A Proposal for the Intercomparison of the Dynamical Cores of Atmospheric General Circulation Models

Isaac M. Held*
and Max J. Suarez**

Abstract

A benchmark calculation is proposed for evaluating the dynamical cores of atmospheric general circulation models independently of the physical parameterizations. The test focuses on the long-term statistical properties of a fully developed general circulation; thus, it is particularly appropriate for intercomparing the dynamics used in climate models. To illustrate the use of this benchmark, two very different atmospheric dynamical cores—one spectral, one finite difference—are compared. It is found that the long-term statistics produced by the two models are very similar. Selected results from these calculations are presented to initiate the intercomparison.

1. Introduction

The careful evaluation and comparison of atmospheric general circulation models (GCMs) is essential if we are to move methodically toward improved climate models. A comparison of the climatic simulations produced by various GCMs using realistic boundary conditions is currently being conducted by the Atmospheric Model Intercomparison Project (Gates 1992). While evaluation of full GCMs is essential, the difficulty of interpreting such intercomparisons is well appreciated by the modeling community. The number of choices that must be made about poorly understood parts of the system—particularly in the parameterization of subgrid-scale processes—is large enough that it is difficult to exhaustively compare even closely related models. Furthermore, important differences often arise from the details of the implementation or from the way in which different parameterizations interact. By the same token, errors in the simulations, as identified by comparison with observations, are difficult to attribute to specific modeling assumptions. Indeed, it is possible for models to produce realistic simulations of some climatic features for the "wrong" reasons, or for the simulation of some features to worsen when other parts of the model are improved.

A logical reaction to these difficulties is to try to think of GCMs as being decomposed into modules that can be tested and compared independently. A good example of this approach is provided by the Intercomparison of Radiation Codes for Climate Models (ICRCCM) project (Ellingson et al. 1991; Fouquart et al. 1991), in which the focus is on testing the broadband parameterizations used in GCMs against the more precise, but expensive, line-by-line calculations. A similar effort is under way comparing land surface parameterizations (Henderson-Sellers et al. 1992). We feel that narrow, well-controlled comparisons like these are the most productive way to proceed and that they should be expanded to all aspects of GCMs, including the dynamics. In this article, we are proposing a way of carrying out such a comparison for what we refer to as the "dynamical core" of the model.

All GCMs solve discrete forms of the equations of motion for the time evolution of the three-dimensional flow field and of the thermodynamic state of the air. There is a need for a careful reexamination of the standard numerical methods used to discretize these equations and, given the advent of parallel computer architectures, for the evaluation of novel approaches.

Williamson et al. (1992) propose a series of tests using the nonlinear shallow water equations on the sphere as surrogates for the three-dimensional equations of motion used in GCMs. They provide analytical or high-resolution solutions to several initial value problems, aiming, for example, at testing the ability of a given numerical scheme to simulate the propagation of a planetary wave without loss of amplitude or to transport a tracer across the pole. While shallow water solutions provide valuable indications of the performance of horizontal discretization schemes, there are two important limitations to this approach: 1) three-dimensional atmospheric flows, in addition to introducing problems associated with vertical discretization, may place distinctive demands on the numerics; and

*Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey.
**Goddard Space Flight Center/NASA, Greenbelt, Maryland.
Corresponding author address: Dr. Isaac M. Held, Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, P.O. Box 308, Princeton, NJ 08542.
In final form 17 May 1994.

2) for climate models, it is important to directly evaluate the long-term statistics rather than focus on the accuracy of short-term, deterministic solutions. Tests like those of Williamson et al. are more applicable to numerical weather prediction models, for which the relevant trade-offs between accuracy, efficiency, and conservation properties may be quite different than for climate models.

We are attempting to define a set of benchmark calculations for the evaluation of statistically steady states produced by the atmospheric dynamical cores used in climate models. In these tests, we replace the detailed radiative, turbulence, and moist convective parameterizations with very simple forcing and dissipation. There are problems with this approach as well. The most obvious is that the true solution is unknown. We presume that different modeling approaches will converge to the true solution as resolution is increased, keeping in mind that if convergence is not clearly obtained in these idealized problems, it may not be easy to obtain in realistic GCMs either. Another problem is that sampling errors, due to very low frequency variability, may interfere with our ability to accurately define the statistically steady state, given the finite time intervals over which the models are integrated, but once again, this problem is also present in realistic climate simulations.

An additional motivation we have in proposing tests focusing on the dynamics is to nudge the modeling community toward the creation of modular dynamical cores that are easy to interchange. The importance of using modular, or "plug-compatible," codes was discussed by Kalnay et al. (1989) for the physical parameterizations. We would like to see these ideas extended, as much as possible, to the design of dynamical cores. In the spirit of having a free exchange of all codes used in climate models, we are making publicly available the Fortran codes for the dynamical cores used to produce the results described below.

The first in our proposed series of benchmark calculations is described in this report. As an example, the benchmark is used to compare two GCM dynamical cores, one spectral and one finite difference. Both are closely related to codes that are being used for climate studies at our laboratories. We present here only a sampling of the two models' climate statistics, but we can make available much more complete results to anyone interested in making a more detailed comparison. In a separate study, we will be looking at this proposed calculation in more depth. In particular, we will study the behavior of these two models as a function of horizontal and vertical resolution and discuss their sensitivity to the choice of dissipation.

2. A first benchmark calculation

In designing the forcing and dissipation, we use simple Newtonian relaxation of the temperature field to a zonally symmetric state and Rayleigh damping of low-level winds to represent boundary-layer friction. Forcing GCMs in this way is relatively common, especially in two-layer models [Hendon and Hartmann (1985) and Suarez and Duffy (1992) are two ex-

$$\frac{\partial \mathbf{v}}{\partial t} = \cdots - k_v(\sigma)\mathbf{v}$$

$$\frac{\partial T}{\partial t} = \cdots - k_T(\phi,\sigma)\left[T - T_{eq}(\phi,p)\right]$$

$$T_{eq} = \max\left\{200\text{K}, \left[315\text{K} - (\Delta T)_y \sin^2\phi - (\Delta\theta)_z \log\left(\frac{p}{p_0}\right)\cos^2\phi\right]\left(\frac{p}{p_0}\right)^{\kappa}\right\}$$

$$k_T = k_a + (k_s - k_a)\max\left(0, \frac{\sigma - \sigma_b}{1 - \sigma_b}\right)\cos^4\phi$$

$$k_v = k_f \max\left(0, \frac{\sigma - \sigma_b}{1 - \sigma_b}\right)$$

$\sigma_b = 0.7$ $\quad k_f = 1 \text{ day}^{-1}$,

$k_a = 1/40 \text{ day}^{-1}$ $\quad k_s = 1/4 \text{ day}^{-1}$

$(\Delta T)_y = 60\text{K}$ $\quad (\Delta\theta)_z = 10\text{K}$

$p_0 = 1000 \text{ mb}$ $\quad \kappa = \frac{R}{c_p} = \frac{2}{7}$ $\quad c_p = 1004 \text{ J kg}^{-1}\text{ K}^{-1}$

$\Omega = 7.292 \times 10^{-5} \text{ s}^{-1}$ $\quad g = 9.8 \text{ m s}^{-2}$ $\quad a_e = 6.371 \times 10^6 \text{ m}$.

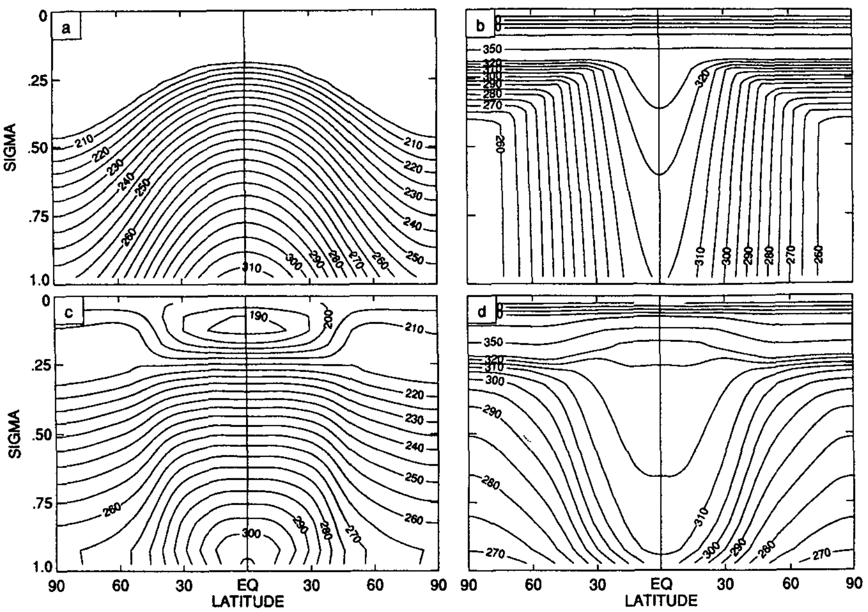

FIG. 1. The upper panels contain the prescribed radiative equilibrium temperature (a) and potential temperature (b) distributions. The lower panels contain 1000-day averages of the zonal mean temperature (c) and potential temperature (d) distributions produced by the G72 gridpoint model.

amples on the sphere]. Recent examples of its use in models with more vertical resolution are James and James (1989) and Yu and Hartmann (1993). Our specifications are detailed in the box on the opposing page.

We start with an ideal gas atmosphere over a rotating spherical surface. There is no topography, in the sense that the surface is at constant geopotential. Nothing is said as to whether the flow is or is not hydrostatic. While most global models assume hydrostatic balance, this is not considered part of the specification of the problem but rather a modeling choice. The choice of upper boundary condition is also left open. The inclusion of a rigid lid at some height or pressure, for example, is again considered a modeling choice. Aside from the forcing parameters, we need to specify only the gas constant R, the specific heat of air at constant pressure c_p, the acceleration of gravity g, the radius of the sphere a_e, and the total mass of the atmosphere p_0/g. The acceleration of gravity is needed only in a nonhydrostatic model.

The only specified dissipation is a simple linear damping of the velocities. The strength of the damping k_v is a function of $\sigma = p/p_s$, where p is the pressure and p_s is the instantaneous surface pressure. We use σ rather than pressure in this expression so that this "boundary layer" will follow the topography in future calculations in this series. This damping is nonzero only in layers near the surface ($\sigma > 0.7$).

Temperatures are relaxed to a prescribed "radiative equilibrium" T_{eq}, which is a function of latitude and pressure. These temperatures and the corresponding potential temperatures are shown in Figs. 1a,b. This radiative equilibrium is given some positive static stability, relatively large in the Tropics and decreasing to zero at the poles. One can think of this tropical static stability as taking into account the effects of moist convection, but this is potentially misleading, and it is better to think of it as simply an artifact that helps minimize the occurrence of gravitational instability. The radiative relaxation time is also a function of latitude and σ. If one uses a long relaxation time everywhere, an unrealistic thin cold layer develops near the surface, particularly in the Tropics. The relaxation time is increased in this region to reduce this effect, but some vestige of this shallow stable layer still remains in the solutions. This is clearly visible in the time-mean temperature and potential temperature distribution (Figs. 1c,d) produced by the "G72" finite-difference model (described below). The potential temperatures also show that the dynamics is maintaining the static stability well above its radiative equilibrium value in the extratropics.

No explicit diffusion is included in the specification of the model. We prefer to think of subgrid-scale diffusivity as part of the numerical scheme. (For the sake of well-posedness, one can think of molecular viscosity and diffusion as being present, but negligible on all scales that we can hope to resolve in a global climate model.) This allows us to treat, evenhandedly, conservative schemes that require explicit subgrid-scale mixing, and schemes that are dissipative by design. By specifying a large diffusivity that produces smooth large-scale solutions, one would be creating a very different problem from the effectively infinite Reynolds number problem that we have posed. Our goal is not simply the accurate simulation of the evolution of smooth flows but the optimum treatment of flows that generate motions on all resolved scales through turbulent cascades, as does the atmosphere.

In the two models described below, we have included only a very scale-selective horizontal mixing and have omitted vertical mixing (diffusion or convective adjustment) altogether. Although gravitationally unstable regions do form on occasion, particularly in

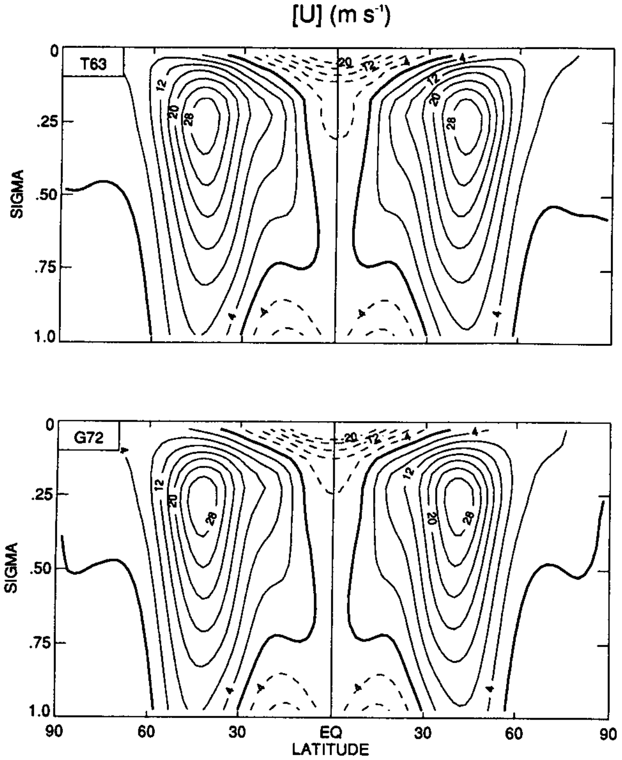

FIG. 2. The zonal-mean zonal wind produced by the T63 spectral model and G72 gridpoint model. Both are 1000-day means. Since the forcing is symmetric about the equator, differences between the hemispheres are indicative of sampling errors.

low latitudes, the models can be integrated stably without enhanced vertical mixing in such regions.

3. A sample intercomparison

To begin the process of developing standards for dynamical cores of atmospheric climate models, we performed long-term integrations with our own models. The models use very different discretization methods and were coded independently by the two authors.

a. The models

The spectral model is a standard hydrostatic, σ-coordinate, semi-implicit, spectral transform model, in the vorticity-divergence form described by Bourke (1974). The transform grid is chosen to ensure alias-free computation of quadratic products, in the usual way. The vertical differencing uses the simplest centered differences, and the hydrostatic equation is integrated analytically assuming that temperature is constant within each layer. This differencing is not energy conserving. There are 20 vertical levels, equally spaced in sigma, with the top of the model formally at zero pressure. A leapfrog scheme is used for time stepping, using the time filter described by Robert (1966) to control the computational mode. The horizontal mixing of vorticity, divergence, and temperature takes the form of a Laplacian raised to the fourth power, with the strength set so that the e-folding time for the smallest wave in the system is always 0.1 days. The truncation is triangular. We made integrations at four resolutions: T21, T30, T42, and T63, where the numeral refers to the maximum number of zonal waves present. In this note, we present results only from T63. The code for this model is available from Isaac Held (e-mail: ih@gfdl.gov).

The finite-difference model is also hydrostatic and uses σ coordinates, with 20 layers equally spaced in sigma. The vertical differencing is that proposed by Arakawa and Suarez (1983). A latitude–longitude grid with Arakawa's C-grid staggering is used for the horizontal discretization. The horizontal finite-differencing scheme is second order in all respects except the horizontal advection of vorticity, which is fourth-order accurate for nondivergent flow, reducing to the fourth-order Arakawa (1966) Jacobian in this case. A Fourier filter is applied to all tendencies to damp short zonal scales poleward of 45° latitude. An explicit leapfrog time step is used, with the pressure gradient averaging suggested by Brown and Campana (1978). The computational mode of the leapfrog is controlled in exactly the same way as in the spectral model. An eighth-order Shapiro (1970) filter controls gridpoint noise. The filter damps the 2Δ wave with an effective e-folding time of 1.5 h. As with the spectral model, we have made integrations at four resolutions: 6° latitude × 7.5° longitude, 4° × 5°, 3° × 3.75°, and 2° × 2.5°. We will show results only for the highest resolution, which we label G72, the numeral referring to half of the number of grid points around a latitude circle. The finite-difference dynamical core is described in Suarez and Takacs (1995). That report and the code are available from Max Suarez (e-mail: suarez@nino.gsfc.nasa.gov).

b. Results

All cases presented were integrated 1200 days. The model's climate in each case is obtained by averaging the last 1000 days of integration. The integration of the spectral model was started from an isothermal state at rest, with some small perturbations added to break the symmetry. The gridpoint integration was started from an earlier run of a lower-resolution model that had already equilibrated to the forcing. By discarding 200 days, we are reasonably certain that we eliminate significant differences due to the different initialization schemes.

Aspects of the climates produced by the two models are displayed in Figs. 2–4: the zonal-mean zonal

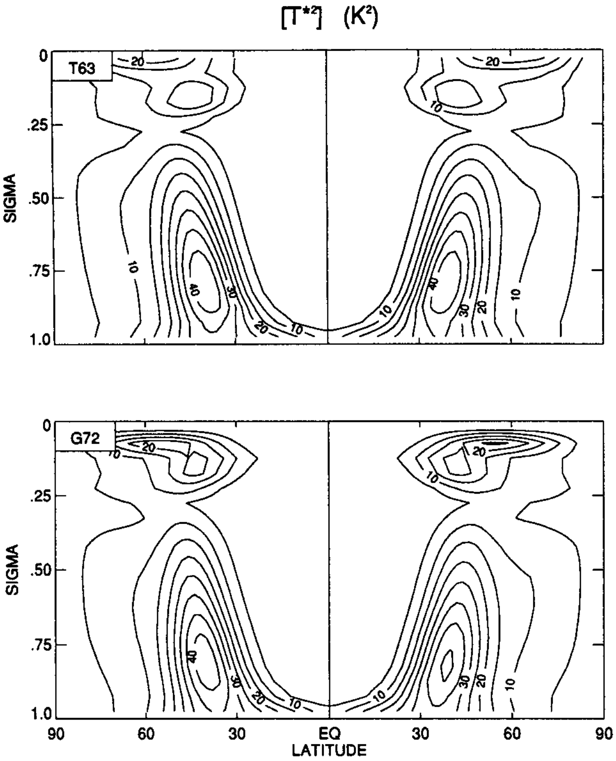

FIG. 3. As in Fig. 2 but for the eddy variance of the temperature.

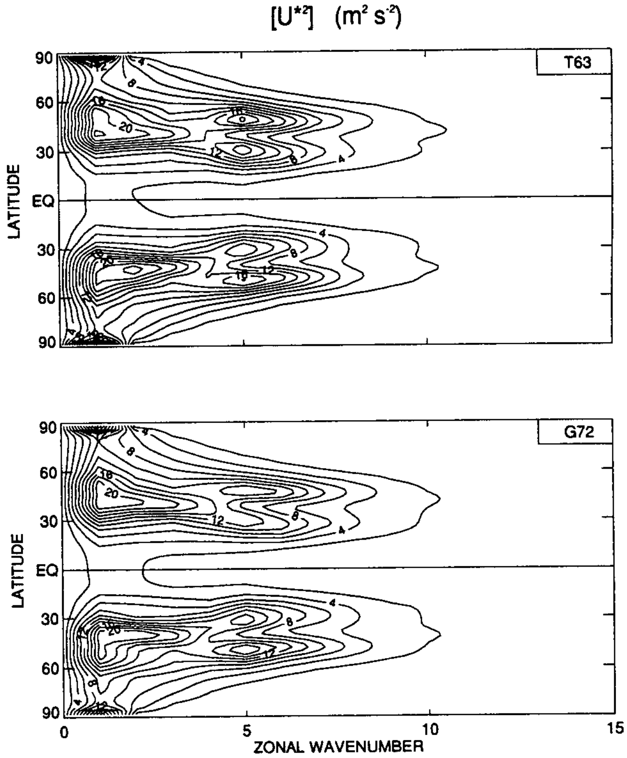

FIG. 4. As in Fig. 2 but for the vertically averaged zonal spectra of the eddy variance of zonal wind.

wind (Figs. 2) as a function of latitude and sigma; the eddy temperature variance (Fig. 3), also as a function of latitude and sigma; and the zonal spectrum of the eddy zonal wind (Fig. 4), as a function of zonal wavenumber and latitude. Since the forcing is symmetric about the equator, we can use the difference between the climates of the two hemispheres for a quick estimate of sampling error in the 1000-day means.

We see from these figures that our simple forcing and dissipation produce a reasonably realistic zonal-mean circulation. A single jet is generated with maximum strength of roughly 30 m s^{-1} near 45° latitude. The jet closes off, with a well-defined reversed shear above $\sigma = 0.2$. The surface westerlies reach almost 8 m s^{-1} near 45° latitude. There are equatorial easterlies in the model stratosphere. The eddy temperature variance shows two midlatitude maxima, one in the lower troposphere and another above the tropopause. An unrealistic feature of the results is the penetration of the temperature variance near the surface well into the Tropics. The spectrum of the eddy zonal winds has two peaks at zonal wavenumber 5 on the flanks of the storm track (with the peak on the polar side somewhat stronger), a third peak at wavenumbers 1 and 2 at the center of the track, and yet another peak, associated with cross-polar flow, in wavenumber 1 at the pole.

These figures indicate an impressive degree of agreement between the two models, despite their very different numerical algorithms. There are large differences in the eddy temperature variance in the stratosphere, which we attribute to the different vertical discretizations and the coarse stratospheric resolution. Another difference, not evident in the figures shown, is that the grid model is noisier than the spectral model near the pole. After completing these integrations, we found this noise could be significantly reduced but not eliminated by a simple modification of the treatment of the momentum equations at the pole. This change has no noticeable effect on other statistics. Details of the finite differencing at the poles are given in Suarez and Takacs (1993).

The agreement between these particular spectral and gridpoint models should not be interpreted as implying that the true solution to our problem has been obtained. Additional calculations, described elsewhere, show that the climates of both models are sensitive to resolution and have not yet converged at the resolution presented here. The solutions may also be more sensitive to the choice of parameters and to the model formulation than the agreement between these two cores suggests. As an example, the same calculations repeated with a second-order version of the finite-difference dynamical core produce a general circulation quite different from the one presented here: the jets are narrower and shifted equatorward, and the surface winds and eddy momentum transports are considerably weaker. Thus, it may be that the current parameters are near a transition in the flow regime. (Hints of the existence of such a transition are also seen when the surface friction is varied.) If the flow regime is sensitive to the model parameters, this test can be a very sensitive indicator of the quality of the model's numerics, although it may be necessary to study the solution as a function of a parameter to understand the differences between models.

4. Future directions

We are interested in the development of tests or standards for the intercomparison of the dynamical cores of atmospheric climate models and have decided that a test should satisfy two conditions to be useful: it should isolate the dynamics as much as possible from the complexity of the physical parameterizations, particularly those for moist convection, clouds, and the planetary boundary layer, and it should evaluate the long-term statistical properties of a realistic three-dimensional global circulation. We feel the benchmark calculation we have proposed satisfies these requirements and will be a good starting point for such an intercomparison. We do recognize that there are deficiencies in the current formulation. For example, as vertical resolution is increased, some vertical mixing may be required, and it is not clear what lower boundary condition is most appropriate given the current drag formulation. Despite these misgivings, the circulation produced is sufficiently meteorological, and the solutions we have obtained with two very different models are sufficiently similar that we feel it provides a useful test. We hope other groups will join in this effort by repeating the calculation with their own models, experimenting with other simple forcing functions, and sharing their results.

The statistics presented here only touch the surface of interesting quantities for intercomparison. Several that interest us are measures of coherent baroclinic wave packet structure, the amplitude and spectral width of the westward propagating external mode resonances, and the gravity wave activity in the model stratosphere. The Lagrangian statistics of the flow are most easily probed by adding passive tracers, which can then be used to evaluate transport algorithms as well. It should also be of interest to test the model's treatment of the pole by rotating the coordinate system or the radiative equilibrium temperature so that the numerical pole resides in the storm track and by checking for asymmetries in the resulting climate. We have begun tests with a second benchmark calculation, in which an idealized midlatitude mountain is introduced. This calculation should be useful for comparing alternatives to sigma coordinates and for evaluating the sensitivity of low-frequency planetary wave activity to the choice of model.

Acknowledgments. MJS was supported in part by the Global Atmospheric Modeling and Analysis Program of the Office of Mission to Planet Earth/NASA Headquarters and by the Earth and Space Science Applications Project of the NASA HPCC Program.

References

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. *J. Comput. Phys.* **1**, 119–143.

——, and M. J. Suarez, 1983: Vertical differencing of the primitive equations in sigma coordinates. *Mon. Wea. Rev.*, **111**, 34–45.

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

Brown, J. A., and K. Campana, 1978: An economical time-differencing system for numerical weather prediction. *Mon. Wea. Rev.*, **106**, 1125–1136.

Ellingson, R. G., J. Ellis, and S. B. Fels, 1991: The intercomparison of radiation codes used in climate models: Longwave results. *J. Geophys. Res.*, **96**, 8929–8953.

Fouquart, Y. B., B. Bonnel, and V. Ramaswamy, 1991: Intercomparing shortwave radiation codes for climate studies. *J. Geophys. Res.*, **96**, 8955–8968.

Gates, W. L., 1992: The atmospheric model intercomparison project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1970.

Hendon, H. H., and D. L. Hartmann, 1985: Variability in a nonlinear model of the atmosphere with zonally symmetric forcing. *J. Atmos. Sci.*, **42**, 2783–2797.

Henderson-Sellers, A., Z.-L. Yang, and R. E. Dickinson, 1992: The project for intercomparison of land-surface parameterization schemes. *Bull. Amer. Meteor. Soc.*, **74**, 1335–1349.

James, I. N., and P. M. James, 1989: Ultra-low frequency variability in a simple atmospheric circulation model. *Nature*, **342**, 53–55.

Kalnay, E., M. Kanamitsu, J. Pfaendtner, J. Sela, M. Suarez, J. Stackpole, J. Tuccillo, L. Umscheid, and D. Williamson, 1989: Rules for the interchange of physical parameterizations. *Bull. Amer. Meteor. Soc.*, **70**, 620–622.

Robert, A. F., 1966: The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan*, **44**, 237–245.

Shapiro, R., 1970: Smoothing, filtering and boundary effects. *Rev. Geophys. Space Phys.*, **8**, 359–387.

Suarez, M. J., and D. G. Duffy, 1992: Terrestrial superrotation: A bifurcation of the general circulation. *J. Atmos. Sci.*, **49**, 1541–1554.

——, and L. T. Takacs, 1995: Documentation of the ARIES/GEOS dynamical core: Version 2. NASA Tech. Memo. 104606, 55 pp.

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

Yu, J.-Y., and D. L. Hartmann, 1993: Zonal flow vacillation and eddy forcing in a simple GCM of the atmosphere. *J. Atmos. Sci.*, **50**, 3244–3259.