A baroclinic instability test case for atmospheric model dynamical cores

By CHRISTIANE JABLONOWSKI* and DAVID L. WILLIAMSON
National Center for Atmospheric Research, Boulder, USA

SUMMARY

A deterministic initial-value test case for dry dynamical cores of atmospheric general-circulation models is presented that assesses the evolution of an idealized baroclinic wave in the northern hemisphere. The initial zonal state is quasi-realistic and completely defined by analytic expressions which are a steady-state solution of the adiabatic inviscid primitive equations with pressure-based vertical coordinates. A two-component test strategy first evaluates the ability of the discrete approximations to maintain the steady-state solution. Then an overlaid perturbation is introduced which triggers the growth of a baroclinic disturbance over the course of several days.

The test is applied to four very different dynamical cores at varying horizontal and vertical resolutions. In particular, the NASA/NCAR Finite Volume dynamics package, the National Center for Atmospheric Research spectral transform Eulerian and the semi-Lagrangian dynamical cores of the Community Atmosphere Model CAM3 are evaluated. In addition, the icosahedral finite-difference model GME of the German Weather Service is tested. These hydrostatic dynamical cores represent a broad range of numerical approaches and, at very high resolutions, provide independent reference solutions. The paper discusses the convergence-with-resolution characteristics of the schemes and evaluates the uncertainty of the high-resolution reference solutions.

KEYWORDS: Analytic initial conditions   Baroclinic wave   Ensemble uncertainty   GCM intercomparison   Reference solutions

1. INTRODUCTION

Tests of atmospheric general-circulation models (GCMs) and, in particular, tests of their dynamical cores are important steps towards future model improvements. They reveal the influence of an individual model design on climate and weather simulations and indicate whether the circulation is described representatively by the numerical approach. However, testing a global three-dimensional (3D) atmospheric model is not straightforward. In the absence of non-trivial analytic solutions, the model evaluations most commonly rely on intuition, experience and model intercomparisons like the Atmospheric Model Intercomparison Project (AMIP) (Gates 1995; Gates *et al.* 1999). In addition, GCM simulation statistics are typically compared to global re-analysis data while numerical weather forecasts are compared to local observations. Such approaches are not applicable to pure dynamical core assessments that isolate the dynamics processes. Note that the term dynamical core generally refers to the numerical approximation to the resolved fluid flow component of an atmospheric model and excludes the subgrid-scale so-called physics parametrizations. In general, three different sets of equations are most commonly used as the basis of dynamical cores. These are the hydrostatic primitive equations, the non-hydrostatic shallow-atmosphere, and non-hydrostatic deep-atmosphere equation sets.

Very few standard test cases for 3D dynamical cores on the sphere have been developed to date. Among them are the long-term idealized benchmarks as suggested by Held and Suarez (1994) and Boer and Denis (1997). They require 1200-day model integrations with simple, pre-defined forcing functions that replace the complex physics parametrizations. Typically, time-mean zonal-mean flow fields are evaluated for both test scenarios that assess the model's idealized climatic state and its variability. But as a consequence, the averaging of the flow patterns essentially eliminates all small-scale flow features in the analysis, although these are the main contributors to the creation of

* Corresponding author, current affiliation: University of Michigan, 2455 Hayward St., Ann Arbor, MI, 48109-2143, USA. e-mail: cjablono@umich.edu

the mean state. Improvements of the statistical mean state are therefore directly related to an improved representation of the fine-scale flow characteristics.

Standard techniques for short-term baroclinic dynamical core assessments can only rarely be found in the literature. This is in contrast to the long history of test cases available for shallow-water models (see Williamson *et al.* (1992) for a collection). One deterministic test technique was suggested by Polvani and Saravanan (2000) who simulated the breakdown of the polar vortex in a 3D hydrostatic dynamical core. The definition of this test case was inspired by earlier 1D and 2D studies of the vortex erosion problem. These were conducted by Juckes and McIntyre (1987) and Bates and Li (1997), respectively, who applied the 1D principles to the 2D shallow-water framework. In addition, Giraldo and Rosmond (2004) used a 3D version of the classical Rossby–Haurwitz test (Phillips 1959) for dynamical core assessments, but note that their description of the initial dataset is imprecise. The 3D extension of this Rossby–Haurwitz wave has first been shown in Monaco and Williams (1975). Other early test techniques from the 1970s, like the baroclinic-wave problem used in Hoskins and Simmons (1975), have not been widely applied by the community. Most often, the lack of a concise test specification make these tests difficult to use and interpret, especially in comparing new schemes to older results.

As modelling groups now move towards the next generation of dynamical cores with new numerical schemes and gridding options (Bonaventura 2004; Fournier *et al.* 2004; Giraldo and Rosmond 2004; Tomita and Sato 2004), a standard test suite for hydrostatic and non-hydrostatic dynamics packages is highly desirable. With the test described here we contribute to this effort. The new test is based on four design principles. In particular, it is deterministic, easy to use, relevant to typical atmospheric phenomena and applicable to a wide variety of model formulations. The test primarily targets dry dynamical cores on the sphere that are either based on the hydrostatic primitive equations or the non-hydrostatic equations with the shallow-atmosphere approximation. For these two types of dynamical cores the initial conditions are a balanced solution to the inviscid adiabatic equation sets. Furthermore, the application of the test to non-hydrostatic deep-atmosphere dynamical cores is encouraged despite the fact that the initial data slightly depart from the formal balance. Note that all modelling choices like the selection of the prognostic variables, coordinate systems, filters, diffusion mechanisms, staggering options, numerical techniques or computational meshes are left to the modelling group. The goal is to test the dynamical cores in their proposed operational configurations. The key question is whether a dynamical core converges with increasing horizontal resolution within the uncertainty of high-resolution reference solutions. This approach also reveals features of the numerical schemes, like Gibbs oscillations, truncation errors, time lags of the solution as well as diffusive properties.

The test assesses the evolution of an idealized baroclinic wave in the northern hemisphere over a short ten-day time period. In addition, the initial conditions can also be used to test the ability of a model to maintain a non-trivial, analytic steady-state solution. Recently, a similar baroclinic-wave test case has been independently developed by Polvani *et al.* (2004) (hereafter POL). The main difference between POL and the test proposed here is that POL prescribe an explicit second-order horizontal diffusion as part of the underlying equations that strongly damps the evolution of the baroclinic wave, in particular the cascade mechanism at small scales. Additionally, the POL initial conditions require the numerical evaluation of 1D integrals and vertical interpolations of horizontal-mean temperatures to define a quasi-realistic balanced flow in the northern hemisphere. In contrast, the test proposed here starts from a reasonably realistic initial dataset in both hemispheres which is completely represented by analytic expressions in

pressure-based vertical coordinates. An accurate method for height-based or isentropic vertical coordinates has also been devised which makes our initial conditions highly versatile. Overall, the evolution in POL is controlled by the diffusion processes that smooth out small-scale features expected to be resolved at higher resolutions. This diffusion is needed to obtain the exact numerical convergence of two models assessed in their study. On the other hand, POL also shows that any other choice of the diffusion mechanism leads to a truly different numerically converged flow field.

Our test is designed to evaluate dynamical cores in the form that will be applied to Numerical Weather Prediction (NWP) and climate simulations at the resolutions typically used for these applications. Thus the aim of the new test presented here is not to compute a numerically converged diffusive solution but to assess the accuracy of solutions at any resolved scale. An analytic solution to the problem is not known. We compute a set of reference solutions with a variety of models at very high resolutions. This set is used to estimate the uncertainty in any of the reference solutions. Throughout this paper, we consider the subgrid-scale diffusivity as a modelling choice and therefore part of the numerical scheme as also argued by Held and Suarez (1994) and Boer and Denis (1997). As a consequence, our models solve different equations but aim at approximating an almost inviscid equation set. Indeed, for some schemes, such as those incorporating nonlinear shape preservation, the inherent diffusion can not be isolated. Thus this test allows the comparison of schemes with very different diffusive behaviours, like a spectral transform semi-Lagrangian method (Collins *et al.* 2004) or a monotonic finite volume technique (Lin 2004), that do not require explicit horizontal diffusion to control the solution spectra at the small scales. In contrast, alternative schemes, like a spectral transform Eulerian method (Collins *et al.* 2004) or a finite-difference approach on an icosahedral grid (Majewski *et al.* 2002), generally apply a rather scale-selective fourth-order horizontal diffusion mechanism to prevent the accumulation of kinetic energy in the smallest scales or to ensure numerical stability. These four types of models are compared herein. The study sheds light on the convergence-with-resolution model characteristics and defines a set of reference solutions along with their uncertainties.

The paper is organized as follows. The design principles of the baroclinic-wave test case as well as the discussion of the initial conditions are presented in section 2. Section 3 briefly summarizes the four hydrostatic dynamical cores used in the study. The models represent a broad range of numerical approaches and provide independent reference solutions. In section 4 the models are assessed in a steady-state scenario without an explicitly imposed trigger for a baroclinic perturbation. This approach reveals how well the dynamical cores maintain the zonally symmetric initial state. The evolution of the baroclinic wave is discussed in section 5. In particular, this section includes a model intercomparison and assesses the uncertainty of the reference solutions. The findings are summarized in section 6.

2.   DEFINITION OF THE TEST CASE

The underlying concept of the initial-value test was inspired by ideas in Hoskins and Simmons (1975) and Simmons and Hoskins (1975) who used a numerically balanced initial flow with a superimposed wave-number-8 perturbation. In contrast to these authors though, the new idealized test is entirely described by analytic initial conditions with a local non-periodic perturbation (Jablonowski 2004). The initial data can therefore be immediately computed on any computational mesh.

The deterministic test is comprised of two parts. First, the dynamical core is initialized with steady-state, balanced initial conditions that are an analytic solution to

the hydrostatic primitive equations. This model set-up reveals how well a dynamical core maintains this initial state before numerical round-off and truncation errors as well as gravity waves degrade the steady state. Second, a relatively large-scale but localized Gaussian hill perturbation is superimposed onto the zonal wind in the northern midlatitudes. This small-amplitude disturbance triggers the evolution of a baroclinic wave over the course of several days.

(a)   The initial state

The baroclinic-wave test case has been developed for dry dynamical cores with pressure-based vertical coordinates like the pure pressure coordinate p, the pure $\sigma = p/p_s$ coordinate (Phillips 1957) or the η (hybrid $\sigma - p$) coordinate (Simmons and Burridge 1981) with p_s symbolizing the surface pressure. The latter two coordinate systems are typically used in GCMs today. The definition of the $\eta = A(\eta)p_0 + B(\eta)p_s$ coordinate together with the specification of all level coefficients A and B for the integrations herein are provided in Jablonowski and Williamson (2006). Note that the reference pressure p_0 is set to 10^5 Pa. For the test described here, the surface pressure p_s is constant and chosen to be $p_s(\lambda, \varphi) = p_0 = 10^5$ Pa. This guarantees that constant η-surfaces coincide with constant σ or pressure surfaces if there are no variations in the surface pressure, which is the case here. If other σ-systems like $\sigma = (p - p_t)/(p_s - p_t)$ are used (Kasahara 1974), the chosen pressure p_t at the model top should be set to zero. In general, the choice of the vertical coordinate system is left to the modelling group despite the fact that each vertical coordinate system implies a different boundary condition for the vertical velocity. In practice, this has been found to be insignificant for the evolution of the baroclinic wave over a ten-day time period (not shown). If other generalized vertical coordinate systems are used, like height-based or hybrid isentropic-σ levels, either an iterative method or vertical interpolations of the initial conditions become necessary. Details of the iterative method, which is the preferred choice, are provided in the appendix.

The initial state is defined by analytic expressions in spherical (λ, φ, η) coordinates where $\lambda \in [0, 2\pi]$ stands for the longitude, $\varphi \in [-\pi/2, \pi/2]$ represents the latitude and $\eta \in [0, 1]$ denotes the position in the vertical direction which is unity at the surface and approaches zero at the model top. The subsequent expressions can also be straightforwardly transformed into different, e.g. Cartesian, coordinate systems. All physical constants used in the test specification are listed later. Users of the test case are encouraged to select the same parameter set in their models to foster future model intercomparisons.

Assuming that a model utilizes η-levels, an auxiliary variable η_v is defined by

$$\eta_v = (\eta - \eta_0)\frac{\pi}{2}, \tag{1}$$

with $\eta_0 = 0.252$. In the case of a vertical σ coordinate, the η-level in Eq. (1) can be simply replaced with the corresponding σ value. This principle applies to all subsequent equations. Equation (1) can also be directly applied to models with pure pressure coordinates if $\eta = \sigma = p/p_s$ is adopted at each pressure level p.

The flow field is comprised of two symmetric zonal jets in midlatitudes. The zonal wind u is defined as

$$u(\lambda, \varphi, \eta) = u_0 \cos^{3/2}\eta_v \sin^2(2\varphi). \tag{2}$$

Here the maximum amplitude u_0 is set to 35 m s^{-1} which is close to the wind speed of the zonal-mean time-mean jet streams in the troposphere. The meridional wind v is set

to zero. In addition, the vertical velocity is set to zero for non-hydrostatic set-ups. This flow field in non-divergent and allows the derivation of the analytic initial data even for models in vorticity-divergence (ζ, δ) form. In particular, the radial outward component of the relative vorticity ζ is given by

$$\zeta(\lambda, \varphi, \eta) = \frac{-4u_0}{a} \cos^{3/2} \eta_v \sin \varphi \cos \varphi \, (2 - 5 \sin^2 \varphi), \tag{3}$$

and $\delta = 0$ is automatically fulfilled; $a = 6.371229 \times 10^6$ m indicates the mean radius of the Earth.

The horizontally averaged temperature $\langle T(\eta) \rangle$ is split into two representations for the lower (Eq. (4)) and middle (Eq. (5)) atmosphere. This introduces the characteristic atmospheric temperature profiles especially at upper levels. They are given by

$$\langle T(\eta) \rangle = T_0 \eta^{R_d \Gamma / g}, \qquad \text{for } \eta_s \geq \eta \geq \eta_t, \tag{4}$$
$$\langle T(\eta) \rangle = T_0 \eta^{R_d \Gamma / g} + \Delta T (\eta_t - \eta)^5, \quad \text{for } \eta_t > \eta, \tag{5}$$

with the surface level $\eta_s = 1$, the tropopause level $\eta_t = 0.2$ and the horizontal-mean temperature at the surface $T_0 = 288$ K. The temperature lapse rate Γ is set to 0.005 K m^{-1} which is similar to the observed diabatic lapse rate. Additionally, the empirical temperature difference $\Delta T = 4.8 \times 10^5$ K is chosen; $R_d = 287.0$ J kg^{-1}K^{-1} represents the ideal gas constant for dry air and $g = 9.80616$ m s^{-2} is the gravitational acceleration. The total temperature distribution comprises the horizontal-mean temperature and a horizontal variation at each level. It is given by

$$\begin{aligned}
T(\lambda, \varphi, \eta) = \langle T(\eta) \rangle &+ \frac{3}{4} \frac{\eta \pi u_0}{R_d} \sin \eta_v \cos^{1/2} \eta_v \\
&\times \left[\left\{ -2 \sin^6 \varphi \left(\cos^2 \varphi + \frac{1}{3} \right) + \frac{10}{63} \right\} 2u_0 \cos^{3/2} \eta_v \right. \\
&\left. + \left\{ \frac{8}{5} \cos^3 \varphi \left(\sin^2 \varphi + \frac{2}{3} \right) - \frac{\pi}{4} \right\} a\Omega \right],
\end{aligned} \tag{6}$$

where $\Omega = 7.29212 \times 10^{-5}$ s^{-1} denotes the earth's angular velocity.

The geopotential $\Phi = gz$ completes the description of the steady-state initial conditions where z symbolizes the elevation of a model level η and g is the acceleration due to gravity. The total geopotential distribution $\Phi = \langle \Phi \rangle + \Phi'$ comprises the horizontal-mean geopotential $\langle \Phi \rangle$ and a horizontal variation Φ' at each level. This is analogous to the description of the temperature field. The geopotential is determined by

$$\begin{aligned}
\Phi(\lambda, \varphi, \eta) = \langle \Phi(\eta) \rangle &+ u_0 \cos^{3/2} \eta_v \\
&\times \left[\left\{ -2 \sin^6 \varphi \left(\cos^2 \varphi + \frac{1}{3} \right) + \frac{10}{63} \right\} u_0 \cos^{3/2} \eta_v \right. \\
&\left. + \left\{ \frac{8}{5} \cos^3 \varphi \left(\sin^2 \varphi + \frac{2}{3} \right) - \frac{\pi}{4} \right\} a\Omega \right],
\end{aligned} \tag{7}$$

with

$$\langle \Phi(\eta) \rangle = \frac{T_0 g}{\Gamma}(1 - \eta^{R_d \Gamma/g}), \qquad \text{for } \eta_s \geq \eta \geq \eta_t, \tag{8}$$

$$\langle \Phi(\eta) \rangle = \frac{T_0 g}{\Gamma}(1 - \eta^{R_d \Gamma/g}) - R_d \Delta T \left\{ \left(\ln\left(\frac{\eta}{\eta_t}\right) + \frac{137}{60} \right) \eta_t^5 - 5\eta_t^4 \eta + 5\eta_t^3 \eta^2 \right.$$
$$\left. - \frac{10}{3}\eta_t^2 \eta^3 + \frac{5}{4}\eta_t \eta^4 - \frac{1}{5}\eta^5 \right\}, \quad \text{for } \eta_t > \eta. \tag{9}$$

This formulation enforces the hydrostatic balance analytically and ensures the continuity of the geopotential at the tropopause level η_t. In hydrostatic models with pressure-based vertical coordinates, it is only necessary to initialize the surface geopotential $\Phi_s = g z_s$. It balances the non-zero zonal wind at the surface with surface elevation z_s and is determined by setting $\eta = \eta_s$ in Eq. (7). Note that Φ_s is actually a function of latitude only. The geopotential equation can fully be utilized for dynamical cores with height-based vertical coordinates. Then, a root-finding algorithm is recommended to determine the corresponding η-level for any given height z. This iterative method, which is also applicable to isentropic vertical coordinates, is outlined in the appendix. The resulting η-level is accurate to machine precision and can consequently be used to compute the initial dataset.

Besides the assessment of the steady-state solution, the test is primarily designed to simulate the evolution of a baroclinic wave. It can be triggered if the initial conditions are overlaid with a perturbation. Here a perturbation with a Gaussian profile is selected and centred at $(\lambda_c, \varphi_c) = (\pi/9, 2\pi/9)$ which points to the location (20°E, 40°N). The perturbation u' of the zonal wind

$$u'(\lambda, \varphi, \eta) = u_p \exp\left\{ -\left(\frac{r}{R}\right)^2 \right\}, \tag{10}$$

with radius $R = a/10$ and maximum amplitude $u_p = 1 \text{ m s}^{-1}$, is superimposed on the zonal wind field (Eq. (2)) by adding u' to u at each grid point at all model levels. Here the great circle distance r is given by

$$r = a \arccos\{\sin \varphi_c \sin \varphi + \cos \varphi_c \cos \varphi \cos(\lambda - \lambda_c)\}. \tag{11}$$

The corresponding overlaying (ζ', δ') perturbations at each level for models in vorticity-divergence form are

$$\zeta'(\lambda, \varphi, \eta) = \frac{u_p}{a} \exp\left\{ -\left(\frac{r}{R}\right)^2 \right\} \times \left\{ \tan \varphi - 2 \left(\frac{a}{R}\right)^2 \arccos(X) \right.$$
$$\left. \times \frac{\sin \varphi_c \, \cos \varphi - \cos \varphi_c \sin \varphi \cos(\lambda - \lambda_c)}{\sqrt{1 - X^2}} \right\}, \tag{12}$$

$$\delta'(\lambda, \varphi, \eta) = \frac{-2u_p a}{R^2} \exp\left\{ -\left(\frac{r}{R}\right)^2 \right\} \arccos(X) \frac{\cos \varphi_c \sin(\lambda - \lambda_c)}{\sqrt{1 - X^2}}, \tag{13}$$

with $X = \sin \varphi_c \sin \varphi + \cos \varphi_c \cos \varphi \cos(\lambda - \lambda_c)$. For both singular points (λ_c, φ_c) and $(\lambda_c + \pi, -\varphi_c)$ with $X^2 = 1$, δ' is identical zero. In addition, $\zeta'(\lambda_c, \varphi_c) = u_p \tan \varphi/a$ is well defined and $\lim_{\lambda \to \lambda_c + \pi, \varphi \to -\varphi_c} \zeta'$ is zero. Similarly, $\lim_{\varphi \to \pm \pi/2} \zeta'$ is zero at the poles. Figures of the perturbation fields u', ζ' and δ' are shown in Jablonowski and Williamson (2006) where the derivation of the initial conditions is documented in more detail.

(b)   Characteristics of the initial data

The balanced initial flow field comprises a zonally symmetric basic state with a jet in the midlatitudes of each hemisphere and a quasi-realistic temperature distribution, which are displayed in Figs. 1(a) and (b). In addition, Fig. 1 shows the profile of the surface geopotential (Fig. 1(c)), three selected vertical temperature distributions on a logarithmic scale (Fig. 1(d)), the initial geopotential height field (Fig. 1(e)) and the unperturbed relative-vorticity distribution (Fig. 1(f)). Overall, the atmospheric conditions resemble the climatic state of a winter hemisphere reasonably well. The centres of the midlatitudinal jets at $45°$N/S are placed at the pressure level $p = \eta_0 p_s = 252$ hPa, which lies just below the tropopause level at $p = 200$ hPa. Furthermore, the globally averaged temperature of this distribution is approximately 256.4 K, which closely matches the observed global temperature of the atmosphere. The horizontally averaged temperature at the surface is selected to be $T = 288$ K, which corresponds exactly to the surface temperature of the so-called US standard atmosphere (US Standard Atmosphere 1976). In addition, the temperature distribution captures an idealized stratospheric and mesospheric temperature profile and prescribes a low-level temperature inversion in polar regions. The test design guarantees static, inertial and symmetric stability properties, but is unstable with respect to baroclinic or barotropic instability mechanisms. These two forms of instabilities result from the vertical and meridional wind shear characteristics. In particular, the growth rate of the baroclinic disturbance highly depends on the strength of the vertical shear. The inertial, symmetric and static stability criteria are analysed in Fig. 2. Figures 2(a) and (b) show the absolute- and potential-vorticity fields. The flow is stable with respect to inertial motions since the absolute-vorticity field $(\zeta + 2\Omega \sin \varphi)$ of the basic flow is positive in the northern hemisphere and negative in the southern hemisphere (Holton 1992). The same argument holds for the potential vorticity that indicates the symmetric stability of the flow. Here the potential vorticity PV is assessed in form of the Ertel PV. Similar distributions of the Ertel potential vorticity have also been found by Simmons and Hoskins (1976).

The initial data are also statically stable. This is displayed in Figs. 2(c) and (d), which show the potential-temperature distribution Θ and the Brunt–Väisälä frequency N. The flow is statically stable if the vertical potential-temperature gradient $\partial \Theta / \partial z$ with respect to the height z, or equivalently the buoyancy frequency N, is positive everywhere. The latter can clearly be seen in Fig. 2(d). Note that the buoyancy frequency at the surface approaches zero at the poles, which defines almost neutrally stratified conditions.

(c)   Test strategy

The following two-step test strategy is suggested. First, the dynamical core is initialized with the balanced initial conditions and run for 30 model days at varying horizontal resolutions. This is a stringent test of the dynamics that not only serves as a debugging tool but also as an assessment tool for the algorithmic design of the numerical scheme and its horizontal grid. For these model runs error norms can be directly assessed since the initial state is the analytic solution (see section 4). If possible, the unperturbed model simulations should be run without horizontal or vertical diffusion. Most often, the latter is a component of the physics package and therefore already inactive in a dynamical core simulation. In addition, no Rayleigh friction near the model top (if included in the model) should be applied. All three aforementioned diffusion and dissipation mechanisms would damp the initial data over time which is consequently reflected in the error statistics.
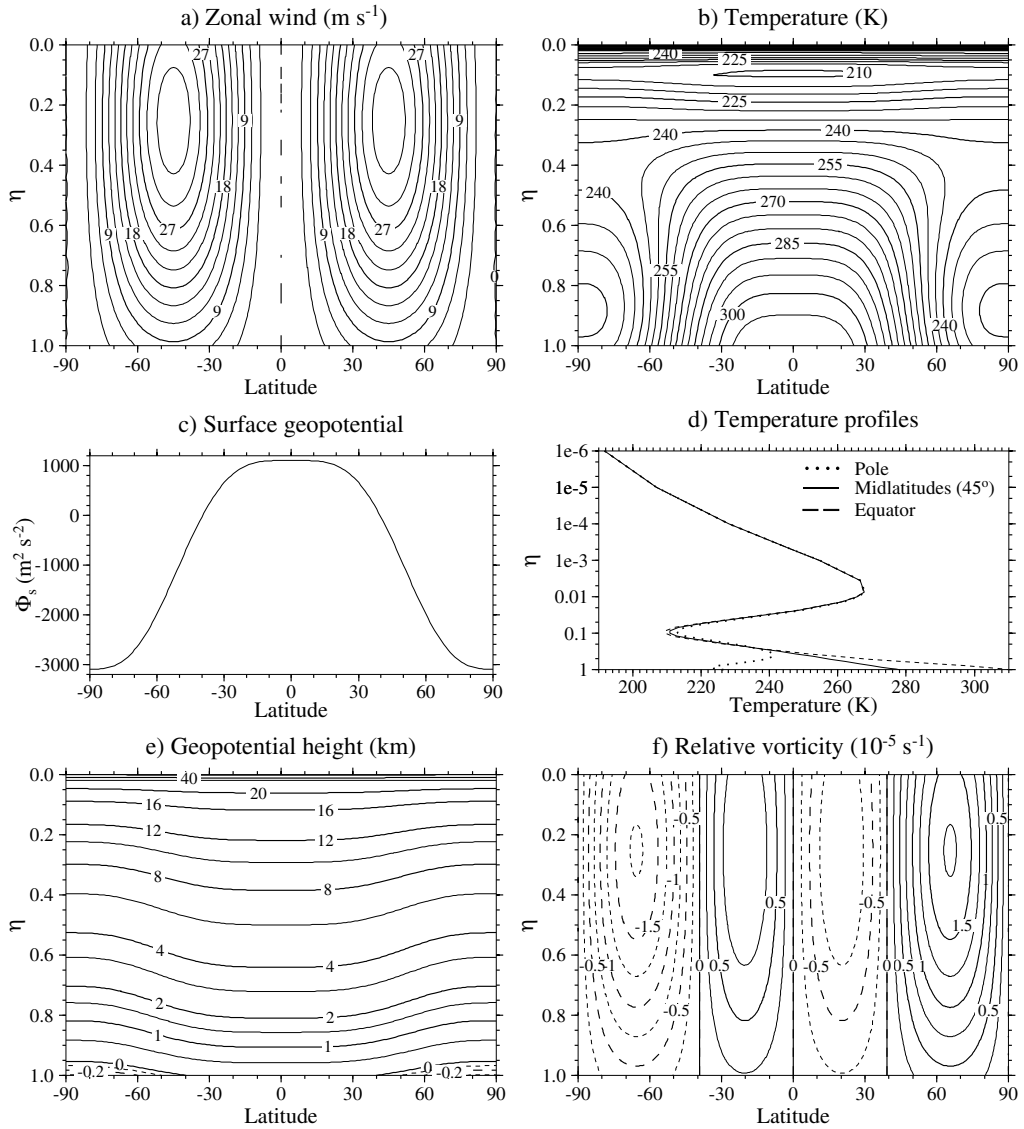
Figure 1. Initial conditions for the (a) unperturbed zonal wind u, (b) temperature T, (c) surface geopotential Φ_s, (d) vertical temperature profiles at the equator, 45°N/S and the poles, (e) geopotential-height field z (km) and (f) unperturbed relative vorticity ζ (in 10^{-5} s^{-1}). The contour levels in (e) are non-uniformly spaced, and negative contours are dashed. Note that the η coordinate coincides with $\sigma = p/p_s$.

Second, the evolution of a baroclinic wave in the northern hemisphere is triggered when using the initial conditions with the overlaid zonal wind perturbation. As before, different horizontal resolutions should be assessed to estimate the convergence characteristics. In general, the baroclinic wave starts growing observably around day 4 and evolves rapidly thereafter with explosive cyclogenesis at model day 8. The wave train breaks after day 9 and generates a full circulation in both hemispheres between day 20 and day 30 depending on the model formulation. Therefore, the simulation should cover at least a ten-day time period that captures the initial and rapid development stages of the baroclinic disturbance. If longer time integrations are performed (e.g. up to 30 days as

a) Absolute vorticity (10^{-4} s^{-1})

b) Potential vorticity (PV units)

c) Potential temperature (K)

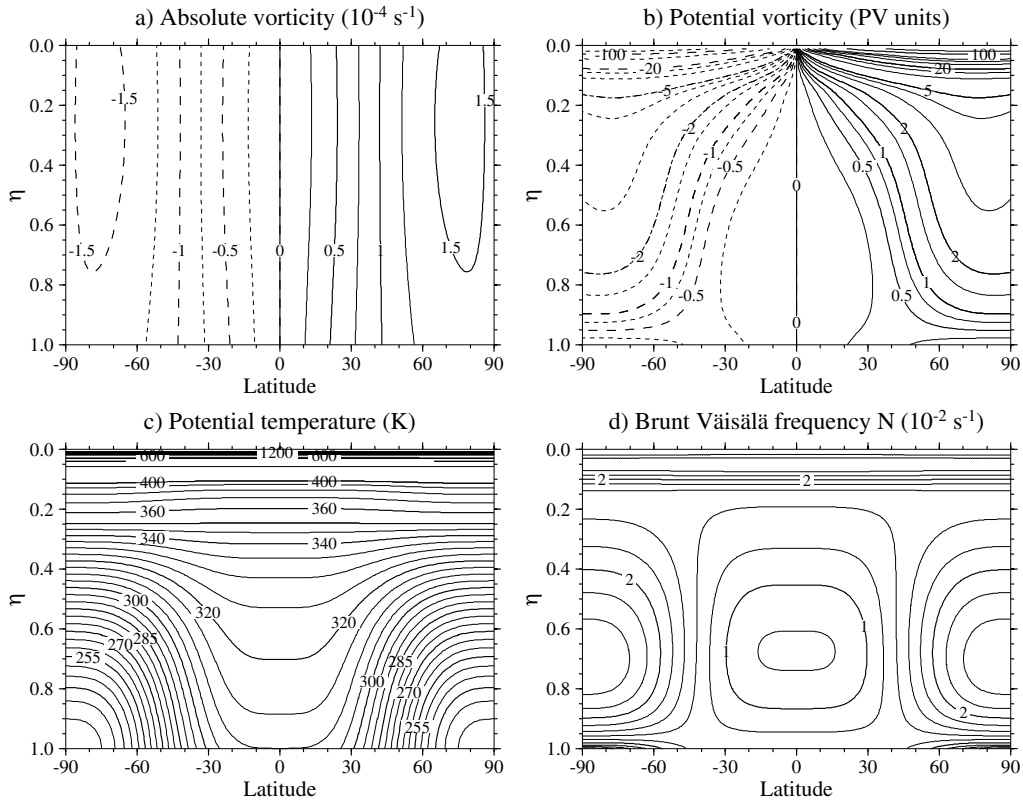d) Brunt Väisälä frequency N (10^{-2} s^{-1})

Figure 2. Assessment of the stability of the balanced initial conditions: (a) absolute vorticity (inertial stability), (b) potential vorticity with varying contour levels in PV units $= 10^{-6}$ K kg^{-1}m^{2}s^{-1} (symmetric stability), (c) potential temperature Θ and (d) Brunt–Väisälä frequency N (static stability).

in the subsequent examples) the spread of the numerical solutions increases noticeably from model day 12 onwards. This indicates the predictability limit of the test case. Nevertheless, the initial development stages of new systems at the leading edge of the baroclinic wave train (compare also to Simmons and Hoskins (1979)) are still predicted reliably until day 16.

The baroclinic wave, although idealized, represents very realistic flow features. Strong temperature fronts develop that are associated with the evolving low and high pressure systems. In addition, the longitudinal wave number with maximum growth rate lies between 5 and 9 (not shown) which agrees well with observations and assessments in the literature (e.g. Simmons and Hoskins 1977). It is important to note that the baroclinic-wave test case does not have an analytic solution. Therefore, high-resolution reference solutions and their uncertainties are assessed in section 5.

3. DESCRIPTION OF THE DYNAMICAL CORES

The baroclinic-wave test sequence has been applied to the three hydrostatic dynamical cores that are part of National Center for Atmospheric Research's Community Atmosphere Model version 3 (NCAR CAM3). In particular, these are a finite-volume dynamical core (Lin 2004) and the spectral transform semi-Lagrangian and Eulerian dynamics packages. They are documented in Collins *et al.* (2004) who also discuss the

exact equation sets. In addition, the dynamical core of the operational weather forecast model GME at the German Weather Service (DWD) has been tested. This hydrostatic finite-difference model is based on an icosahedral mesh with a triangulated grid structure (see also the documentation in Majewski *et al.* (2002)). As mentioned earlier, our test is intended to evaluate the accuracy and diffusive aspects of dynamical cores. It targets dynamical cores that are configured for their NWP or climate simulation applications. This includes any diffusion or filtering operators with typical coefficients even if they are not needed for this test. We briefly summarize the four dynamical cores below.

(*a*) *Eulerian*

The Eulerian dynamical core in vorticity-divergence form is based on the traditional three-time-level, semi-implicit spectral transform approximations applied on a quadratically unaliased transform grid with horizontal triangular truncation (Machenhauer 1979). The vertical coordinate is terrain-following hybrid pressure. The core includes ∇^4 horizontal diffusion on temperature, divergence and vorticity to control the energy at the smallest resolved scales. The coefficient has been chosen for each resolution to yield a reasonably straight tail for the kinetic-energy spectra in realistically forced simulations. The model also includes a ∇^2 horizontal diffusion on the top three levels of the model which serves as a top boundary condition to control upward-propagating waves. The temperature equation includes a frictional heating term corresponding to the momentum diffusion. The Eulerian core includes an a posteriori mass fixer applied at every time step. An energy fixer is also available with the Eulerian core but was was not invoked for our experiments. The three-time-level core includes a time filter to control the $2\Delta t$ time computational modes.

(*b*) *Semi-Lagrangian*

The semi-Lagrangian dynamical core is based on two-time-level, semi-implicit semi-Lagrangian spectral transform approximations with quasi-cubic Lagrangian polynomial interpolants. Triangular truncation is adopted. The core is based on the same terrain-following vertical coordinate as the Eulerian core. For our experiments we use a quadratically unaliased transform grid. We do not include the ∇^4 and ∇^2 horizontal diffusions which are available in the core; the interpolants control the energy at the smallest resolved scales for the time-scales considered here. Since the energy loss due to the damping by the interpolants is not explicitly known, a term corresponding to the ∇^4 diffusive heating in the Eulerian core cannot be derived. Thus an a posteriori energy fixer is applied every time step in the model which acts as a globally uniform heating term. In addition, the a posteriori mass fixer is invoked every time step as in the Eulerian core. Note that the semi-Lagrangian dynamical core applies a decentring technique to damp the noise induced by orographic resonance (Collins *et al.* 2004). Although not needed for stability reasons in idealized studies, the standard CAM3 decentring parameter $\epsilon = 0.2$ is used, here in the spirit of evaluating the core as it would be applied in practice.

(*c*) *Finite volume*

The mass-conservative finite-volume (FV) dynamical core in flux form is built upon a 2D shallow-water approach in the horizontal plane (Lin and Rood 1996, 1997). The vertical discretization utilizes a 'Lagrangian control-volume' principle, which is based on a terrain-following 'floating' Lagrangian coordinate system and a fixed 'Eulerian' reference frame. In particular, the vertically stacked finite volumes are allowed to float

for a duration of several time steps before they are mapped back monotonically and conservatively to the fixed reference system (Lin 2004). The latter coincides with the hybrid levels of CAM3's Eulerian and semi-Lagrangian dynamical cores. The advection algorithm makes use of the monotonic third-order piecewise parabolic method (Colella and Woodward 1984) with an explicit time-stepping scheme. Lower-order approximations are applied near the model top to add extra diffusion at the upper boundary. This configuration has been used for both the steady-state and baroclinic-wave simulations. The FV dynamics package is built upon a regular latitude–longitude grid and employs both a weak three-point digital filter in midlatitudes as well as a fast Fourier transform (FFT) filter in polar regions. These control the unstable waves in the zonal direction that result from the convergence of the computational grid near the poles. The FV dynamical core does not include explicit horizontal ∇^4 or ∇^2 diffusion terms. The smallest scales are controlled by the monotonicity constraints of the numerical scheme. In addition, a divergence damping mechanism is applied that damps the divergence δ in the form of $\nabla^2(\nu\delta)$ where ν is a spatially varying divergence damping coefficient. As with the semi-Lagrangian core, the energy loss due to the numerical damping is not explicitly known. Therefore, an a posteriori energy fixer is applied that has the same form as that in the semi-Lagrangian model.

(d)   GME

The dynamical core of the weather prediction system GME applies a finite-difference approximation with local spherical basis functions at each grid point. The horizontal grid is based on an icosahedron which is further subdivided into smaller triangles. An Arakawa-A grid staggering is chosen that places the prognostic variables at the vertices of the triangles. In the vertical direction GME utilizes the same terrain-following η coordinate as the CAM3 Eulerian dynamics package. The semi-implicit numerical scheme is second-order accurate and applies a classical leapfrog three-time-level approach with an Asselin time filter. Due to the quasi-uniformity of the computational grid no longitudinal FFT-filtering is needed. The smallest scales are controlled by a ∇^4 horizontal diffusion mechanism that is applied to the prognostic wind and temperature variables u, v and T. Near the model top the ∇^4 diffusion is replaced with a stronger ∇^2 diffusion operator that serves as a top boundary condition. Neither a mass fixer nor an energy fixer is applied.

(e)   Resolutions and model parameters

For the basic experiments here, which examine the effect of horizontal resolution, we adopt the standard 26 vertical levels (L26) used to produce the CAM3 control simulations (Collins *et al.* 2006). For mid-range horizontal resolutions we also calculate the solutions with increased and decreased vertical resolution: 49 levels (L49), which essentially halve the thickness of the standard 26-level grid intervals, and the 18 levels (L18) used for the standard simulations of the predecessor of CAM3, the Community Climate Model version 3 (CCM3) (Kiehl *et al.* 1998). In all level set-ups, the centres of the topmost levels lie between 3 and 5 hPa (see the appendix in Jablonowski and Williamson (2006) for the exact level placement).

(i)   *Eulerian (EUL).*   Table 1 lists the horizontal resolutions, time steps and ∇^4 diffusion coefficients K_4 for the spectral transform Eulerian dynamical core. The abbreviation T symbolizes a triangular truncation which is followed by the maximum

TABLE 1. Horizontal grid resolutions, time steps and diffusion coefficients for the spectral Eulerian (EUL) dynamical core in CAM3. The ∇^2 horizontal diffusion coefficient, $K_2 = 2.5 \times 10^5$ m²s⁻¹, is independent of horizontal resolution.

Spectral resolution	No. grid points lat × lon	Grid distance at the equator (km)	Time step Δt (s)	Diffusion coefficient K_4 (m⁴s⁻¹)
T21	32 × 64	625	2400	2.0×10^{16}
T42	64 × 128	313	1200	1.0×10^{16}
T85	128 × 256	156	600	1.0×10^{15}
T170	256 × 512	78	300	1.5×10^{14}
T340	512 × 1024	39	150	1.5×10^{13}

TABLE 2. Horizontal grid resolutions and time steps for the finite-volume (FV) dynamical core in CAM3. The subcycled dynamics time step $\Delta\tau = \Delta t/m$ with $m = 10$ is shown.

Resolution $\Delta\varphi \times \Delta\lambda$	No. grid points lat × lon	Grid distance at equator $\Delta y \times \Delta x$ (km × km)	Time step (subcycled) $\Delta\tau = \Delta t/m$ (s)
4° × 5°	46 × 72	435 × 556	720
2° × 2.5°	91 × 144	220 × 278	360
1° × 1.25°	181 × 288	111 × 139	180
0.5° × 0.625°	361 × 576	55 × 69	90
0.25° × 0.3125°	721 × 1152	28 × 35	45

resolved wave number. The ∇^2 horizontal diffusion coefficient, $K_2 = 2.5 \times 10^5$ m²s⁻¹, is independent of horizontal resolution.

(ii) *Semi-Lagrangian (SLD).* The same resolutions as in the Eulerian case are tested (see Table 1). In contrast though, the time step Δt is three times the corresponding Eulerian value and both diffusion coefficients K_4 and K_2 are set to zero.

(iii) *Finite volume (FV).* The finite-volume dynamical core is run at the horizontal resolutions listed in Table 2, which also shows the approximate equatorial grid distances in physical space to ease the comparison to the other dynamical cores. The time steps represent the lengths of the subcycled time steps in the dynamical core (equivalent to the notation $\Delta\tau$ in Lin (2004)). In our set-up, the vertical re-mapping algorithm was invoked every $m = 10$ subcycled time steps.

(iv) *GME.* The horizontal grid of GME is based on an icosahedron with 20 equilateral triangles. The sides of these base triangles are further subdivided into *ni* equal intervals. The corresponding maximum and minimum grid distances, time steps and diffusion coefficients are listed in Table 3.

(f) Run time

The run times of the four dynamical cores at their mid-range and second-highest resolutions and 26 levels are listed in Table 4. This table serves as a general guide for the computational costs of each model, while acknowledging that such costs are hardware-dependent and vary with the ingenuity of the programmer. The run-time data represent the wallclock time needed to complete one model day on a 32-processor node of an IBM Power4 architecture when using a pure Message Passing Interface parallelization approach. Identical compiler optimization flags were used for all models. No effort was made to optimize the numerical schemes or to configure the models in their optimal

TABLE 3. HORIZONTAL GRID RESOLUTIONS, TIME STEPS AND DIFFUSION COEFFI-
CIENTS FOR GME WITH AN ICOSAHEDRAL GRID

Resolution ni	No. grid points	Min. grid distance (km)	Max. grid distance (km)	Time step Δt (s)	Diffusion coefficients	
					K_4 $(m^4 s^{-1})$	K_2 $(m^2 s^{-1})$
16	2562	440	526	1600	5.0×10^{16}	2.0×10^6
32	10242	220	263	800	6.0×10^{15}	1.5×10^6
64	40962	110	132	400	1.0×10^{15}	1.0×10^6
128	163842	55	66	200	1.2×10^{14}	2.0×10^5
256	655362	26	33	100	1.2×10^{13}	2.5×10^4

TABLE 4. WALLCLOCK TIME (s) FOR ONE MODEL DAY MEASURED
ON A 32-PROCESSOR NODE OF AN IBM POWER4 ARCHITECTURE. THE
RUN TIMES ARE LISTED FOR THE MID-RANGE AND SECOND-HIGHEST
HORIZONTAL RESOLUTIONS AND 26 LEVELS.

Horizontal resolution	Wallclock time (s) for 1 day			
	EUL	SLD	FV	GME
T85/$1° \times 1.25°$/ni64	44	24	66	48
T170/$0.5° \times 0.625°$/ni128	483	271	625	325
Increase in run time by a factor of	11.0	11.3	9.5	6.8

set-ups, such as selecting an optimal time step or switching from a quadratic to a
linear truncation technique in case of SLD. The dynamical cores represent the standard
versions in CAM3 and GME.

Table 4 shows that the semi-Lagrangian dynamical core clearly outperforms EUL,
FV and GME. At T85 the SLD dynamical core is roughly twice as fast as the other
three models at the comparable resolutions. At the higher-resolution T170 though, the
computational advantage of the SLD model diminishes especially in comparison to
GME. This is mainly due to a better cache utilization and reduced parallel overhead
in GME. GME shows a superlinear parallel speed-up as the horizontal resolution is
doubled and the time step is halved. In general, this corresponds to an increase in the
workload by a factor of 8 when disregarding any additional work for the parallelization.
Here, the increase in the wallclock time for GME ni128 lies around a factor of 7, whereas
the EUL T170, SLD T170 and FV $0.5° \times 0.625°$ dynamical cores show an increase
in the run time by a factor between 9.5 and 11.3 compared to the lower resolutions.
Nevertheless, the SLD T170 run is still the most cost-effective approach, while giving
similar simulation results as shown later.

4. STEADY-STATE TEST CASE

All four dynamical cores have been initialized with the steady-state initial condi-
tions and run for 30 model days at varying horizontal resolutions and 26 levels. In an
ideal case the models are expected to perfectly maintain the initial state which is the
analytic solution. In practice though, the initial state degrades over time as discussed
later. The errors are quantified via l_2 (root-mean-square) error norms. In particular, two
types of error measures are used that are applied to the zonal-wind field u. Note that we
do not normalize the error norms by the integral of the square of the analytic solution
which is customary for l_2 analyses of the shallow-water tests (Williamson *et al.* 1992).

The first l_2 error norm evaluates the symmetry-deviations from the zonal average. It is defined as

$$l_2(u(t) - \overline{u}(t)) = \left[\frac{1}{4\pi} \int_0^1 \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \{u(\lambda, \varphi, \eta, t) - \overline{u}(\varphi, \eta, t)\}^2 \cos\varphi \, d\lambda \, d\varphi \, d\eta \right]^{1/2}$$

$$\approx \left[\frac{\sum_k \sum_j \sum_i \{u(\lambda_i, \varphi_j, \eta_k, t) - \overline{u}(\varphi_j, \eta_k, t)\}^2 w_j \Delta\eta_k}{\sum_k \sum_j \sum_i w_j \Delta\eta_k} \right]^{1/2}, \quad (14)$$

where the overbar $\overline{(\,)}$ denotes the zonal average and the sums with indices (i, j, k) are taken over all longitude points λ_i, latitude points φ_j and vertical levels η_k of the global grid. The summation weights w_j are the Gaussian weights for the calculations on the spectral transform grids. On uniform grids, $w_j = |\sin(\varphi_{j+1/2}) - \sin(\varphi_{j-1/2})|$ is used where the half indices denote the location of the cell interfaces in the meridional direction. The weights

$$\Delta\eta_k = (\eta_{k+1/2} - \eta_{k-1/2})$$

indicate the thickness of a model layer. Here the half indices $k \pm 1/2$ point to the locations of the level interfaces. Equation (14) assumes that the longitudinal grid points are equally spaced.

The second l_2 norm assesses the degradation of the zonal average with respect to the analytic solution. It is defined by

$$l_2(\overline{u}(t) - \overline{u}(t=0)) = \left[\frac{1}{2} \int_0^1 \int_{-\pi/2}^{\pi/2} \{\overline{u}(\varphi, \eta, t) - \overline{u}(\varphi, \eta, t=0)\}^2 \cos\varphi \, d\varphi \, d\eta \right]^{1/2}$$

$$\approx \left[\frac{\sum_k \sum_j \{\overline{u}(\varphi_j, \eta_k, t) - \overline{u}(\varphi_j, \eta_k, t=0)\}^2 w_j \Delta\eta_k}{\sum_k \sum_j w_j \Delta\eta_k} \right]^{1/2}. \quad (15)$$

For both analyses the GME data are biquadratically interpolated to a regular latitude–longitude grid that is identical to the FV grid at a comparable resolution (see Tables 2 and 3). In general, alternative numerical approximations to the integrals can also be chosen if models with non-regular computational meshes are assessed.

Figure 3 compares the zonal-symmetry characteristics (Eq. (14)) of the dynamical cores. The simulations with EUL and FV at all horizontal resolutions show that both models maintain the zonal symmetry exactly up to machine precision (no figure shown). This is in contrast to SLD (Fig. 3(a)) and GME (Fig. 3(b)) which develop zonal asymmetries immediately after the start of the simulation. In the case of SLD the asymmetries grow continuously and become clearly visible after approximately four weeks, especially in surface maps of the meridional wind (not shown). Here it is interesting to note that the growth of the zonal asymmetry is slightly faster at increased resolution because of the decreased damping from the interpolations at higher resolution. The asymmetries are introduced at the machine precision level by calculations that involve the longitudes of the grid points such as the departure point, which is calculated as an offset from the arrival grid point. Thus the relative location of the departure point within the grid interval will differ at the rounding level for the points on a latitude circle even though the offset is the same to machine precision. In contrast, the Eulerian spectral and finite-volume methods use only the grid interval and not the actual longitudes of grid points and thus the calculations at all longitudes on a given latitude circle are identical. Other calculations in the semi-Lagrangian core that introduce rounding-level perturbations along latitude circles are the local geodesic trajectory calculations, the rotation of the
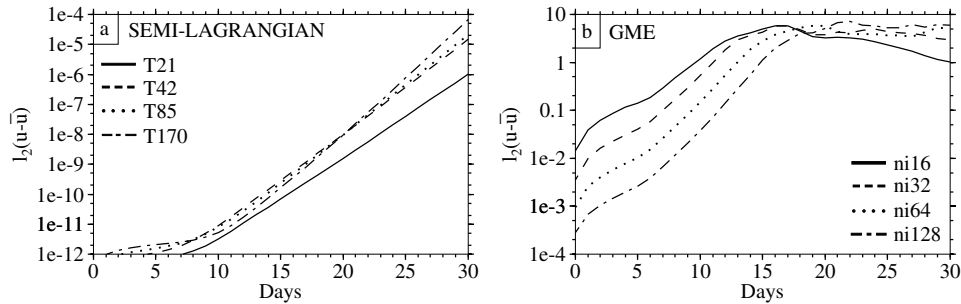
Figure 3. Root-mean square l_2 norm of $u - \overline{u}$ (m s^{-1}) for (a) the semi-Lagrangian and (b) the GME dynamical cores at varying horizontal resolutions and 26 levels.

unit vectors associated with the advection of a vector wind, and the calculation of a wave-number-one polar wind vector used for meridional interpolation which should be zero in this case but is only so to rounding (see Williamson and Olson (1994) and Collins *et al.* (2004) for details of the algorithms). We note that these are details of the particular implementation in the CAM3 SLD core. It is possible to devise coding strategies that would not introduce rounding-level longitudinal asymmetries, such as by interpolating differences from arrival point values rather than full fields. However, the elimination of rounding-level longitudinal asymmetries seems unimportant in practical atmospheric applications. In the case of GME (Fig. 3(b)) the initial errors as plotted at $t = 0$ are completely determined by the interpolation to the regular grid which sets the lower bound of the detectable signal. During the simulation this noise level is then overlaid with true non-zonal numerical truncation errors. These have a built-in wave-number-5 structure which arises from the variations in the icosahedral grid intervals. The wave-number-5 signal grows considerably over time due to the magnitude of the truncation errors and unstable mean state which leads to a visible breakdown of the zonal symmetry after several days. This effect is postponed at higher resolutions as the truncation error perturbations decrease.

Figure 4 shows the evolution of the zonal-mean zonal-wind error norms for the 30-day time period (Eq. (15)). In general, the degradation of the zonal-mean is mainly attributable to the initial generation of gravity waves in the φ-η plane which slowly decay over time. They adjust the geostrophic balance in the discrete system since the analytically balanced initial state is not entirely balanced in a discrete representation. Both the EUL (Fig. 4(a)) and FV (Fig. 4(c)) dynamical cores modulate the zonal mean initially and exhibit flat profiles thereafter. There are only minor (FV) or almost no (EUL) improvements of the l_2 errors with increasing horizontal resolution. In contrast, the l_2 norms of the SLD dynamical core (Fig. 4(b)) increase linearly over the 30-day forecast period. As resolutions increase and time steps decrease the zonal mean is better maintained. This behaviour is primarily determined by the decentring parameter $\epsilon = 0.2$, the length of the time step and the above-mentioned meridional position of the non-zonal geodesic trajectory calculations in polar regions where the algorithm is switched from a spherical-coordinate-based scheme (Williamson and Rasch 1989). Note that the SLD curves closely match the EUL results (Fig. 4(a)) if the spherical-coordinate-based trajectory calculation is used everywhere and if the decentring parameter ϵ is set to zero (Jablonowski and Williamson 2006, Fig. 4.3). The former is, of course, only reasonable with zonal advection as in this case, and the latter in the absence of mountains. We deliberately do not include those variations in the tests here so as to
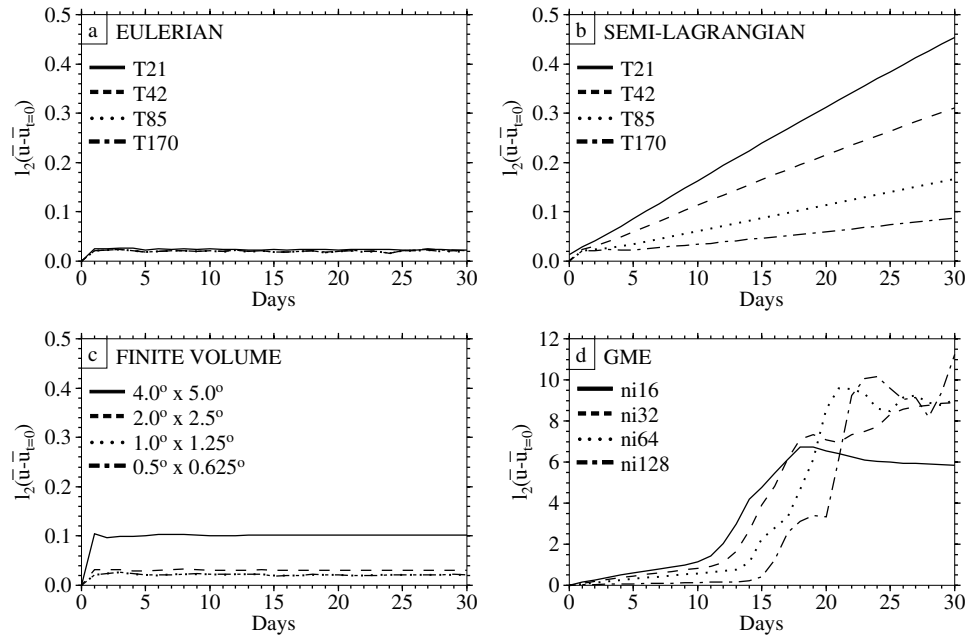
Figure 4.   Root-mean square l_2 norm of $\overline{u} - \overline{u}_{t=0}$ (m s^{-1}) for the four dynamical cores at varying horizontal resolutions and 26 levels. Note that the l_2 error scale in (d) differs from that in (a)–(c).

evaluate the scheme as it would be applied in practice. Again, the GME errors arise from the built-in wave-number-5 truncation errors which lead to the breakdown of the steady-state solution after several days (Fig. 4(d)). In general, such model behaviour is an artifact of the computational grid and the numerical method, but despite the clear exposure of this phenomenon it is somewhat artificial and unimportant unless one is trying to model a problem dominated by a zonal flow. For atmospheric applications this is clearly not the case since the forcing is strongly non-zonal. The wave-number-5 effect is therefore overly enhanced by the idealized set-up. Similar conclusion would also apply to models that utilize a cubed-sphere computational grid with a built-in wave-number-4 structure.

5. BAROCLINIC-WAVE TEST CASE

The baroclinic wave is triggered when overlaying the steady-state initial conditions with the zonal-wind perturbation. As before, 30-day simulations at varying resolutions are performed with all four dynamical cores. In general, the nonlinear analytic solution for finite perturbations is not known. Therefore, we determine a set of reference solutions numerically by applying all the dynamical cores at very high horizontal resolutions. Because the basic state is unstable, small differences in addition to our imposed initial perturbation will also grow with time. In particular small differences between the initial discrete representations of the different cores will grow with time. Such differences can be introduced by different truncation errors between the schemes or between different resolutions of the same scheme, or even by different relative locations of the grid points with respect to the imposed perturbation for a given scheme. Thus the reference solutions differ from each other and the differences grow with time. These differences provide an estimate of the uncertainty in the reference solutions.
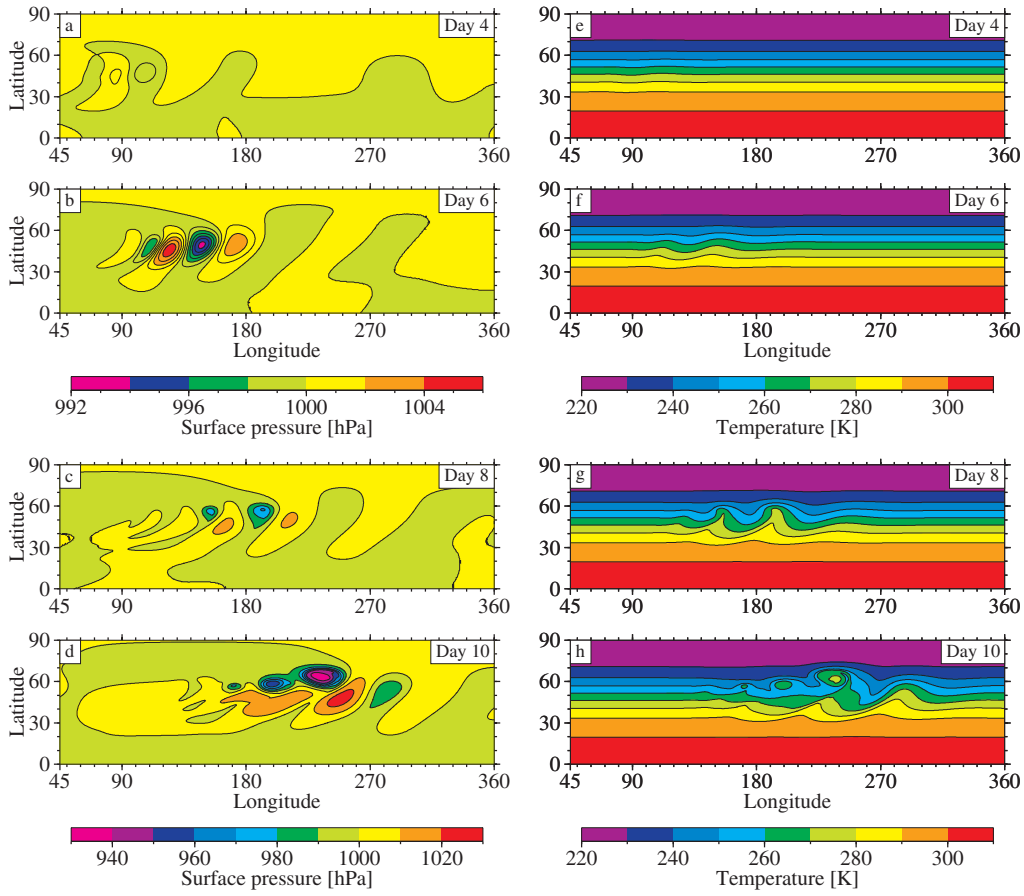
Figure 5. Evolution of the baroclinic wave from day 4 until day 10: Surface pressure (left column) and temperature at 850 hPa (right column) modelled with the finite-volume (FV) dynamical core at the horizontal resolution 0.5° × 0.625° and 26 levels.

This will be quantified in section 5(e). First we examine the detailed evolution of the growing wave.

(a) Time evolution

Figure 5 shows a time sequence of the evolving reference solution surface pressure and 850 hPa temperature field* from the 0.5° × 0.625° L26 FV integration. The wave grows very slowly until day 4. At day 6 the surface pressure shows two weak high and low pressure systems (Fig. 5(b)) and two small-amplitude waves are visible in the temperature field (Fig. 5(f)). At day 8 the highs and lows have deepened significantly (Fig. 5(c), note the change in the scale). In addition, the two waves in the temperature field have almost peaked and are beginning to wrap around with fronts trailing from them. A third upstream wave is now visible. At day 10 wave breaking has set in. Three closed cells with relatively higher temperatures (Fig. 5(h)) have formed with the leading front quite sharp.

* In Figs. 5–8, the 850 hPa temperature or relative-vorticity field is computed via a linear interpolation in ln p coordinates using the two surrounding model levels.
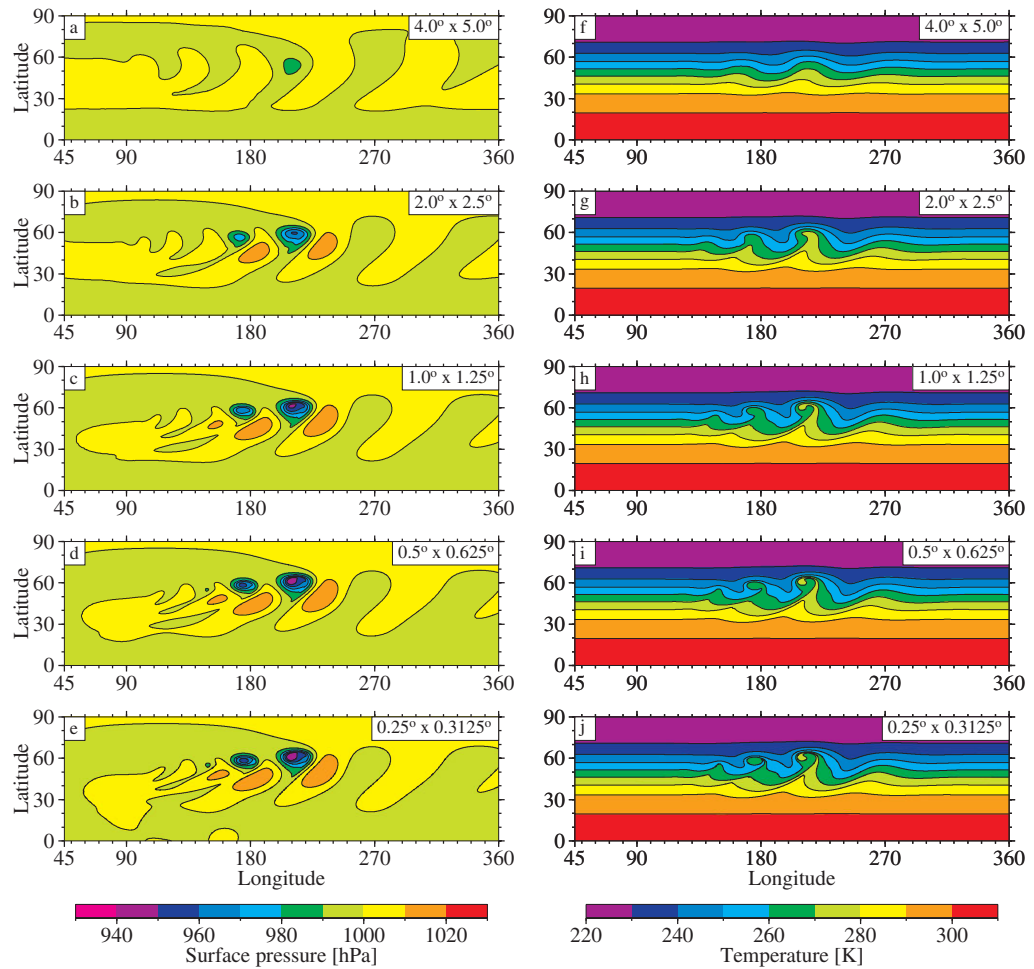
Figure 6.   Convergence of the finite-volume (FV) dynamical core with increasing horizontal resolution and 26 levels. The snapshots show the surface pressure (left column) and temperature at 850 hPa (right column) at day 9.

(*b*)   *Convergence*

Figure 6 compares the FV L26 solutions at day 9 for the range of horizontal resolutions considered. The 2° run does not capture the strength of the closed cells in the surface pressure and the fronts in the temperature field are not as sharp as in the higher-resolution solutions. The 1° simulation is visually very similar to the 0.5° run although the closed cell in the surface pressure at 170°E is not quite as deep. The 0.5° and 0.25° fields are visually almost indistinguishable, although differences at the smallest scales can be seen as expected.

(*c*)   *Model comparison*

Figure 7 compares the surface pressure and 850 hPa temperature fields at day 9 from each dynamical core at its second-highest horizontal resolution and 26 levels. The solutions are visually very similar and almost identical to the simulations at the highest or mid-range resolutions (Jablonowski and Williamson 2006). Note that the mid-range
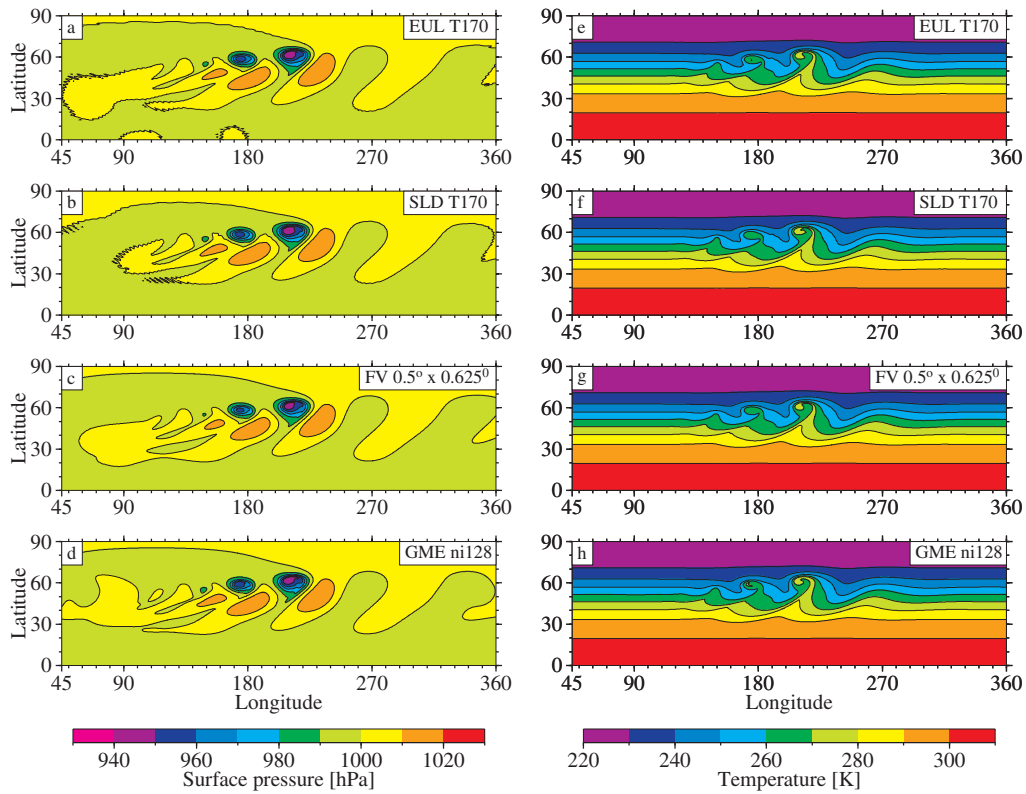
Figure 7. Comparison of the four dynamical cores at their second-highest horizontal resolutions and 26 levels. The snapshots show the surface pressure (left column) and temperature at 850 hPa (right column) at day 9. The models and resolutions are indicated in the legends.

resolutions (around T85) or even coarser resolutions (around T42) are typically used for climate simulations today. However, the closed cells in the surface pressure of the Eulerian, semi-Lagrangian and GME solutions are slightly deeper than those from the FV core. The two spectral cores also exhibit spectral noise where the fields become nearly flat. This is an artifact of the quadratic unaliased grid adopted for these tests. When the fields are reconstructed from spectral coefficients on a linear unaliased grid the spectral noise is basically eliminated. Similarly, when the semi-Lagrangian core is run with a linear truncation the spectral noise is minimal. However, the fields here are plotted on the quadratic grid as that is the grid on which the physics parametrizations are calculated in the complete Eulerian model, and thus such noise could be important.

Figure 8 displays the 850 hPa relative-vorticity fields of the four dynamical cores at days 7 and 9. As before, the models are shown at their second-highest resolutions with 26 levels. In these plots, the overall vorticity patterns of all dynamical cores agree reasonably well despite the fact that small-scale differences in the vorticity fields grow more rapidly. Therefore, the differences among the models become more evident in comparison to Fig. 7. At day 7, the FV dynamical core (Fig. 8(c)) exhibits a slightly weaker vorticity pattern in comparison to EUL, SLD and GME. The latter are visually very similar and, as argued before, almost identical to the simulations at the highest or mid-range resolutions (Jablonowski and Williamson 2006). The slightly weaker

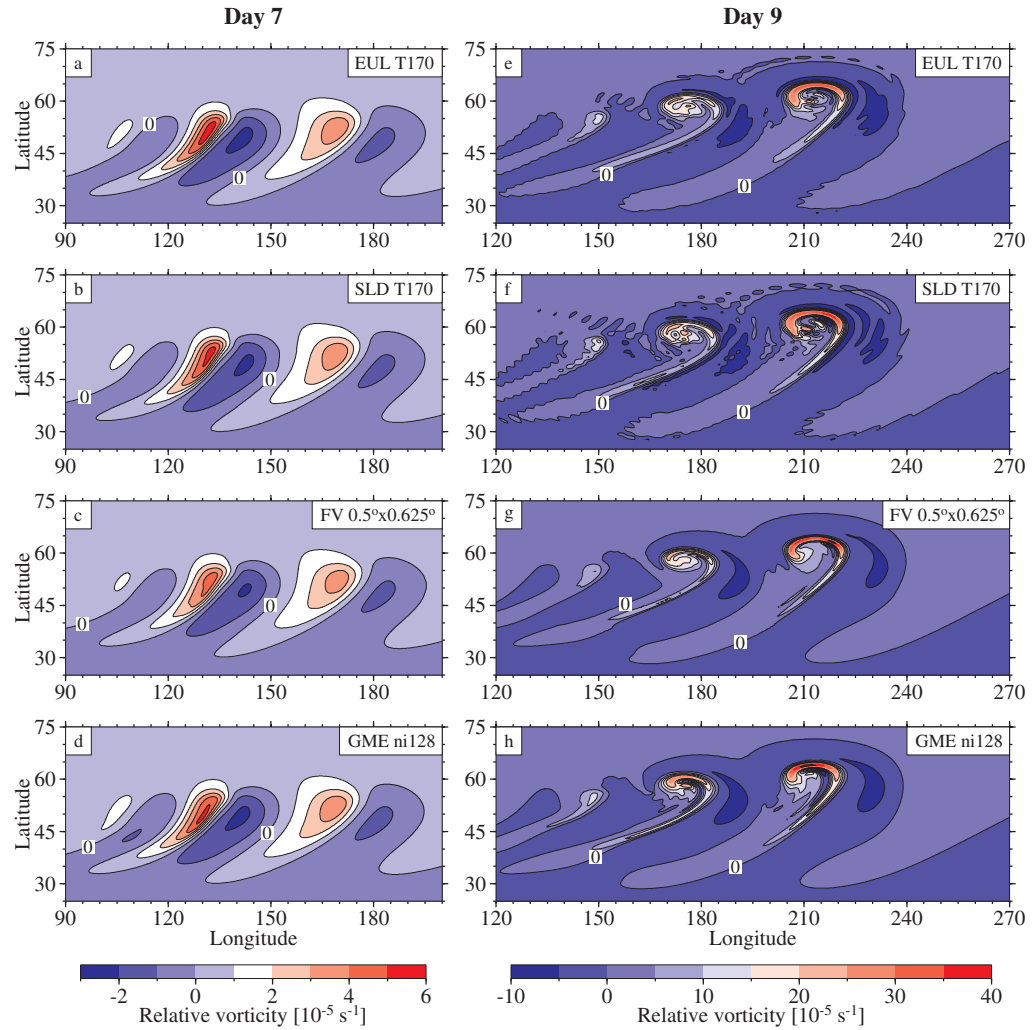**Day 7**   **Day 9**

Figure 8.   As in Fig. 7, but for the 850 hPa relative-vorticity fields at day 7 (left column) and day 9 (right column).

vorticity fields are caused by the monotonicity constraint in the FV dynamical core. This constraint adds nonlinear intrinsic dissipation in the regions where the monotonicity principles are locally violated. At day 9 the waves have almost reached their breaking point. Here, the details of the vorticity patterns differ, especially in the vicinity of the rolled up tongues of the vortices. Again, because of the quadratic unaliased grid the EUL and SLD dynamical cores (Figs. 8(e) and (f)) show spectral noise that disrupts the vorticity contours. Note that the vorticity gradients become extremely sharp after day 7 since we do not apply strong diffusion mechanisms. This enables us to assess the uncertainties in the high-resolution reference solutions for models in their operational configurations (section 5(e)). Any additional diffusion smoothes the contours in the EUL and SLD simulations (Jablonowski and Williamson 2006) but has the potential to eliminate the interesting small-scale steep-gradient features of the solution.

(*d*)   *Forecast skills after model day 10*

In this paper, the discussion of the baroclinic-wave test is focused on the first ten model days although integrations until day 30 have been performed. In general, it can be observed that the solutions of the four dynamical cores start diverging from each other after wave-breaking events set in around day 9. In particular, the differences in the solutions after day 10 become evident in the regions where the baroclinic waves have reached maturity and generated sharp fronts. Nevertheless, in other regions the forecasts after day 10 are still reliable. This is especially true for the newly developing systems downstream of the baroclinic wave. In these regions the initial development is in essence a linear phenomenon. As shown in Fig. 9, new systems develop every 2–3 days 60–90° downstream of the leading edge of the wave train. Figure 9 depicts the surface pressure maps of the four dynamical cores at the second-highest resolution at days 11, 14 and 16 in a northpolar-stereographic projection. The high pressure centres are marked with the symbol '+'. The evolution of the new downstream systems agrees well across the four model simulations even until day 16. This is in particular true for the EUL, SLD and FV model runs whereas the GME simulation is noticeably perturbed by the growing wave number 5 from day 14 onwards (cf. to section 4). Such a succession of new systems downstream of a baroclinic wave is also discussed in Simmons and Hoskins (1979).

(*e*)   *Surface pressure l_2 differences*

In this section we quantify the differences between solutions from different dynamical cores and at different resolutions using the l_2 difference norm of the surface pressure. Surface pressure is chosen since the norm can be calculated without vertical interpolation to a specified pressure surface. Also it does not require the calculation of a secondary variable such as vorticity which is not a prognostic variable for many dynamical cores. We have also calculated the l_2 norm for temperature and vorticity interpolated to selected pressure levels and the general conclusions are the same as those presented in the following. In addition, we examined the l_1 and l_∞ norms and again the general conclusions are the same (see Jablonowski and Williamson 2006). The qualitative difference between the different norms and different variables is the speed of growth in the uncertainty of the reference solutions, which will be discussed later. However, the relative errors for the cores compared to the reference solution uncertainties, which is emphasized in the following, is the same for all measures.

The l_2 difference norm is defined by

$$l_2(p_s(t)) = \left[\frac{1}{4\pi} \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \{p_{s_1}(\lambda, \varphi, t) - p_{s_2}(\lambda, \varphi, t)\}^2 \cos \varphi \, d\varphi \, d\lambda \right]^{1/2}$$
$$\approx \left[\frac{\sum_i \sum_j \{p_{s_1}(\lambda_i, \varphi_j, t) - p_{s_2}(\lambda_i, \varphi_j, t)\}^2 w_j}{\sum_i \sum_j w_j} \right]^{1/2}, \tag{16}$$

where the sums are taken over all points (λ_i, φ_j) of the global grid and p_{s_1} and p_{s_2} represent the two cases being compared. The integration weights w_j are given following Eq. (14). As before Eq. (16) assumes a tensor product grid with a uniform longitudinal grid interval. The model solutions must be interpolated to a common grid in order to calculate the norm. For each pair compared we interpolate from the coarser grid to the finer grid of the pair with horizontal spectral interpolation with no truncation. As outlined in section 4, the GME data are also biquadratically interpolated to a regular latitude–longitude grid before further spectral interpolations are made for the l_2 norm calculations. All spectral interpolations are computed with SPHEREPACK, a

Day 11   **Day 14**   **Day 16**

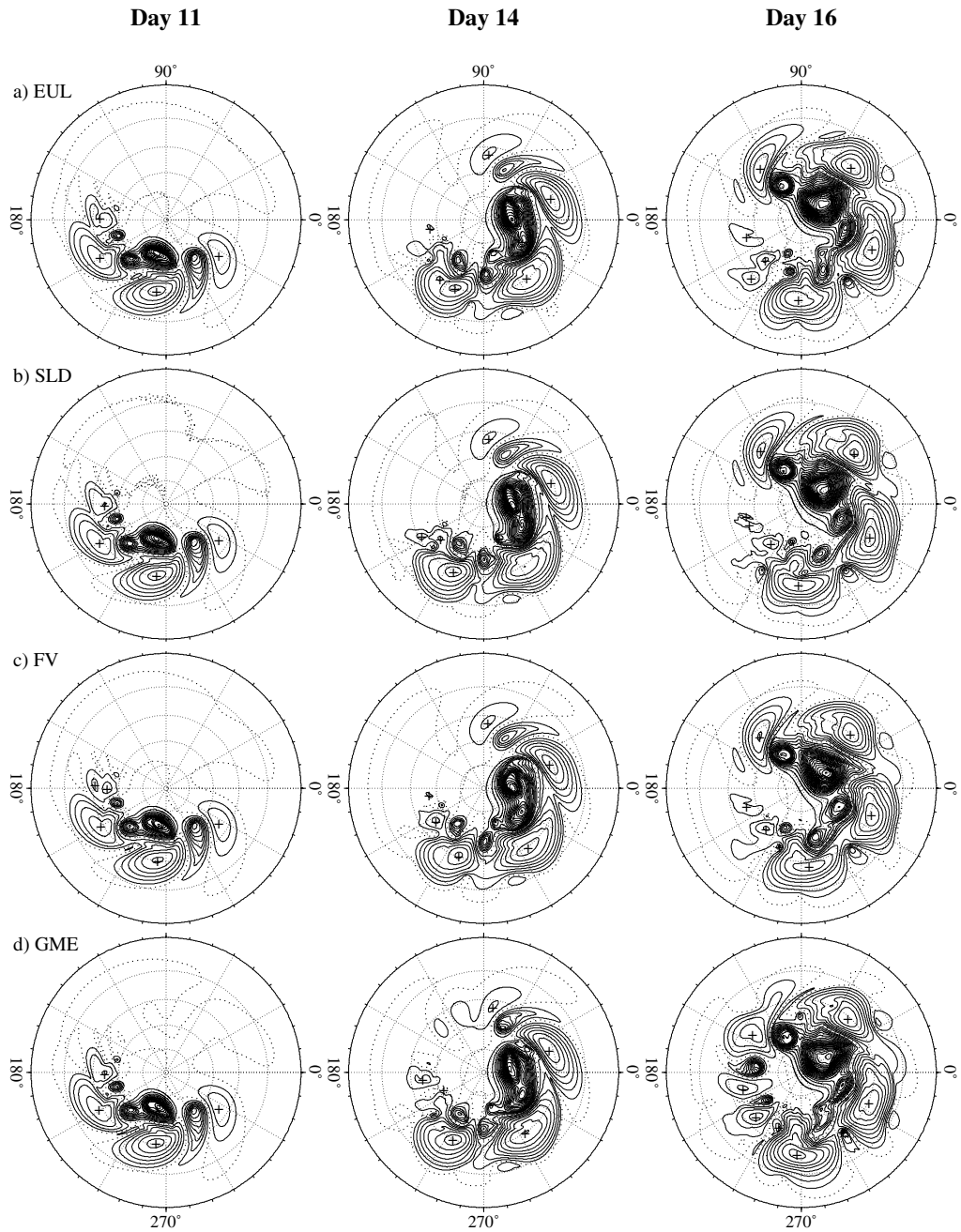a) EUL

b) SLD

c) FV

d) GME

Figure 9.   Northpolar-stereographic projection of the surface pressure at day 11 (left), day 14 (middle) and day 16 (right) for the four dynamical cores at their second-highest resolution (a) EUL T170, (b) SLD T170, (c) FV 0.5° × 0.625° and (d) GME ni128 with 26 levels. The contour interval is 5 hPa, the 1000 hPa contour line is dotted. The high pressure centres are labelled with the symbol '+'.
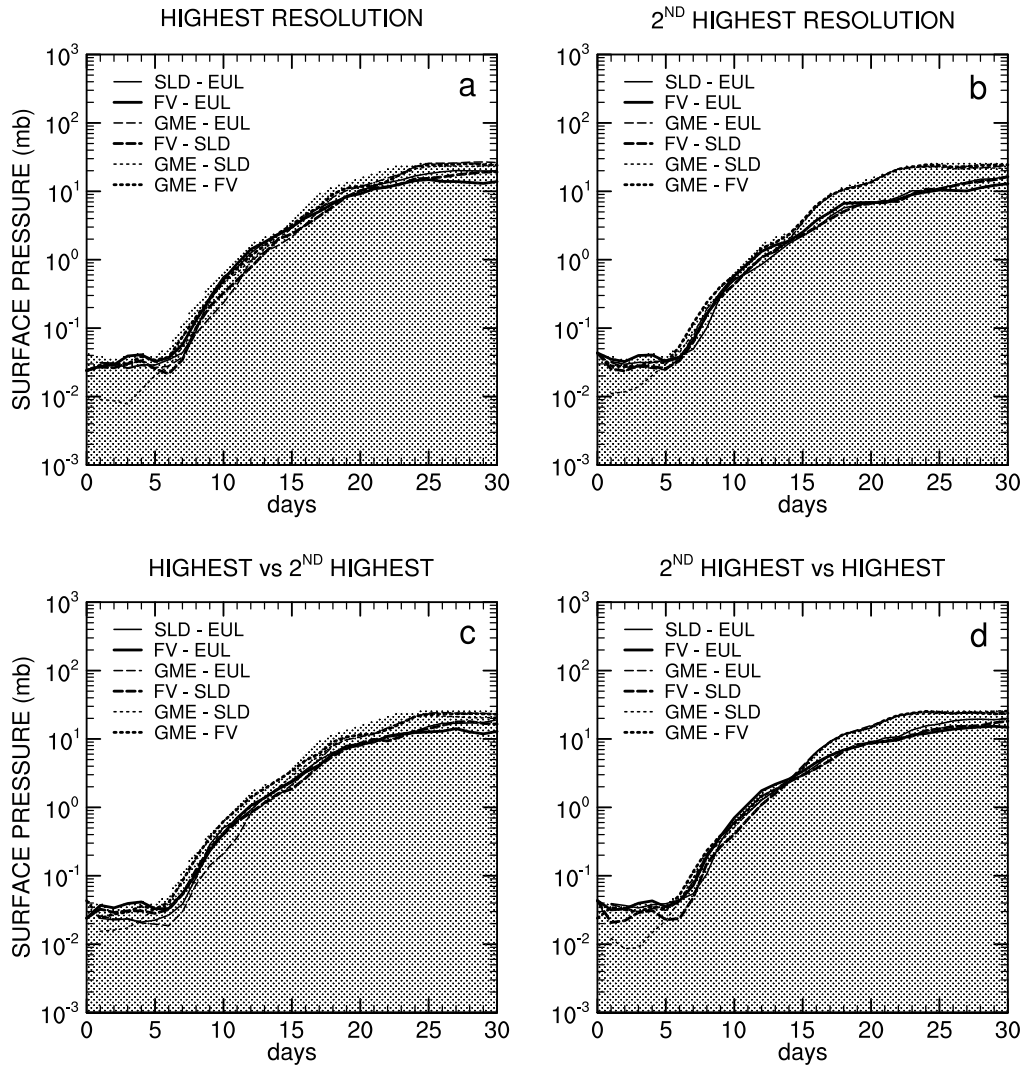
Figure 10.   Root-mean square l_2 norm of the surface pressure differences (hPa) between the four dynamical cores with 26 levels: (a) highest horizontal resolutions, (b) second-highest horizontal resolutions, (c) and (d) highest of one model compared to second-highest of the other.

publicly available analysis package for problems in spherical geometry (Adams and Swarztrauber 1997, 1999).

(i)   *Uncertainty in the reference solutions.* Figures 10(a) and (b) show the l_2 difference norms of the surface pressure for different models from the highest resolution class and from the second-highest resolution class, respectively. Many of the individual curves overlay each other and cannot be distinguished. Figures 10(c) and (d) show the differences between a model from the highest resolution class with a model from the second-highest resolution class. Again, many of the individual curves cannot be distinguished. The region below the maximum of all curves in Figs. 10(a)–(d) is stippled in all the plots. The significance of this stippled region will be discussed later.

The differences in the l_2 norms for days 0 to 6 are very small and are easily affected by the interpolations to a common grid needed to calculate the difference norms. The differences from day 0 to day 6 are relatively flat, rather than growing as would be expected. This indicates that the differences are affected by something in addition to the growth of the perturbation itself. During this period dispersion and damping of the gravity waves affect the initial growth characteristics. In addition, the interpolation adds to the difference and can dominate. Thus the l_2 norm does not represent the actual difference in the unstable growing mode between the models during this period. That difference is smaller. Before day 7 the differences are very small in practical terms as seen in Fig. 5 where, for example, the temperature contours on day 4 are practically straight lines.

For days 7 through 14 the l_2 differences in Fig. 10(b), which compare the different models in the second-highest resolution class, are similar to those in Fig. 10(a), which compare models in the highest resolution class. Since the uncertainties are not decreased going from second-highest to the highest resolution, these reference solutions are the best that can be obtained for the problem as posed here. The highest curves in Figs. 10(a) and (b) provide an indication of the uncertainty in the reference solutions. This is further verified by Figs. 10(c) and (d), which show the l_2 differences between a model from the highest resolution class with one from the second-highest class. The largest l_2 differences in Figs. 10(c) and (d) are very similar to the largest in Figs. 10(a) and (b), again indicating the level of uncertainty in the reference solutions. Thus the top of the stippled region defines the uncertainty of our reference solutions.

Eventually the l_2 differences grow until they saturate and the two solutions are related to each other as if they were chosen randomly from all possible states. Figure 10 indicates that that occurs around day 25 where the curves begin to level off. The asymptotic levels appear to differ somewhat between models. This apparent different level of saturation occurs because the perturbations in EUL, SLD and FV are in fact not fully developed in the southern hemisphere until after day 30, in contrast to GME. In general, the basic zonal state in the test case is symmetric about the equator. Thus any perturbation in the southern hemisphere will also grow. However, we impose an initial perturbation in the northern hemisphere only. Perturbations are introduced into the southern hemisphere by truncation errors and by gravity waves which arise from the geostrophic adjustment associated with the imposed unbalanced perturbation in the northern hemisphere and which propagate into the southern hemisphere. The perturbations introduced by the gravity waves are of similar amplitude for the different schemes at comparable resolutions, but the truncation errors are not. As indicated earlier by Figs. 3 and 4, the EUL and FV schemes approximate the zonal balanced state in the southern hemisphere rather well without any zonal variations in the truncation error. This is contrary to GME which truncation errors exhibit the built-in wave-number-5 grid structure. For practical application these errors are small at the higher resolutions. Nevertheless, in the southern hemisphere they grow fast and saturate sooner than the numerical errors introduced by the EUL, SLD and FV schemes. Notice that in fact the l_2 norms for those schemes which saturate later continue to grow to day 30 (the lower curves in Fig. 10) and eventually reach the same level as the others. As mentioned above, for reference in later figures, the region below all lines of Fig. 10 is stippled in all figures. The top of that stippled region provides an estimate of the uncertainty in the reference solution.

(ii)   *Horizontal resolution.* Figure 11 shows the convergence with increasing resolution for the (a) EUL, (b) SLD, (c) FV and (d) GME dynamical cores. Each is compared with
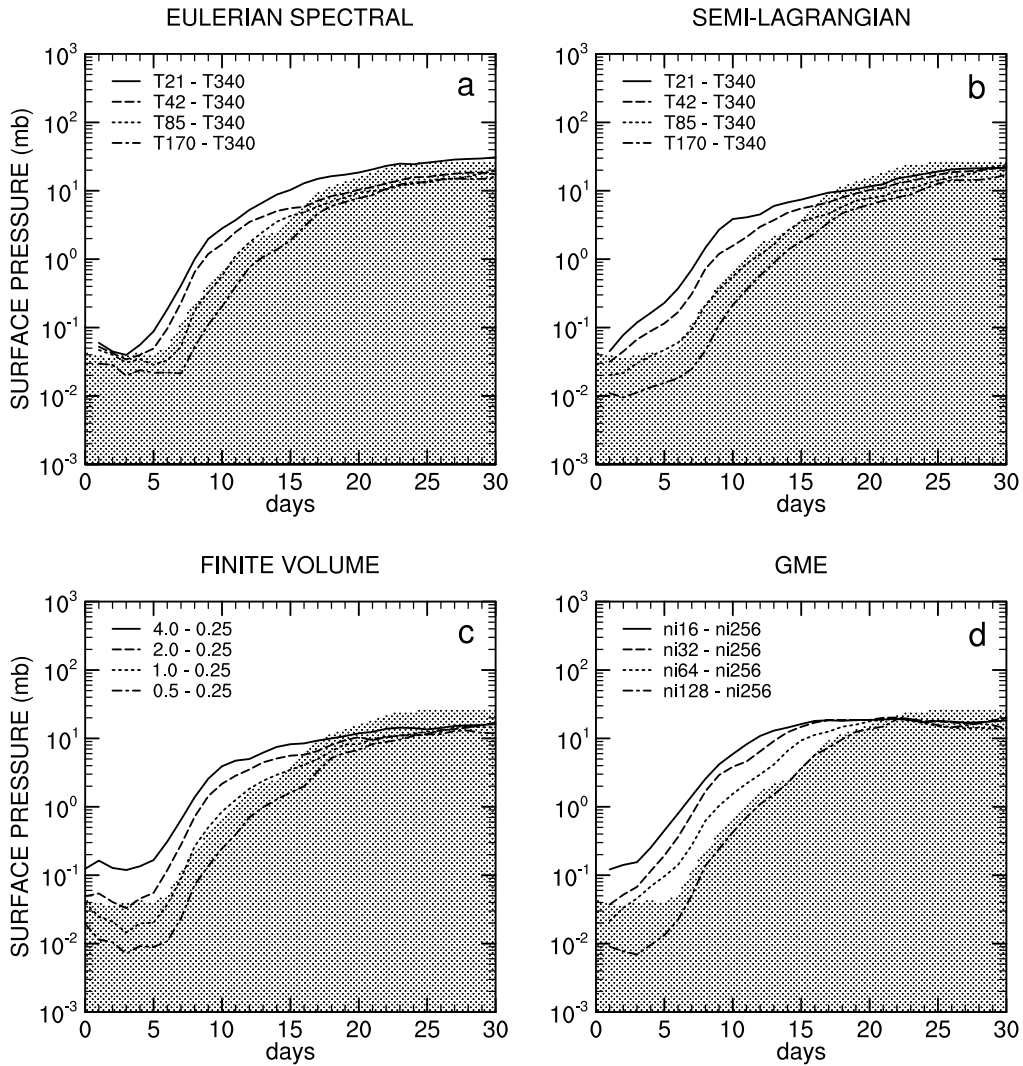
Figure 11. Root-mean square l_2 norm of the surface pressure differences (hPa) for each horizontal resolution compared to the highest horizontal resolution of the same model with 26 levels: (a) EUL, (b) SLD, (c) FV and (d) GME.

the reference solution produced by the highest resolution version of the same model. Figures 11(a) and (b) show that the T42 EUL and SLD do not capture the reference solution to within the uncertainty, but the T85 EUL and SLD do. Likewise, the 2.0 FV does not capture the reference solution but the 1.0 FV does. For the GME, ni64 does not capture the reference solution and the resolution must be at least ni128 to capture it to within the uncertainty.

Notice that the differences between higher-resolution solutions of the same model are well below the uncertainty obtained from Fig. 10, e.g. T170–T340 EUL and SLD, and 0.5–0.25 FV. The decrease in differences with increasing horizontal resolution is the expected behaviour for a single model. As horizontal resolution is increased, the errors decrease so the differences between solutions also decrease. On first consideration this

might be taken to imply our estimate of the uncertainty in the reference solutions is too large. Rather it implies a danger in using a single model to determine a reference solution for problems involving geostrophic adjustment and unstable states.

Figure 10 shows that the solutions of different models do not become closer to each other with increasing resolution once the resolutions exceed a certain limit, whereas in Fig. 11 the differences between solutions with the same model continue to decrease with increasing resolution. A possible explanation is that the truncation errors for a single model at different resolutions are more similar than the errors between models. Thus the differences involving a single model are initially smaller and remain smaller for a fixed elapsed time until saturation is reached. Although the perturbations introduced by the truncation errors of a single model are more similar and the evolving solutions are closer to each other, those solutions are not necessarily more accurate.

A more likely explanation for the convergence of inter-model differences is the different way each numerical method handles the geostrophic adjustment of the un-balanced initial perturbation. The models have different vertical approximations and different effective top boundary conditions, although the nominal level placement is the same in our experiments. With different vertical approximations the models are likely to project the imposed initial unbalanced perturbation differently onto the vertical modes, leading to a slightly different initial partition between the slow modes which grow with time, and the fast modes which are damped out, or at least remain small. This partition will be less affected by the horizontal resolution of a single scheme once it is fine enough with respect to the scales of the imposed initial wind perturbation.

The surface pressure differences represented by Figs. 10 and 11 are relatively large scale. The differences for all cases truncated to a T42 representation lie on top of the curves in the Figs. 10 and 11 (not shown). Thus at least T85 resolution is needed to calculate the T42 component of the solution to within the uncertainty of the reference solution for the EUL and SLD, while at least 1° is needed for the T42 component of the FV and ni128 for the T42 component of the GME.

(iii) *Phase error.* Figure 12 shows the surface pressure phase error in degrees longitude for the various models, positive for phase lag of the coarser resolution model. We define the phase error to be the distance the field must be shifted in longitude with respect to the reference solution to produce the minimum l_2 difference. The reference solution is taken to be the highest resolution solution from the same model. The phase error is determined to an integral number of grid intervals of the reference solution. A meaningful phase error can only be calculated when there is a reasonable correlation between the growing unstable structures in the two fields. In our experience, the perturbation is neither large enough nor well organized enough to yield a reasonable phase error before day 4. After day 10, the differences involving the coarser resolutions are too close to saturation (and randomly related) to yield a consistent phase error.

The finer-resolution solutions have insignificant phase error which is consistent with their matching the reference solution to within its uncertainty. The T21 EUL has about 2° phase error at day 10, and the 4.0 and 2.0 FV phase errors range from 1.5° to 2°. The T21 SLD phase error reaches 6.5° by day 10. The GME shows larger phase errors than the other three schemes at similar resolutions. The ni16, ni32 and ni64 GME phase errors are about 25°, 10°, and 3°, respectively, at day 10. The phase error is responsible for much of the total error in the ni64 GME solution, but not all. The 'amplitude' error (not shown), i.e. the minimum l_2 used to determine the phase error, is close to the reference solution uncertainty but still does not fall within it. Some of the total error is also attributable to the growth of the wave-number-5 perturbation.
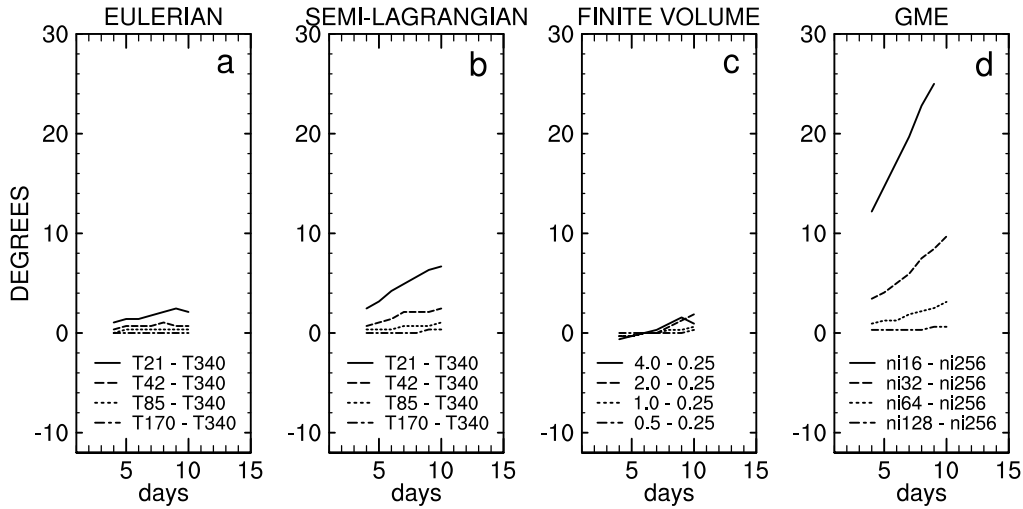
Figure 12. Surface pressure phase error (degrees longitude) for each horizontal resolution compared to the highest horizontal resolution for each model with 26 levels: (a) Eulerian, (b) semi-Lagrangian, (c) finite volume and (d) GME. The phase error is positive for a phase lag of the coarser resolution run.

(iv) *Vertical resolution.* In addition to the standard 26-level configuration, calculations have also been performed with 18 and 49 vertical levels (see Jablonowski and Williamson 2006) to address the impact of the vertical resolution on the reference state. For these simulations the mid-range horizontal resolutions are selected: T85 EUL and SLD, 1.0 FV and ni64 GME. Recall that the 26-level T85 EUL and SLD, and 1.0 FV have converged to within the uncertainty of the reference solutions, but the ni64 GME has not.

Figure 13 displays the sensitivity of the solution to the vertical-level distribution. Figure 13(a) shows that the differences between all 49-level and 26-level model simulations are within the uncertainty of the reference solutions. The same conclusion applies to the differences between the 26-level and 18-level model runs that are presented in Fig. 13(b). However, these differences are significantly smaller than the differences between the 49-level and 26-level model set-ups. The effect is attributable to the similar placement of the levels in the 18- and 26-level configurations. This is especially true for the tropospheric levels since the 26-level distribution was created from the 18-level set-up by adding levels around the tropopause with most of the new levels added between 200 hPa and 50 hPa (Williamson *et al.* 1998). Above 200 hPa the 26-level grid layers are therefore about half the thickness of the 18-level ones (in ln p). In the lower troposphere, however, the levels of the two grids are very close to each other. As a consequence, the truncation errors of the 18- and 26-level set-ups are indeed very similar since the growing modes of our test case have most of their amplitude at lower levels. In contrast, the level distribution of the 49-level configuration clearly differs from the 26-level set-up in all regions which is reflected by the slightly elevated levels of the l_2 error norms (Fig. 13(b)). For practical purposes though, these differences stay well below the uncertainty limit so that the 26-level distribution is an adequate representative of the reference state.

The discussion above is concentrated on the l_2 difference norm of surface pressure. As mentioned earlier, we have also calculated the l_1 and l_∞ norms for surface pressure and all three norms for 850 hPa temperature and vorticity. These are all presented in

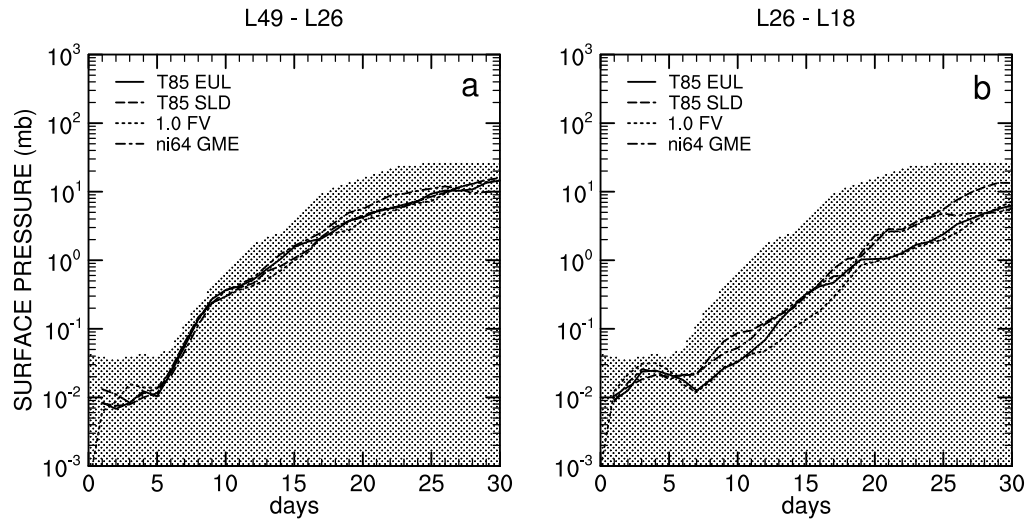L49 - L26   L26 - L18

Figure 13.   Root-mean square l_2 norm of the surface pressure differences (hPa) between each model (at its mid-range horizontal resolution) with different vertical resolutions: (a) 49 levels versus 26 levels, and (b) 26 levels versus 18 levels.

Jablonowski and Williamson (2006). The general characteristics of Figs. 10–13 are observed with each norm and variable. The qualitative difference between the different norms and different variables is the speed of growth in the uncertainty of the reference solutions. The temperature difference growth is comparable to that of surface pressure while the vorticity difference growth is faster than either. The l_1 difference grows more slowly than the l_2 while the l_∞ grows faster. However, the relative errors for the cores compared to the reference-solution uncertainties is the same for all measures and the conclusions drawn above are the same with all norms. The only differences are the time-scales in approaching saturation.

6.   CONCLUSIONS

An idealized baroclinic-wave test case for dry hydrostatic and non-hydrostatic dynamical cores of GCMs has been developed that is deterministic, easy to use, relevant to typical atmospheric phenomena and applicable to a wide variety of model formulations and grids. The two-step test strategy first assesses the ability of the models to maintain a steady state before an overlaid perturbation triggers the evolution of a baroclinic wave over the course of several days. The initial conditions are given in form of analytic expressions in pressure-based vertical coordinates. They can also be accurately derived for models with height-based or isentropic vertical coordinates. The zonally symmetric initial state is a steady-state solution of the adiabatic inviscid primitive equations that resembles the climatic conditions of a winter hemisphere.

The test has been applied to four very different hydrostatic dynamical cores at varying horizontal and vertical resolutions. In particular, two spectral transform models (Eulerian (EUL) and semi-Lagrangian (SLD)) and two grid-point models (Finite Volume (FV) and DWD's finite difference, icosahedral model (GME)) with different numerical schemes and computational grids have been tested. All models provide independent reference solutions. The uncertainty in the reference solutions is established by comparing the solutions of all four dynamical cores at high horizontal resolutions.

This approach not only takes the convergence-with-resolution characteristics of a single model into account but also assesses the cross-model differences to gain confidence in the reference states. Here, it is strongly emphasized that the maximum spread of all high-resolution reference simulations defines the uncertainty of the reference state despite the fact that all single-model assessments stay well below the uncertainty limit. Once a single model converges to within the uncertainty, finer horizontal resolutions do not provide a better estimate of the reference solution. Thus the test case tells the user at what resolution a new scheme is as good as those considered here and something about its behaviour at coarser resolutions, but not at finer resolutions.

For the assessment of the steady-state solution, we suggest computing two error norms for the zonal-wind field which can be immediately compared to the analytic solution. For the baroclinic-wave simulations, we recommend comparing the readily available surface pressure fields along with their l_2 difference norms. The same convergence analyses with other derived and interpolated quantities, like the temperature or vorticity fields on the 850 hPa pressure level, lead to identical conclusions (Jablonowski and Williamson 2006). We also have assessed the convergence characteristics of the eddy momentum and heat fluxes but do not show them in this paper. We found that such an analysis leads to the same conclusions as drawn from the surface pressure here. In addition, it is informative to visually compare the aforementioned flow fields, especially shortly before wave-breaking events set in around day 9. This reveals possible noise issues related to the numerical methods.

In summary, it has been shown that all four dynamical cores with 26 levels converge to within the uncertainty at the horizontal resolutions EUL T85, SLD T85, FV $1.0° \times 1.25°$ and GME ni128. The error at coarser resolutions are similar to each other, i.e. EUL T42, SLD T42, FV $2.0° \times 2.5°$ and GME ni64. Still coarser resolutions are not capable of capturing the growing wave at all. The higher resolution for GME was partly needed to reduce the phase error and wave-number-5 signal that is evident in the coarser GME simulations. From the computational viewpoint, the semi-Lagrangian method has been the most cost-effective approach while matching the results within the uncertainty limit. In addition, the test reveals some intrinsic diffusive characteristics of the numerical schemes since we aim at keeping the effects of the explicit diffusion mechanisms (if any) small. Note that the test also exposes programming or design flaws that might be hidden in real climate and weather prediction applications. In particular, we discovered a less than ideal energy fixer and vector projection method in an earlier version of the semi-Lagrangian dynamical core which led to clear outliers in the convergence diagrams. These findings will be documented in a forthcoming paper.

In the future, we would like to establish a standard test suite for dynamical cores that fosters the developments and intercomparisons of new numerical schemes and gridding options. We hope that this test finds broad acceptance in the modelling community to become part of such a test series. As a starting point for model intercomparions we make our reference solutions in netCDF format available online (contact the authors for details). In particular, the datasets contain the surface pressure fields for all four models at the highest and second-highest resolutions so that modelling groups can readily compute the l_2 norms against the established reference states.

<div align="center">APPENDIX</div>

Iterative method for dynamical cores with z- or Θ-based vertical coordinates

In this paper the initial conditions are shown in closed form for pressure-based vertical coordinates with $\eta = \sigma = p/p_s$. For dynamical cores with z- or Θ-based vertical coordinates a root-finding technique (Newton's method) is suggested. It determines the corresponding η-level for any given height z or isentropic level Θ with high accuracy. This technique therefore avoids vertical interpolations of the initial dataset. The iterative method is given by

$$\eta^{n+1} = \eta^n - \frac{F(\lambda, \varphi, \eta^n)}{(\partial F/\partial \eta)(\lambda, \varphi, \eta^n)}, \tag{A.1}$$

where the superscript $n = 0, 1, 2, 3, \ldots$ indicates the iteration count.

For a given z-level the functions F and $\partial F/\partial \eta$ are determined by

$$F(\lambda, \varphi, \eta^n) = -gz + \Phi(\lambda, \varphi, \eta^n), \tag{A.2}$$

$$\frac{\partial F}{\partial \eta}(\lambda, \varphi, \eta^n) = -\frac{R_d}{\eta^n} T(\lambda, \varphi, \eta^n), \tag{A.3}$$

where T and Φ are shown in Eqs. (6) and (7), respectively.

If isentropic vertical coordinates with given Θ-levels are utilized, the function F is defined by

$$F(\lambda, \varphi, \eta^n) = -\Theta + (\eta^n)^{-\kappa} T(\lambda, \varphi, \eta^n), \tag{A.4}$$

with $\kappa = R_d/c_p$. The specific heat of dry air at constant pressure c_p is set to 1004.5 J kg^{-1}K^{-1}. Using $R_d = 287$ J kg^{-1}K^{-1}, κ is identical to 2/7. For $\partial F/\partial \eta$ it follows that

$$\frac{\partial F}{\partial \eta}(\lambda, \varphi, \eta^n) = G(\eta^n) - \frac{3\pi u_0}{4R_d(\eta^n)^\kappa} \cos^{1/2} \eta_v^n$$

$$\times \left[\frac{3\pi}{2} u_0 \eta^n A \sin^2 \eta_v^n \cos^{1/2} \eta_v^n - (2Au_0 \cos^{3/2} \eta_v^n + Ba\Omega) \right. \tag{A.5}$$

$$\left. \times \left\{ (1 - \kappa) \sin \eta_v^n + \frac{\pi \eta^n}{4 \cos \eta_v^n} (2 - 3\sin^2 \eta_v^n) \right\} \right],$$

with

$$A = -2\sin^6 \varphi \left(\cos^2 \varphi + \frac{1}{3} \right) + \frac{10}{63},$$

$$B = \frac{8}{5} \cos^3 \varphi \left(\sin^2 \varphi + \frac{2}{3} \right) - \frac{\pi}{4},$$

$$\eta_v^n = (\eta^n - \eta_0)\pi/2.$$

As before, η_0 is set to 0.252. The vertical derivative of the horizontal-mean potential temperature $\langle \Theta(\eta) \rangle = \eta^{-\kappa} \langle T(\eta) \rangle$ is symbolized by G. It is given by

$$G(\eta^n) = T_0 \left(\frac{R_d \Gamma}{g} - \kappa \right) (\eta^n)^{(R_d \Gamma / g - \kappa - 1)}, \qquad \text{for } \eta_s \geq \eta^n \geq \eta_t, \quad \text{(A.6)}$$

$$G(\eta^n) = T_0 \left(\frac{R_d \Gamma}{g} - \kappa \right) (\eta^n)^{(R_d \Gamma / g - \kappa - 1)}$$
$$- \frac{\Delta T}{(\eta^n)^{(\kappa+1)}} (\eta_t - \eta^n)^4 \{ \kappa \eta_t + \eta^n (5 - \kappa) \}, \quad \text{for } \eta_t > \eta^n. \quad \text{(A.7)}$$

The starting value $\eta^0 = 10^{-7}$ is recommended for all Newton iterations following Eq. (A.1). Then Newton's method converges for any given height z below $z_{max} = 100$ km or any given potential temperature Θ below $\Theta_{max} = 15\,000$ K. Note that the convergence is only achieved if η^0 is greater than zero and physically lies above the uppermost model level. If models with higher model tops are used, η^0 needs to be decreased. We have observed that this iterative technique converges within a maximum of 25 iterations. Then the absolute error $|\eta - \eta^n|$ is decreased to machine precision (10^{-14} for double-precision arithmetics). Most often, the convergence is already achieved with fewer iterations. The resulting η-level can now be used for the computation of the analytic initial conditions at the location (λ, φ, η).

REFERENCES

Adams, J. C. and Swarztrauber, P. N. — 1997 — SPHEREPACK 2.0: A model development facility. NCAR Technical Note NCAR/TN-436+STR. National Center for Atmospheric Research, PO Box 3000, Boulder CO80307-3000, USA

— 1999 — SPHEREPACK 3.0: A model development facility. *Mon. Weather Rev.*, **127**, 1872–1878

Bates, J. R. and Li, Y. — 1997 — 'Simulation of stratospheric vortex erosion using three global shallow water numerical models'. Pp. 55–74 in Numerical methods in atmospheric and oceanic modelling: The André J. Robert Memorial Volume. *Atmos.-Ocean.*, Eds. C. A. Lin, R. Laprise and H. Ritchie

Boer, G. J. and Denis, B. — 1997 — Numerical convergence of the dynamics of a GCM. *Clim. Dyn.*, **13**, 359–374

Bonaventura, L. — 2004 — 'Development of the ICON dynamical core: Modelling strategies and preliminary results'. Pp. 197–213 in Proceedings of the ECMWF/SPARC workshop on modelling and assimilation for the stratosphere and tropopause, 23–26 June 2003. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK

Colella, P. and Woodward, P. R. — 1984 — The Piecewise Parabolic Method (PPM) for gas-dynamical simulations. *J. Comput. Phys.*, **54**, 174–201

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Kiehl, J. T., Briegleb, B. P., Bitz, C. M., Lin, S.-J., Zhang, M. and Dai, Y. — 2004 — Description of the NCAR Community Atmosphere Model (CAM3.0). NCAR Technical Note NCAR/TN-464+STR. National Center for Atmospheric Research, PO Box 3000, Boulder CO80307-3000, USA

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Briegleb, B. P., Bitz, C. M., Lin, S.-J., Zhang, M. and Dai, Y. — 2006 — The formulation and atmospheric simulation of the Community Atmosphere Model: CAM3. *J. Climate*, **19**, 2144–2161

Fournier, A., Taylor, M. A. and Tribbia, J. J. 　2004　The Spectral Element Atmospheric Model: High-resolution parallel computation and response to regional forcing. *Mon. Weather Rev.*, **132**, 726–748

Gates, W. L. 　1995　'Proceedings of the first international AMIP scientific conference'. 15–19 May 1995, Monterey, CA, USA. World Climate Research Program Report WCRP-92, WMO/TD-No. 732, Geneva

Gates, W. L., Boyle, J. S., Covey, C. C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler, P. J., Hnilo, J. J., Marlais, S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E. and Williams, D. N. 　1999　An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Am. Meteorol. Soc.*, **80**, 29–55

Giraldo, F. X. and Rosmond, T. E. 　2004　A Scalable Spectral Element Eulerian Atmospheric Model (SEEAM) for NWP: Dynamical Core Tests. *Mon. Weather Rev.*, **132**, 133–153

Held, I. M. and Suarez, M. J. 　1994　A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteorol. Soc.*, **75**, 1825–1830

Holton, J. R. 　1992　*An introduction to dynamic meteorology*, third edition. Academic Press, London

Hoskins, B. J. and Simmons, A. J. 　1975　A multi-layer spectral model and the semi-implicit method. *Q. J. R. Meteorol. Soc.*, **101**, 637–655

Jablonowski, C. 　2004　'Adaptive grids in weather and climate modeling'. PhD thesis, University of Michigan

Jablonowski, C. and Williamson, D. L. 　2006　A baroclinic wave test case for dynamical cores of general circulation models: Model intercomparisons. NCAR Technical Note NCAR/TN-469+STR, available online at http://www.library.ucar.edu/uhtbin/hyperion-image/ DR000790 or http://www.engin.umich.edu/~cjablono/ publications/TN-469+STR.pdf

Juckes, M. N. and McIntyre, M. E. 　1987　A high-resolution one-layer model of breaking planetary waves in the stratosphere. *Nature*, **328**, 590–596

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

Kiehl, J. T., Hack, J. J., Bonan, G. B., Boville, B. A. and Williamson, D. L. 　1998　The National Center for Atmospheric Research Community Climate Model: CCM3. *J. Climate*, **11**, 1131–1149

Lin, S.-J. 　2004　A 'vertically Lagrangian' finite-volume dynamical core for global models. *Mon. Weather Rev.*, **132**, 2293–2307

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

　1997　An explicit flux-form semi-Lagrangian shallow-water model on the sphere. *Q. J. R. Meteorol. Soc.*, **123**, 2477–2498

Machenhauer, B. 　1979　'The spectral method'. Pp. 121–275 in *Numerical methods used in atmospheric models*, Vol. 2. Ed. A. Kasahara. GARP Publications Series No 17, WMO and ICSU, Geneva

Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W. and Baumgardner, J. 　2002　The Operational Global Icosahedral-Hexagonal Gridpoint Model GME: Description and high-resolution tests. *Mon. Weather Rev.*, **130**, 319–338

Monaco, A. V. and Williams, R. T. 　1975　An atmospheric global prediction model using a modified arakawa differencing scheme. Technical Report NPS-51WU75041, Dept. of Meteorology, Naval Postgraduate School, Monterey, CA

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

　1959　Numerical integration of the primitive equations on the hemisphere. *Mon. Weather Rev.*, **87**, 333–345

Polvani, L. M. and Saravanan, R. 　2000　The three-dimensional structure of breaking Rossby waves in the polar wintertime stratosphere. *J. Atmos. Sci.*, **57**, 3663–3685

Polvani, L. M., Scott, R. K. and Thomas, S. J.	2004	Numerically converged solutions of the global primitive equations for testing the dynamical core of atmospheric GCMs. *Mon. Weather Rev.*, **132,** 2539–2552
Simmons, A. J. and Burridge, D. M.	1981	An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Weather Rev.*, **109,** 758–766
Simmons, A. J. and Hoskins, B. J.	1975	A comparison of spectral and finite-difference simulations of a growing baroclinic wave. *Q. J. R. Meteorol. Soc.*, **101,** 551–565
	1976	Baroclinic instability on the sphere: Normal modes of the primitive and quasi-geostrophic equations. *J. Atmos. Sci.*, **33,** 1454–1477
	1977	Baroclinic instability on the sphere: Solutions with a more realistic tropopause. *J. Atmos. Sci.*, **34,** 581–588
	1979	The downstream and upstream development of unstable baroclinic waves. *J. Atmos. Sci.*, **36,** 1239–1254
Tomita, H. and Sato, M.	2004	A new dynamical framework of nonhydrostatic global model using the icosahedral grid. *Fluid Dyn. Res.*, **34,** 357–400
US Standard Atmosphere	1976	US Government Printing Office, Washington D.C.
Williamson, D. L. and Olson, J. G.	1994	Climate simulations with a semi-Lagrangian version of the NCAR Community Climate Model. *Mon. Weather Rev.*, **122,** 1594–1610
Williamson, D. L. and Rasch, P. J.	1989	Two-dimensional semi-Lagrangian transport with shape-preserving interpolation. *Mon. Weather Rev.*, **117,** 102–129
Williamson, D. L., Drake, J. B., Hack, J. J., Jakob, R. and Swarztrauber, P. N.	1992	A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.*, **102,** 211–224
Williamson, D. L., Olson, J. G. and Boville, B. A.	1998	A comparison of semi-Lagrangian and Eulerian tropical climate simulations. *Mon. Weather Rev.*, **126,** 1001–1012