Wea. Rev.*, **117,** 1439–1457.

Semtner, A. J., Jr., and R. M. Chervin, 1992: Ocean general circulation from a global eddy-resolving model. *J. Geophys. Res.*, **97C,** 5493–5551.

Staniforth, A. N., and H. L. Mitchell, 1977: A semi-implicit finite-element barotropic model. *Mon. Wea. Rev.*, **105,** 154–169.

Williams, R. T., 1981: On the formulation of finite-element prediction models. *Mon. Wea. Rev.*, **109,** 463–466.

Winninghoff, F. J., 1968: On the adjustment toward a geostrophic balance in a simple primitive equation model with application to the problems of initialization and objective analysis. Ph.D. thesis, University of California, Los Angeles.