A new spherical shallow-water model: building in wave–vortex decomposition

David G. Dritschel

Mathematical Institute, University of St Andrews
St Andrews KY16 9SS, United Kingdom
dgd@mcs.st-and.ac.uk

ABSTRACT

A new approach to simulating spherical shallow-water flows is described and illustrated. The approach exploits the distinction between slow vortical motions and fast gravity waves at leading order. Contour advection is used for the vortical part, specifically for the potential vorticity. This allows an accurate representation of the vortical dynamics at modest grid resolution. New variables are used to represent the gravity-wave part, at leading order in Froude number. These variables represent the departure from geostrophic balance. Their use requires the solution of several spatially elliptic equations for the original variables (fluid depth and velocity). These equations come from the definitions of the new variables and the potential vorticity. They simplify to balance conditions for the vortical flow when the "wave" variables are neglected. Simulations of rotating turbulence at moderate to high Froude numbers demonstrate that these new variables help to preserve the underlying vortical flow in these flows. Moreover, they reduce the erroneous numerical generation of gravity waves arising from a misrepresentation of the underlying vortical flow.

1 Introduction

The effects of rotation and stratification in the Earth's atmosphere and oceans cause large and intermediate-scale motions flows to be dominated by sub-inertial frequency vortical motions. Superimposed on this state are higher frequency inertia–gravity oscillations or 'gravity waves', of generally much weaker amplitude. The origins of these two types of motion are fundamentally distinct, the first induced by, and entirely determined by, the distribution of potential vorticity, and the second whatever is *not* induced by potential vorticity, i.e. whatever is left over.

A fundamental problem is that this 'wave–vortex' decomposition is non-unique and is mathematically subtle. Nevertheless, progress has been made employing approximate decomposition techniques, e.g. based on the linearised equations about a state of rest, or based on reductions of the full nonlinear equations which filter the gravity waves (cf. Norbury & Roulestone 2002a,b). Much is made of these techniques in, for example, weather forecast initialisation and in interpreting flow behaviour.

Little if any use, however, is made of these techniques in simulating atmospheric and oceanic flows. Models which filter the gravity waves also filter a significant portion of the desired vortical motions, simply because the wave–vortex decomposition is not sufficiently precise. Such models are generally not suitable for weather forecasting or ocean modelling, and instead the full or 'gravity-wave-permitting' equations are employed. While these models are capable of modelling both the vortical and the wave parts of the flow, the numerical approaches taken often unnecessarily limit the accuracy of both parts, and in particular the interaction between the two.

This problem can be overcome, it appears, by weaving wave–vortex decomposition directly into the numerical approach (Dritschel & Mohebalhojeh 2000, Mohebalhojeh & Dritschel 2000, 2001, 2004, Dritschel & Viúdez 2003). A significant gain in numerical accuracy, for *both* the vortical and the wave parts of the flow, results from (1) directly using potential vorticity (PV) as one of the prognostic variables, and (2) transforming the remaining variables to new variables which vanish in the limit of rapid rotation and strong stratification (small

Rossby and Froude numbers). These new variables represent the *departure* from the 'balance' between flow variables assumed in the wave–vortex decomposition. They represent the 'imbalance' to leading order.

In the single-layer rotating shallow-water context considered here, if one wants robustness as well as accuracy across the whole range of Froude numbers, then the choice of new variables which appears to work best — for constant background rotation — is $\delta = \boldsymbol{\nabla} \cdot \boldsymbol{u}$, the velocity divergence, and $\gamma = \boldsymbol{\nabla} \cdot \boldsymbol{a}$, the acceleration divergence (Dritschel & Mohebalhojeh 2000, Mohebalhojeh & Dritschel 2000). This choice also appears to work best in multi-layer shallow-water flows (Mohebalhojeh & Dritschel 2004). Both δ and γ vanish in the limit of small Froude number for the vortical component of the flow.

Here, we extend this approach to variable background rotation, specifically to single-layer shallow-water flows on a sphere. The approach is applied to rotating turbulence at moderate to high Froude numbers. Using a novel wave–vortex decomposition procedure called the *Optimal PV* (OPV) balance (Viúdez & Dritschel 2004), we demonstrate that these turbulent flows remain close to balance if they start close to balance. That is, there is practically no generation of gravity waves.

In the next section, the mathematical and numerical structure of the approach is described. It is then applied to rotating shallow-water turbulence in §3. Here, we focus on the degree to which the balance is preserved, and the properties (magnitude, form, frequency spectra) of the imbalance, i.e. of the gravity waves. Conclusions are given in §4.

2 Mathematical and numerical set up

2.1 Re-casting the shallow-water equations on a sphere

In vector form, the shallow-water equations may be written

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + f\boldsymbol{k} \times \boldsymbol{u} = -c^2 \boldsymbol{\nabla}\tilde{h} \tag{1}$$

$$\frac{\partial \tilde{h}}{\partial t} + \boldsymbol{\nabla} \cdot (\tilde{h}\boldsymbol{u}) = 0 \tag{2}$$

(where $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{u} \cdot \boldsymbol{\nabla}$) and have the same form in both planar and spherical geometry. In spherical geometry, $f = 2\Omega \sin \phi$, where Ω is the planetary rotation rate and ϕ is latitude. Use is made of the dimensionless depth anomaly $\tilde{h} = (h - H)/H$, where H is the mean fluid depth, while $c^2 = gH$ (a mean-square gravity wave speed).

For now we have neglected all non-conservative and diabatic processes to focus on the basic dynamical properties of shallow-water flows, and on our ability to model them accurately.

Under these conditions, the potential vorticity

$$\Pi = \frac{\zeta + f}{1 + \tilde{h}} \tag{3}$$

is materially conserved,

$$\frac{\mathrm{D}\Pi}{\mathrm{D}t} = 0. \tag{4}$$

The connection between potential vorticity (PV) and the material or Lagrangian description of fluid flow makes PV the natural variable for representing vortical motion (Hoskins *et al.* 1985, Mohebalhojeh 2002, Norbury & Roulestone 2002a,b, Dritschel & Viúdez 2003), though precisely how it does so depends on the balance relations chosen in the wave–vortex decomposition to carry out the 'PV inversion'. Here, therefore, we take PV as one of the prognostic variables for computing the flow evolution. But then we must decide what the other two prognostic variables should be.

We could suppress any one of the three scalar equations implied in (1) & (2), but there are other possibilities. A common one is to take the divergence of (1), and thus form a single equation for $\delta = \boldsymbol{\nabla} \cdot \boldsymbol{u}$. Our prognostic variable set would then be (h, δ, Π). This is close to the set (h, δ, ζ) commonly used in spectral models, except ζ does not have the special material conservation property enjoyed by Π, nor the intimite connection with the vortical flow.

The set (h, δ, Π), used previously in the f-plane shallow-water context ($f = $ constant; Dritschel *et al.* 1999), directly exploits these properties but does not exploit the underlying near-geostrophic balance expected at small Froude numbers (Mohebalhojeh & Dritschel 2000). To do this also, we must avoid evolving h. Instead, we *diagnose h* from the definition of Π (3) together with the definition of the variable chosen to replace h. We can also replace both h and δ. In previous work (Dritschel & Mohebalhojeh 2000, Mohebalhojeh & Dritschel 2000, 2001, 2004, Dritschel & Viúdez 2003), we have advocated choosing the new variables from associated balance conditions. These conditions come in pairs, and may take the form $F(h, \boldsymbol{u}) = 0$ and $G(h, \boldsymbol{u}) = 0$. When applied, they eliminate two time derivatives from the equations of motion, leaving only the PV evolution equation. Solving $F(h, \boldsymbol{u}) = 0$ and $G(h, \boldsymbol{u}) = 0$, together with the definition of PV, itself of this form, provides the original variables h & $\boldsymbol{u}$ in terms of Π. This is a balanced flow. However, we need not apply $F(h, \boldsymbol{u}) = 0$ and $G(h, \boldsymbol{u}) = 0$ to get the benefits of balance — instead we can choose the new variables to be F and G, and write evolution equations for them. The original equations are then not approximated, but merely transformed.

In the present work, concerned with the shallow-water equations in spherical geometry, we take $F = \boldsymbol{\nabla} \cdot \boldsymbol{u} = \delta$ and $G = \boldsymbol{\nabla} \cdot \boldsymbol{a} = \gamma$, where $\boldsymbol{a} = \mathrm{D}\boldsymbol{u}/\mathrm{D}t = -f\boldsymbol{k} \times \boldsymbol{u} - c^2 \boldsymbol{\nabla}\tilde{h}$ from (1). This choice is motivated by the excellent performance of the set (δ, γ, Π) in the f-plane single- and multi-layer contexts.

The evolution equations for δ and γ are found by substituting (1) and (2) into $\delta_t = \boldsymbol{\nabla} \cdot \boldsymbol{u}_t$, and $\gamma_t = \boldsymbol{\nabla} \cdot \boldsymbol{a}_t$, using $\boldsymbol{a}_t = -f\boldsymbol{k} \times \boldsymbol{u}_t - c^2 \boldsymbol{\nabla}\tilde{h}_t$. The result is

$$\delta_t = \gamma - |\boldsymbol{u}|^2 - 2[u_\phi(u_\phi + \zeta) + v_\phi(v_\phi - \delta)] - \boldsymbol{\nabla} \cdot (\delta\boldsymbol{u}) \tag{5}$$

$$\gamma_t = c^2 \nabla^2 \{\boldsymbol{\nabla} \cdot [(1+\tilde{h})\boldsymbol{u}]\} + 2\Omega B_\lambda - \boldsymbol{\nabla} \cdot (Z\boldsymbol{u}) \tag{6}$$

where λ is longitude, u & v are the eastwards and northwards velocity components, $B \equiv c^2\tilde{h} - \frac{1}{2}|\boldsymbol{u}|^2$ (Bernoulli pressure), and $Z = f(\zeta + f)$.

The equation for δ is written in a non-standard way to avoid any spatial derivatives on ζ (which are infinite for a PV jump, allowed in the numerical method described below), and second derivatives on $\boldsymbol{u}$ (which are also infinite for a PV jump). These derivatives cancel, mathematically, but may not numerically. It is unsafe to count on cancellations between large terms in a numerical model. The same strategy was also adopted in the f-plane context (Dritschel *et al.* 1999).

These equations, coupled with PV conservation (4), constitute the three evolution equations for the (δ, γ, Π) set. However, to evaluate the tendencies, we also need the fields $\tilde{h}$, u and v. These are provided by the definitions of δ, γ and Π, as follows. As usual, we express the velocity in terms of a streamfunction ψ and a divergence potential χ via

$$\boldsymbol{u} = \boldsymbol{k} \times \boldsymbol{\nabla}\psi + \boldsymbol{\nabla}\chi \tag{7}$$

and find the potentials by inverting

$$\nabla^2\psi = \zeta \quad \& \quad \nabla^2\chi = \delta. \tag{8}$$

Since δ in one of the prognostic variables, χ can be found directly. However, ζ depends on $\tilde{h}$ through the definition of PV (3):

$$\zeta = (1+\tilde{h})\Pi - f. \tag{9}$$

So, we need to find $\tilde{h}$ before we can invert ζ. To this end, consider the definition of γ,

$$\begin{aligned} \gamma &= -\boldsymbol{\nabla} \cdot (f\boldsymbol{k} \times \boldsymbol{u} + c^2\boldsymbol{\nabla}\tilde{h}) \\ &= f\zeta - \beta u - c^2\nabla^2\tilde{h} \end{aligned} \tag{10}$$

where $\beta = f_\phi = 2\Omega \cos\phi$. If we substitute ζ from (9) and re-arrange terms, we get

$$c^2 \nabla^2 \tilde{h} - f\Pi\tilde{h} = f(\Pi - f) - \beta u - \gamma, \qquad (11)$$

which appears to be a Helmholtz equation for $\tilde{h}$, except that u depends also on $\tilde{h}$ implicitly through (7), (8) and (9). Hence, it is seen that the equations to determine ψ and $\tilde{h}$ are coupled — they must be solved simultaneously. However, they are linear, implying there is a unique solution for any given δ, γ and Π. In practise, these equations are solved iteratively and convergence is rapid even in cases where $f\Pi < 0$ over parts of the domain.

In summary, given the fields δ, γ and Π at a given instant of time, the first step is to invert (8), (9), and (11) for ψ, χ (hence u, v by (7)) and $\tilde{h}$. This allows one to compute the tendencies and thus to advance δ, γ and Π to the next instant of time.

Details of numerical method will be provided in a forthcoming paper. The main novel features are (1) a mixed spatial representation, spectral in longitude and grid point in latitude, with *equal* numbers of grid points in longitude and latitude $n_\lambda = n_\phi$; (2) 2nd-order finite-differences in latitude, and (3) a contour representation for PV. The latter feature has been implemented in a wide variety of contexts (cf. Dritschel & Ambaum (1997), Dritschel & Viúdez (2003) & refs. therein), and spherical geometry poses no special difficulties if one represents points on contours in a (3D) Cartesian basis (Dritschel, 1988). Features (1) and (2) give rise to tridiagonal problems (for each longitudinal wavenumber) for inverting the operators ∇^2 and $c^2\nabla^2 - f^2$, and for discontinuous PV, the tridiagonal solution proves to be as accurate as a corresponding fully spectral solution — at a fraction of the cost. Moreover, $c^2\nabla^2 - f^2$, which is used to solve (11) iteratively, does not permit a direct spectral inversion.

All latitudinal derivatives on the velocity components are avoided by using the definitions of δ and ζ and longitudinal derivatives instead, i.e.

$$
\begin{aligned}
u_\phi &= r^{-1}(zu + v_\lambda) - \zeta \\
v_\phi &= r^{-1}(zv - u_\lambda) + \delta
\end{aligned} \qquad (12)
$$

where $z = \sin\phi$ and $r = \cos\phi$. The longitudinal derivatives are carried out spectrally, and more accurately than can be done by finite differences.

Time stepping is done by a standard semi-implicit leap-frog scheme with a minimal Robert-Asselin time filter. Here, we have kept the time step below the CFL value to accurately treat gravity waves.

3 Shallow-water turbulence

3.1 Initial conditions

Three simulations are described next, which may be classified as weak, moderate and strong turbulence. The strength of the turbulence, or of the circulation, is controlled by the r.m.s. PV anomaly ϖ_{rms}, where $\varpi \equiv \Pi - 1$. For the three simulations, $\varpi_{rms}/2\Omega = \frac{1}{6}$, $\frac{1}{3}$, and $\frac{1}{2}$. In all other respects, the simulations are identical. We take $\Omega = 2\pi$ so that a unit of time corresponds to one day, and the mean wave speed $c = 4\pi/3$ (giving a mean 'Rossby radius' $L_R = c/2\Omega = 1/3$). The basic grid resolution is $n_\lambda = n_\phi = 128$, and the time step is $\Delta t = 0.004$ (while $\Delta t_{CFL} = \Delta\phi/c \approx 0.0058$). For the PV contours, a large-scale length $L = 1/3$ and a point spacing parameter of $\mu = 0.1716171$ is used to control the point spacing. PV contours are regularised by 'surgery' on a scale of $\Delta\phi/10$ (well below the grid scale). A fine-grid to coarse-grid ratio $m_g = 4$ is used for converting the PV contours to gridded values. These contour parameters are detailed in Dritschel & Ambaum (1997).

Each simulation starts with the same spatial structure of ϖ, randomly and isotropically distributed but correlated over a scale $L_c = 0.1$. This anomaly is then ramped up over a 20-day ramp period ($\Delta\tau_I = 20$), during which significant advection of the original PV contours occurs. This establishes the initial flow fields at '$t = 0$'
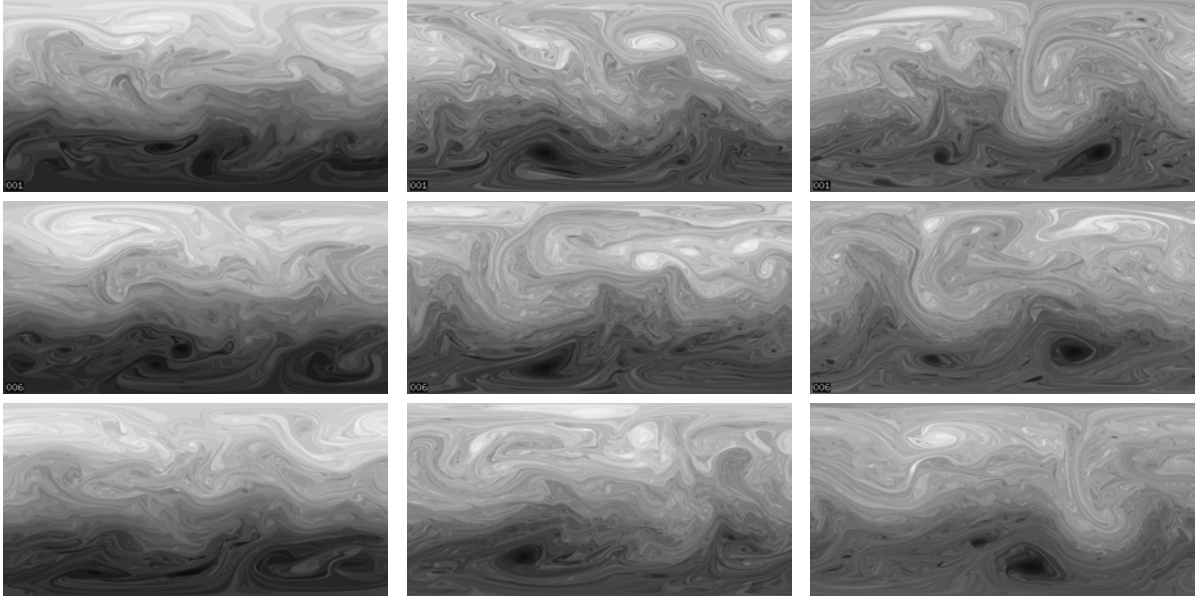
Figure 1: *Evolution of the PV field (downward) for the three cases $\varpi_{rms}/2\Omega = \frac{1}{6}$ (left column), $\varpi_{rms}/2\Omega = \frac{1}{3}$ (middle column), $\varpi_{rms}/2\Omega = \frac{1}{2}$ (right column). Times shown are $t = 0$, 2.5 and 5 days. The PV field is rendered in a linear grey scale, with dark shading corresponding to negative values and light shading to positive values. The PV contours used in the simulations are too dense to be sensibly visualised. Each rectangle corresponds to the full domain, $-\pi \leq \lambda \leq \pi$ and $-\frac{1}{2}\pi \leq \phi \leq \frac{1}{2}\pi$.*

(see Dritschel & Viúdez (2003) for details of this method). In the strong turbulence case, the turbulence is already decaying (measured by the number of nodes representing the PV contours), while it is just before peak complexity in the moderate turbulence case, and still growing in complexity in the weak turbulence case.

3.2 Vortex dynamics

Each simulation was run for a 5 day period, which is analysed in detail next. We first describe the evolution of the basic flow statistics and the PV field. The key conserved quantities, mass, angular momentum, and energy, vary in the numerical simulations by no more than a few percent at this resolution (128×128). This level of variation is expected in a complex turbulent flow. The mass, by construction, is exactly conserved.

The maximum Froude number Fr_{max}, where $\text{Fr} \equiv |\boldsymbol{u}|/c(1+\tilde{h})^{\frac{1}{2}}$, varies only weakly and exhibits time-averaged values of 0.202, 0.375 and 0.570 for the three cases, $\varpi_{rms}/2\Omega = \frac{1}{6}$, $\frac{1}{3}$ and $\frac{1}{2}$. The maximum Rossby number Ro_{max}, where $\text{Ro} \equiv |\zeta|/2\Omega$, varies more strongly and exhibits time-averaged values of 0.511, 0.878 and 1.33. At these Froude and Rossby numbers, one might not expect gravity waves to be insignificant. Yet the results presented next indicate the the vortical flow is highly dominant in all cases.

The evolution of the PV field is shown in figure 1 for the three cases (left, middle, right columns). The flows range from wavy-turbulence, consisting of nonlinear waves and few vortices except near the poles (left), to strong turbulence, with many active vortices entangling the background PV field associated with the planetary rotation (right). This complexity is the result of repeated vortex interactions and breaking nonlinear Rossby waves, and is typical of weakly-dissipative flows.

As deduced from the wave–vortex decomposition (see next subsection), a major part of the other fields like the height anomaly $\tilde{h}$, divergence δ and ageostrophic vorticity γ may be attributed to the PV. Figure 2 shows the evolution of $\tilde{h}$ for the three cases, corresponding to the PV in figure 1. The depth field exhibits only broad-scale features. This structure may be traced back to (11), in which the effect of inverting $c^2\nabla^2 - f\Pi$
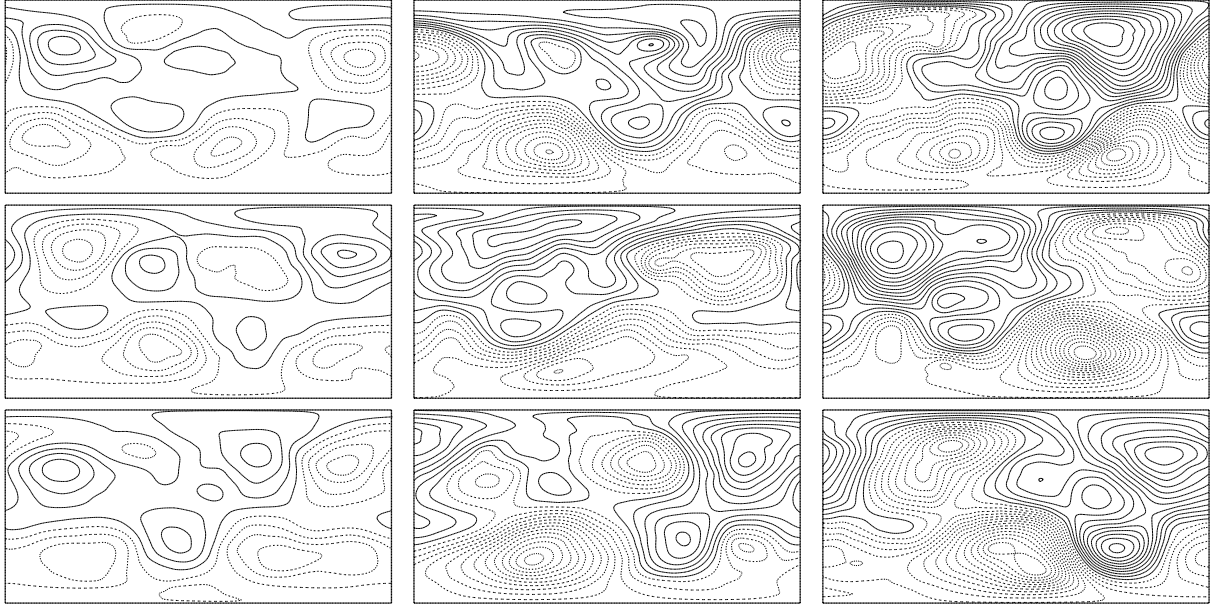
Figure 2: Evolution of the anomalous depth field $\tilde{h}$ corresponding to the PV evolution shown in figure 1. The contour interval is 0.02 *(dimensionless).*

smears any fine-scale structure which may be present on the right-hand side. If we scale (11) by $(2\Omega)^2$ to make it dimensionless, then $c^2\nabla^2 \to L_R^2\nabla^2$, where $L_R = c/2\Omega$ is the Rossby radius. Hence, scales smaller than L_R are smoothed by the operator inversion. On the right-hand side, the source is dominated by terms involving the PV anomaly, $f(\Pi - f)/(2\Omega)^2 = (\varpi/2\Omega)\sin\phi = \mathscr{O}(\text{Ro})$, and the zonal velocity, $\beta u/(2\Omega)^2 = (u/2\Omega)\cos\phi = \mathscr{O}(\text{Ro})$. The dimensionless ageostrophic vorticity, $\gamma/(2\Omega)^2$ contributes only weakly: in the time mean, $\gamma_{\text{rms}}/(2\Omega)^2 = 0.0144, 0.0388$ and 0.0717 for the three cases, whereas for the Rossby number, $\text{Ro}_{\text{rms}} = \zeta_{\text{rms}}/2\Omega = 0.137, 0.224$ and 0.291. These values imply $\gamma_{\text{rms}}/(2\Omega)^2 \approx \alpha\text{Ro}_{\text{rms}}^2$, with $\alpha = 0.81 \pm 0.04$ for the simulation parameters chosen. No frontal structures in $\tilde{h}$ can be seen in these simulations (at least over these time scales), evidence of an underlying balanced evolution.

The importance of ageostrophic motions can be measured by the magnitude of the acceleration $\boldsymbol{a} = \text{D}\boldsymbol{u}/\text{D}t = -f\boldsymbol{k} \times \boldsymbol{u} - c^2\boldsymbol{\nabla}\tilde{h}$, whose divergence gives γ and whose curl gives $f\delta + \beta v$. *A priori*, one may expect $\boldsymbol{a}$ to be largest at the equator since there $f = 0$. This is confirmed in figure 3, showing the dimensionless field $|\boldsymbol{a}|/(2\Omega)^2$. Although $|\boldsymbol{a}|/(2\Omega)^2 \ll 1$, using $\boldsymbol{a} = \boldsymbol{0}$ to define a balanced flow is mathematically inconsistent. This is because the meridional velocity field v cannot be found from $fv = c^2r^{-1}\tilde{h}_\lambda$ at the equator in general. Alternatively, using $\delta = \gamma = 0$ to define the balanced flow *is* consistent. Then one must solve (11) together with (7), (8) and (9) for $\tilde{h}$ and $\boldsymbol{u}$ (with $\delta = \gamma = 0$). The resulting meridional velocity field is compared with the exact field for case 2 ($\varpi_{\text{rms}}/2\Omega = \frac{1}{3}$) in figure 4 — note the contour interval is 20 times less for the difference field. The latter is remarkably broad-scale, and mainly associated with the divergence field, which is missing in the 'crude' balanced assumed. Overall, the time-mean r.m.s. difference is 5.68% of v_{rms}. In case 1, this drops to 4.72% while in case 3, this rises to 6.35%. How this compares with the 'optimal PV balance' mentioned in §2.4 is discussed below.

3.3 Gravity waves

The wave–vortex decomposition or optimal PV (OPV) balance procedure (Viúdez & Dritschel 2004) is used next to examine the gravity-wave or imbalanced component in the three turbulence simulations above. The OPV balance procedure involves integrating the full equations — only for a ramped PV anomaly — over a
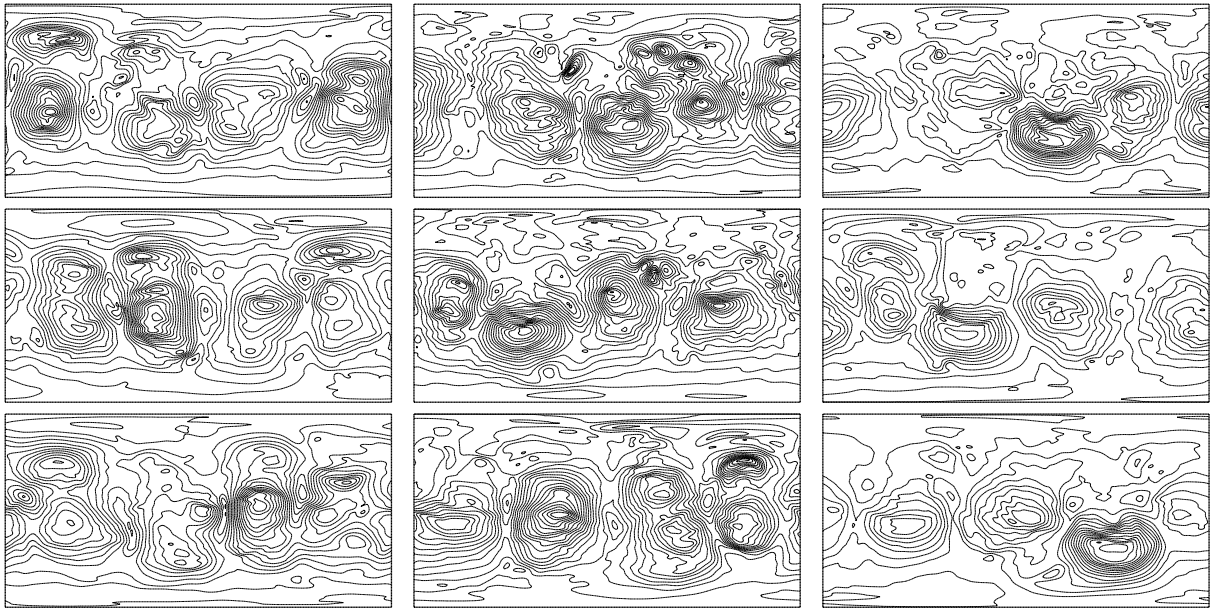
Figure 3: Evolution of the dimensionless acceleration magnitude $|\boldsymbol{a}|/(2\Omega)^2$ corresponding to the PV evolution shown in figure 1. The contour interval is 0.001 in case 1 (left column), 0.002 in case 2 (middle column) and 0.005 in case 3 (right column).
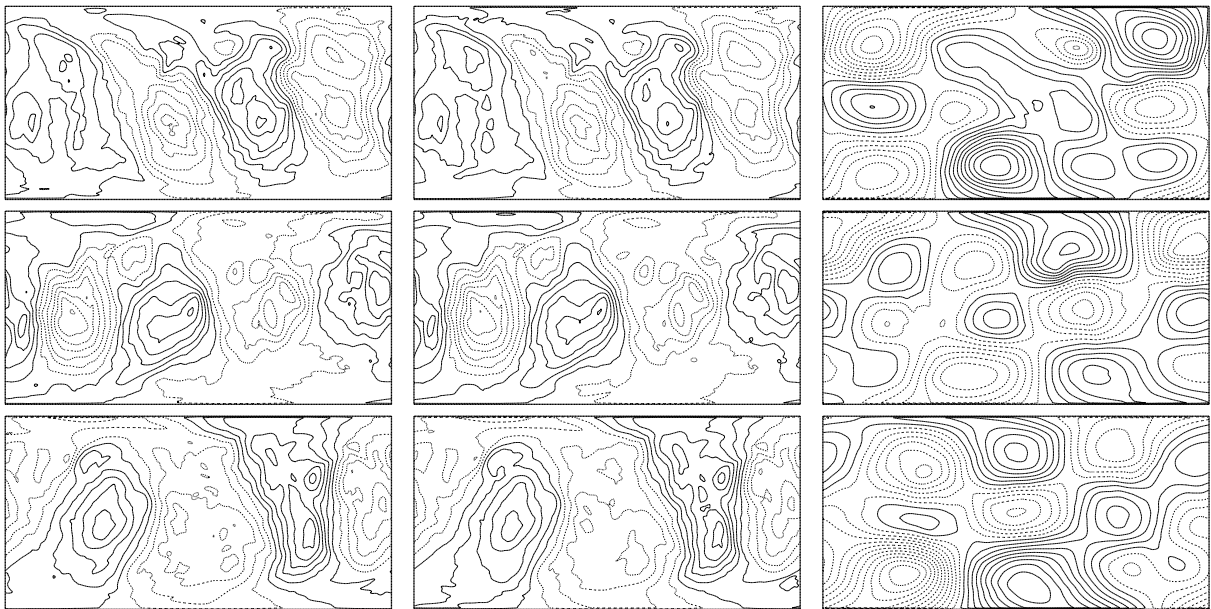

Figure 4: Evolution of the meridional velocity field v for case 2; the left column shown the full field at times $t = 0$, 2.5 and 5; the middle column shows the 'balanced' field associated with the conditions $\delta = \gamma = 0$; and the right column shows the difference between the two, or the 'imbalanced' field. The contour interval is 0.2 in the left and middle columns, and just 0.01 in the right column.
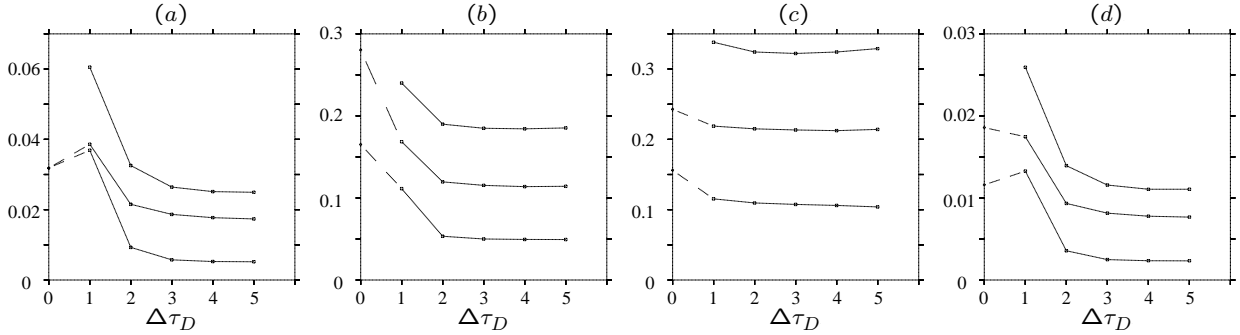
Figure 5: Measures of imbalance as a function of the ramp period $\Delta\tau_D$ for the three simulations (lower to upper curves for cases 1, 2 and 3 in order). (a) $\tilde{h}_{i\,rms}/\tilde{h}_{rms}$, (b) $\delta_{irms}/\delta_{rms}$, (c) $\gamma_{irms}/\gamma_{rms}$, and (d) $|\boldsymbol{u}_i|_{rms}/|\boldsymbol{u}|_{rms}$. All results are for OPV balance except those plotted as diamonds along $\Delta\tau_D = 0$. The latter are for first-order δ–γ balance ($\delta_t = \gamma_t = 0$).

time interval $\Delta\tau_D$. The procedure seeks to find the 'base configuration' of PV contours which evolves, while ramping up the PV anomaly as in initialisation, *into* the current or actual set of PV contours. In practise, this has to be done iteratively, in a cycle of backward and forward integrations, but convergence is strong in all cases.

We begin by considering how the diagnosed imbalance depends on $\Delta\tau_D$ for the different flows. The results are also compared with more traditional assessments of imbalance using specific balance relations (McIntyre & Norton 1999, Mohebalhojeh & Dritschel 2001). Here we obtain results for the first two members of the δ–γ hierarchy, whose balance relations are $\delta = \gamma = 0$ (cf. figure 4 at the end of the last section) and $\delta_t = \gamma_t = 0$. Next, frequency spectra for both the balanced and imbalanced components are computed from highly-sampled data over a 2-day time window.

We first assess the dependence on the diagnostic period $\Delta\tau_D$. The diagnosis of balance was carried out using $\Delta\tau_D = 1$, 2, 3, 4 and 5 days. Typically, 6-8 forward and backward integration cycles were required for convergence of the OPV balance procedure (here when the maximum pointwise difference in $\tilde{h}$ at the end of successive cycles is less than 10^{-7}). Each case was analysed at 0.5 day intervals from $t = 0$ to 5 (11 samples). This produces a balanced flow, indicated by a subscript b, for each $\Delta\tau_D$. The imbalanced component, indicated by a subscript i, is just the difference between the actual and the balanced flows, e.g. $\delta_i = \delta - \delta_b$. This also depends on $\Delta\tau_D$. The effectiveness of the wave–vortex decomposition procedure may be judged by the degree to which the flow can be attributed to the PV alone. This is equivalent to minimising the imbalance.

The magnitude of the imbalance is quantified by measuring the time-mean r.m.s. values of $\tilde{h}_i$, δ_i, γ_i and $\boldsymbol{u}_i$ relative to the actual flow, i.e. $\delta_{irms}/\delta_{rms}$. The results are summarised in figure 5, which shows these measures of imbalance as a function of the ramp period $\Delta\tau_D$, for all three simulations. Included also are the results for the first-order δ–γ balance ($\delta_t = \gamma_t = 0$) plotted along $\Delta\tau_D = 0$ (when available). The results for the zeroth-order δ–γ balance ($\delta = \gamma = 0$) are far off the scale and are given instead in Table 1. Across the board, the OPV balance procedure proves most effective, at least for $\Delta\tau_D > 1$ day. And, the best results are found for $\Delta\tau_D = 4$ to 5 days, though there is only marginal improvement over $\Delta\tau_D = 3$ days. The dependence on $\Delta\tau_D$ is remarkably similar for all three flows, which span a wide range of Froude numbers from 0.17 to 0.65. The first-order δ–γ balance performs as well as the OPV balance for $\Delta\tau_D = 1$ day, apart from the divergence measure. However, the balancing procedure fails to converge for case 3, despite using the OPV balanced fields as initial guesses. This lack of convergence also occurs in the related f-plane context (Mohebalhojeh & Dritschel, 2001), where it was shown that the order of the balance which minimises measures of imbalance is inversely related to the Froude number. In other words, higher-order balance relations such as $\delta_{tt} = \gamma_{tt} = 0$ might perform better than first-order balance for small Froude numbers as in case 1, but may not for moderate Froude numbers. The OPV balance does not appear to be as limited. Furthermore, its convergence with increasing $\Delta\tau_D$ supports the assertion that the residual flow, the imbalance, can be regarded as pure gravity waves to a high-degree of

$\varpi/2\Omega$	$\tilde{h}_{\mathrm{i\,rms}}/\tilde{h}_{\mathrm{rms}}$	$|\boldsymbol{u}_{\mathrm{i}}|_{\mathrm{rms}}/|\boldsymbol{u}|_{\mathrm{rms}}$
1/6	0.1076	0.0608
1/3	0.1419	0.0740
1/2	0.2616	0.1040

Table 1: *Measures of imbalance for zeroth-order δ–γ balance ($\delta = \gamma = 0$) for the three cases. All values are ratios of time-mean r.m.s. field values. Note that $\delta_{\mathrm{irms}}/\delta_{\mathrm{rms}}$ and $\gamma_{\mathrm{irms}}/\gamma_{\mathrm{rms}}$ are both equal to 1 and so are not listed.*
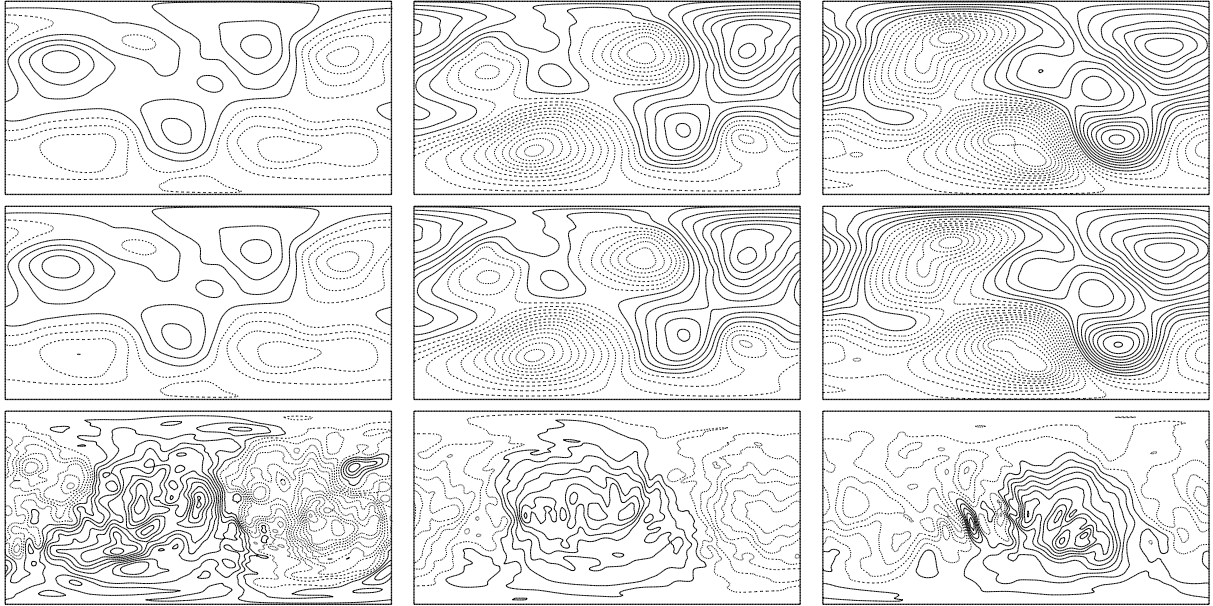

Figure 6: *Fields of $\tilde{h}$ (top), $\tilde{h}_{\mathrm{b}}$ (middle) and $\tilde{h}_{\mathrm{i}}$ (bottom) at $t = 5$ days for case 1 (left), case 2 (middle) and case 3 (right). The contour interval is 0.02 for $\tilde{h}$ and $\tilde{h}_{\mathrm{b}}$, while it is 0.00005, 0.0005 and 0.005 for $\tilde{h}_{\mathrm{i}}$ from left to right in the bottom row.*

accuracy. This residual flow cannot be attributed to PV.

We examine next the spatial structure of the balanced and imbalanced fields. The balanced fields generally differ little from the full fields shown in the previous section with the exception of δ and γ in case 3. This is illustrated in figures 6 and 7 for the depth anomaly and divergence fields, together with their balanced and imbalanced components (top, middle and bottom rows) for cases 1, 2 and 3 (left, middle and right columns) at $t = 5$ days. Only in case 3 and only for δ do differences between the full and balanced fields become significant. The imbalanced fields reveal wave-train like structures, radiating from particular locations, together with what appears to be a large-scale equatorial Kelvin wave in $\tilde{h}_{\mathrm{i}}$. This is consistent with the frequency spectra analysed next.

Frequency spectra for the full, (OPV) balanced, and imbalanced fields were obtained over the first two days of each simulation at a sampling frequency of 0.02 days. The spectra of $\tilde{h}$ and δ are shown in figure 8 — note the logarithmic scale. The full and balanced depth spectra are virtually indistinguishable at all frequencies, even for case 3 (right). The imbalanced spectrum (lower curve) is several orders of magnitude weaker, though it exhibits a peak at the inertial frequency (1 day^{-1}), which is probably related to the presence of an equatorial Kelvin wave in $\tilde{h}_{\mathrm{i}}$. Perhaps surprisingly, the balanced spectra have long tails extending into high frequencies — the notion of a frequency separation between balanced and imbalanced motions is clearly not supported. That is, vortical motions are not confined to low frequencies, nor are the gravity waves confined to high frequencies.
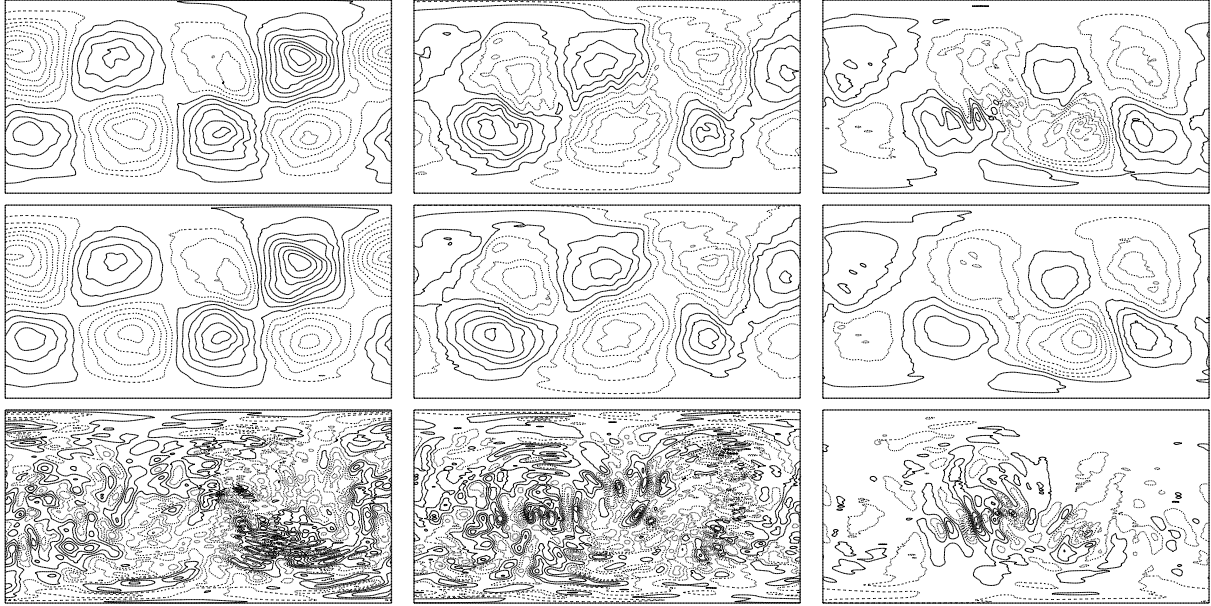
Figure 7: Fields of $\delta/2\Omega$ (top), $\delta_b/2\Omega$ (middle) and $\delta_i/2\Omega$ (bottom) at $t = 5$ days for case 1 (left), case 2 (middle) and case 3 (right). The contour intervals are 0.002, 0.002 *and* 0.0002 *in the left column,* 0.005, 0.005 *and* 0.001 *in the middle column, and* 0.01, 0.01 *and* 0.005 *in the right column.*

Depending on the field, all frequencies can be dominantly balanced (as in the case of $\tilde{h}$). For the divergence, low frequencies are dominantly balanced, while high frequencies are dominantly imbalanced for cases 2 and especially 3. This is why gravity waves are visible in δ but not in $\tilde{h}$ (cf. figures 7 and 6).

4   Conclusions

This work began with the development of a spherical shallow-water model employing the more standard variables (h, δ, Π), which were known in the f-plane context to perform poorly as regards the gravity-wave component of the flow. Guided by ideas of balance, Mohebalhojeh & Dritschel (2001,2002) demonstrated that better accuracy for *both* the gravity-wave and the vortical components could be achieved by using a transformed set of prognostic variables tailored to the expected underlying balance. For efficient computation and robust performance, the set (δ, γ, Π) proved most useful. The present work has used these same variables but now in spherical geometry. Comparisons with the original code using the set (h, δ, Π) have also shown the new set strongly reduces the erroneous numerical generation of gravity waves. It is noteworthy that the two variable sets are closely comparable in numerical efficiency.

A new procedure for diagnosing gravity waves has demonstrated that these waves remain weak (at least over moderate time) if they start weak, even for flows whose Froude number is close to unity. Future work will explore how the gravity-wave component of the flow varies with Rossby and Froude number.
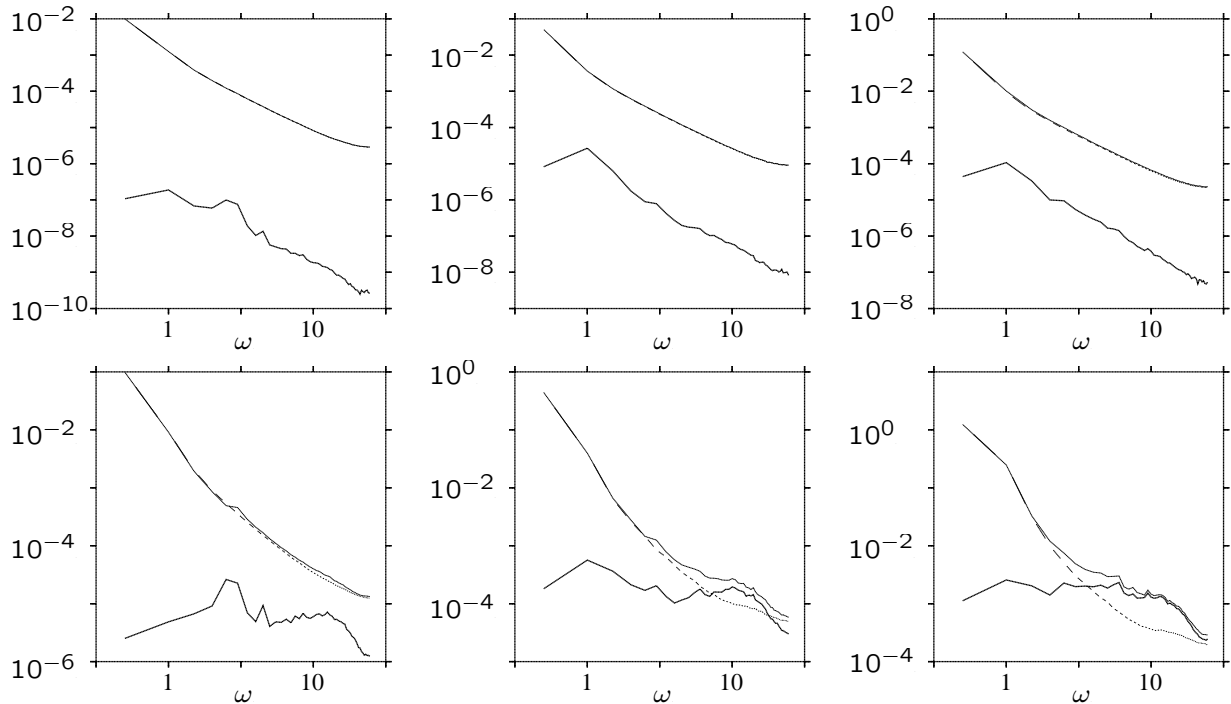
Acknowledgements

Figure 8: Depth (top) and divergence (bottom) frequency spectra versus frequency ω for for case 1 (left), case 2 (middle) and case 3 (right).

References

[1] DRITSCHEL, D. G. 1988 Contour dynamics/surgery on the sphere. *J. Comput. Phys.* **78**, 477–483.

[2] DRITSCHEL, D. G., & AMBAUM, M. H. P. 1997 A contour-advective semi-Lagrangian algorithm for the simulation of fine-scale conservative fields. *Quart. J. Roy. Meteor. Soc.* **123**, 1097–1130.

[3] DRITSCHEL, D. G., POLVANI, L. M. & MOHEBALHOJEH, A. R. 1999 The contour-advective semi-Lagrangian algorithm for the shallow-water equations. *Mon. Wea. Rev.* **127(7)**, 1151–1165.

[4] DRITSCHEL, D. G., & MOHEBALHOJEH, A. R. 2000 The contour-advective semi-Lagrangian algorithm: keeping the balance. ECMWF Workshop Proceedings on *Developments in numerical methods for very high resolution global models*, also available at http://www-vortex.mcs.st-andrews.ac.uk/∼arm/ecmwf00/index.html.

[5] DRITSCHEL, D. G., & VIÚDEZ, A. 2003 A balanced approach to modelling rotating stably-stratified geophysical flows. *J. Fluid Mech.* **488**, 123–150.

[6] HOSKINS, B. J., MCINTYRE, M. E., & ROBERTSON, A. W. 1985 On the use and significance of isentropic potential-vorticity maps. *Quart. J. Roy. Meteorol. Soc.* **111**, 877–946.

[7] MCINTYRE, M. E., & NORTON, W. A. 1999 Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.* **57**, 1214–1235, Corrigendum **58**, 949.

[8] MOHEBALHOJEH, A. R., & DRITSCHEL, D. G. 2000 On the representation of gravity waves in numerical models of the shallow water equations. *Quart. J. Roy. Meteor. Soc.* **126**, 669–688.

[9] MOHEBALHOJEH, A. R., & DRITSCHEL, D. G. 2001 Hierarchies of balance conditions for the *f*-plane shallow water equations. *J. Atmos. Sci.* **58(16)**, 2411–2426.

[10] MOHEBALHOJEH, A. R. 2002 On shallow-water potential-vorticity inversion by Rossby-number expansions. *Quart. J. Roy. Meteor. Soc.* **128**, 679–694.

[11] MOHEBALHOJEH, A. R., & DRITSCHEL, D. G. 2004 Contour-advective semi-Lagrangian algorithms for many-layer primitive equation models. *Quart. J. Roy. Meteor. Soc.* **130**, 347–364.

[12] NORBURY, J. & ROULSTONE, I. (EDS.) 2002a Large-scale Atmosphere–Ocean Dynamics: Vol. I: Analytical Methods and Numerical Models. Cambridge University Press.

[13] NORBURY, J. & ROULSTONE, I. (EDS.) 2002a Large-scale Atmosphere–Ocean Dynamics: Vol. II: Geometric Methods and Models. Cambridge University Press.

[14] VIÚDEZ, A., AND DRITSCHEL, D. G. 2003 Vertical velocity in mesoscale geophysical flows. *J. Fluid Mech.,* **483**, 199–223.

[15] VIÚDEZ, A., AND DRITSCHEL, D. G. 2004 Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.,* (under review).