Q. J. R. Meteorol. Soc. (1996), **122**, pp. 291–322

Potential vorticity inversion and balanced equations of motion for rotating and stratified flows

By GEOFFREY K. VALLIS*
University of California, Santa Cruz, USA

(Received 25 August 1994; revised 9 May 1995)

SUMMARY

Balanced equations of motion based on potential vorticity evolution and inversion for the shallow water and stratified primitive equations are derived and, in some shallow-water cases, numerically tested. The schemes are based on asymptotic expansions in Rossby or Froude number, or rational scaling-based truncations of the equations of motion, assuming that the dynamics are determined by the advection of potential vorticity. Thus, regimes of validity are rapidly rotating and/or highly stratified flow. Both new and familiar results are straightforwardly obtained, in a unified framework in both height and isentropic coordinates. For both shallow-water and stratified equations, Rossby number expansions schemes give quasi-geostrophy at lowest order. Both gradient-wind balance and the nonlinear terms in the potential vorticity enter at next order. A low Froude number expansion for non-rotating flow gives two-dimensional flow, uncoupled in the vertical at lowest order. A single consistent inversion scheme can be derived that is valid at lowest order in Froude number for all Rossby numbers, for both shallow-water and the stratified equations. It may be a particularly appropriate model for the atmospheric mesoscale and oceanic submesoscale, where rotation and stratification can both be important in defining balanced motion. A model is also proposed that is valid at both planetary and synoptic scales, combining the familiar planetary geostrophic and quasi-geostrophic equations. Most of the models derived require the solution only of linear or near linear elliptic equations, possibly with varying coefficients.

Numerical experiments indicate that a higher-order inversion can be quantitatively better than quasi-geostrophy, if Rossby number and divergence are sufficiently small. In some other cases, no noticeable improvement over quasi-geostrophy is found, even when the Rossby number is quite small. However, the balanced model valid for both planetary and synoptic scales shows a significant qualitative and quantitative improvement over both planetary geostrophy and quasi-geostrophy for large-scale flows, and its evolution is in good agreement with a primitive equation model.

KEYWORDS: Balanced flow Potential vorticity inversion Shallow-water equations

1. INTRODUCTION

The large-scale and mesoscale motion of both ocean and atmosphere are observed to be, to some degree, 'balanced'. That is to say, the motion is largely devoid of gravity waves, and the gravity waves which do exist are not primarily responsible for determining the evolution of the large-scale features of the flow. The flow field is called 'balanced' because the velocity and pressure fields approximately satisfy some diagnostic relationship, of which geostrophic balance is perhaps the simplest and most well known to meteorologists and oceanographers.

The primitive equations—that is, the Navier–Stokes equations plus hydrostatic balance—are usually considered the appropriate set of equations that describe the large-scale and mesoscale motion, although for some situations the requirement of hydrostatic balance is relaxed. The primitive equations are closely related to the shallow-water equations, especially if the former are written in isentropic co-ordinates for then a modal expansion in the vertical leads to equations for each mode with similar structure to the shallow-water equations. Thus, the latter are a very useful model problem; their basic equations are:

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + f\mathbf{k} \times \mathbf{u} = -g\nabla h \tag{1}$$

$$\frac{\mathrm{D}h}{\mathrm{D}t} + h\nabla \cdot \mathbf{u} = 0. \tag{2}$$

* Corresponding address: Department of Marine Science, University of California, Santa Cruz, CA 95064, USA.

TABLE 1. NOTATION AND PARTIAL DESCRIPTION OF SYMBOLS

Symbol	Meaning
ψ	Streamfunction.
Q	Potential Vorticity.
q	Potential Vorticity, minus contribution from planetary rotation. See text.
$\mathbf{u}$	Two dimensional velocity, $u\mathbf{i} + v\mathbf{j}$.
u, v, w	Components of velocity.
$\psi; \zeta$	Streamfunction; vorticity.
$\chi; \delta$	Velocity potential; divergence.
$h; \tilde{h}$	Height field; height divided by reference height (h/H).
H	Reference (normally mean) height field (shallow water), or height scale (stratified equations).
θ	Potential temperature.
$\hat{\theta}$	Reference potential temperature.
p	Pressure.
ϱ	$\partial p/\partial\theta$.
M	Montgomery streamfunction, $c_pT + gz$.
Π	Exner function, $c_p(p/p_s)^{R/c_p}$.
$R,\ c_p,\ c_v$	Gas constant and heat capacities at constant pressure and volume.
f	Variable Coriolis parameter.
f_0	Reference value of f.
β	$\partial f/\partial y$.
U, L, H	Representative (scaling) magnitudes of velocity, length and height.
N	Brunt-Väisällä frequency, $(g/\hat{\theta})\partial\hat{\theta}/\partial z \sim g\Delta\theta/(\theta H)$
F	Froude number, $U/\sqrt{gH}$ or U/NH. See text.
F_R	Square rotational Froude number. f^2L^2/gH.
D	Deformation radius. $\sqrt{gH}/f$.
ε	Rossby number. U/fL.
∇^2	Two-dimensional Laplace operator.
$J(a,\ b)$	Jacobian. $J(a,\ b) = (\partial a/\partial x)(\partial b/\partial y) - (\partial a/\partial y)(\partial b/\partial x)$.

(See Table 1 for notation). Geostrophic balance is *two* constraints, namely:

$$fu = -g\frac{\partial h}{\partial y}, \qquad fv = g\frac{\partial h}{\partial x}. \tag{3}$$

It is a characteristic of this kind of balanced system that there is one evolving variable, coupled with two constraints. An equivalent way of expressing this is that the linear normal modes of the system contain two gravity modes and one 'Rossby' mode. Some balanced systems nevertheless apparently contain two evolution equations. In some cases, for example the 'Balance Equations' or the 'Charney Balance' system (Gent and McWilliams 1983), this is only superficially true, since one evolving variable is sometimes tied to the other via the constraint. In any case, because there is only one variable which evolves on a slow timescale, a fully balanced system should be expressible in terms of one evolving variable, plus constraints or diagnostic relationships between the variables, and it therefore seems appropriate to explicitly exploit this.

Is there a preferred evolution variable? Examining (2) and (3) it appears that one could simply evolve the height (or potential temperature or density in the stratified equations) and diagnoze the velocity from geostrophy, say. This is what is done in the planetary geostrophic or Phillips type II equations, (Phillips 1963), valid only for very large-scale flow. But height field is not an advected variable, and the possibility exists that some terms will be large and some negligible; the scaling of the time-derivative is then not obvious and the model dynamics are thereby complicated. On the other hand potential vorticity is evolved purely through advective effects; it does not partake of (linear) gravity-wave oscillations, and is unambiguously a 'slow' variable (Warn *et al.* 1994). In three-dimensional stratified flow, both potential vorticity and potential temperature are slow, advected, variables. Here, it is appropriate for a balanced system to comprise potential vorticity advection in the interior, with boundary conditions provided by potential temperature (or density) evolution. For adiabatic flow in isentropic coordinates 'vertical' advection is eliminated and all the advection terms scale in a similar way, making potential vorticity advection in isentropic coordinates the simplest and most natural evolution equation. Potential temperature is then used as a 'coordinate' rather than an independent variable in the interior. For these, and most likely other phenomenological reasons, potential vorticity has come to play a central role in oceanic and atmospheric dynamics—see for example Charney (1948), Charney and Stern (1962), Green (1970), Rhines and Young (1982), Hoskins *et al.* (1985), Raymond (1992), and many others.

Given that the dynamics are to be determined by one evolution equation, plus constraints, the question arises as to how to actually construct equations of motion for balanced dynamics. The problem becomes one of 'inversion' of the chosen variable; that is, constructing a representation of the other variables—velocity, temperature and so on—in terms of the evolving variable (c.f. Norton 1988). In the presence of a small parameter, such as a Rossby or Froude number, one way to do this is to systematically expand all variables, except the evolution variable, in a series in that small parameter (Warn *et al.* 1994). This differs somewhat from the most straightforward asymptotic procedure, which is to assume that all variables depend on the small parameter, and thus expand everything. Rather, it is equivalent to expanding the *relationship* between the evolving, independent, variable and the dependent fields. This approach may be regarded just as a means of avoiding secularities in higher-order expansions of the equations of motion. Alternatively, the reason that one field cannot be expanded is that the field to be inverted is essentially ' given.' All of it must then be considered to be lowest order. It makes little sense to expand a given field, just as it makes little sense to expand an independent variable.

The advantage of such a procedure, aside from its rational base and 'algorithmic' nature, is that a 'proper' balanced system of one evolution operator is assured. It gives a well founded scheme for inverting a potential-vorticity field to an accuracy limited, in principle, by the dynamics themselves and not by the limitations of the inversion scheme, provided the assumed scaling is satisfied and a small parameter present. As a way of constructing models for long term numerical integration it may be less appropriate: a possible disadvantage is that exact integral conservation properties of energy and mass are not assured. Thus, the procedure may be more useful as a diagnostic tool for analysis of primitive equation models and observations than as a procedure for model construction. However, the importance of maintaining conservation properties (at least for initial value problems) is not obvious (see for example Allen and Newberger 1993; Snyder *et al.* 1991). Strict asymptotic procedures can be rather restrictive too, especially when considering transitions between different flow regimes or scales where balance is maintained in different ways (e.g. low Froude number and/or low Rossby number; planetary/synoptic-scale flow), a common circumstance in both ocean and atmosphere, and one wishes to have a single valid model for both regimes. A less formal approach may be called for, perhaps mixing formal asymptotic orders or retaining lowest order terms from two expansion procedures, but potential-vorticity advection and evolution can still be central.

The semi-geostrophic (or geostrophic momentum) system (Hoskins 1975) is, when written in physical space, of rather different form than potential-vorticity advection/inversion. However, it does preserve a form of potential vorticity, and in transformed space has a natural physical interpretation. Xu (1994) discusses connections between semi-geostrophy and other balanced models, and presents a 'semi-balance' model with an invertible potential vorticity. See also Craig (1993).

In the rest of this paper we directly attack the problem of constructing potential-vorticity inversion schemes for both rapidly rotating and stratified flow—the two conditions which more clearly than any define the structure of the atmosphere and ocean. The Rossby number and Froude number, respectively, are then the natural small parameters. We aim to make rational approximations throughout, whether using asymptotic or other techniques, and to keep models as austere as possible. Section 2 is concerned with the shallow-water equations and section 3 with the stratified primitive equations, in both physical space and isentropic coordinates. For both classes of flow, balanced models valid at low Rossby number and/or low Froude number are derived. In particular, schemes accurate to first order in Rossby number (i.e. one order beyond quasi-geostrophy which will be referred to as zeroth order) are explicitly derived, as are schemes valid for small Froude number at all Rossby number. Balanced models valid for both synoptic and planetary-scale flow are also presented. Throughout, we discuss where models are rigorously valid, where they appear phenomenologically valid, and where they are not applicable.

2. The shallow-water equations

(*a*) *Rapidly rotating flow*

A Rossby number expansion for the shallow-water equations is the simplest example of the asymptotic method. The derivation, following Warn *al.* (1995), extended slightly to include the β-plane, is given here by way of an introduction. The equations of motion (1) and (2) are re-written in a linear normal-mode form, with evolution equations for potential vorticity, height and 'imbalance'. The potential vorticity equation is:

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

where

$$Q = \frac{f + \zeta}{H + h} \tag{5}$$

where ζ is vorticity, H and h are mean and perturbation fluid depths respectively. The divergence equation is:

$$\frac{\partial \delta}{\partial t} + J(\chi,\ \nabla^2\psi) + \frac{1}{2}\nabla^2[(\nabla\chi)^2] - \nabla^2 J(\chi,\ \psi) - 2J(\psi_x,\ \psi_y) + \beta u = f\nabla^2\psi - g\nabla^2 h \tag{6}$$

where

$$\delta = \nabla \cdot \mathbf{u} = \nabla^2\chi. \tag{7}$$

The imbalance equation is:

$$\begin{aligned}\frac{\partial \Omega}{\partial t} + J(\psi,\ \nabla^2\psi) + \nabla\cdot(\nabla^2\psi\nabla\chi) - (g/f)\nabla^2[J(\psi,\ h) + \nabla\cdot(h\nabla\chi)]&\\ +\beta v = -f(1 - D^2\nabla^2)\delta&\end{aligned} \tag{8}$$

where

$$\Omega = \nabla^2\psi - (g/f)\nabla^2 h. \tag{9}$$

Under small Rossby number scaling (i.e. a scaling from geostrophic balance), the following holds:

$$\frac{h}{H} \sim \frac{fUL}{gH} = \varepsilon F_{\mathrm{R}} = \varepsilon \frac{L^2}{D^2} = \frac{F^2}{\varepsilon} \tag{10}$$

$$\zeta \sim \frac{U}{L} \sim \varepsilon f \tag{11}$$

where ε is the Rossby number U/fL; F_{R}, a square rotational Froude number, is f^2L^2/gH; $D = \sqrt{gH}/f$ is a deformation radius; and F is the Froude number $U/\sqrt{gH}$. Also, for all but planetary-scale flow the divergence is small because, for $L^2 \leqslant D^2$, the height or the imbalance equation yields the estimate

$$\delta = \nabla \cdot \mathbf{u} \sim \frac{U}{L}\varepsilon F_{\mathrm{R}} = \frac{U}{L}\frac{F^2}{\varepsilon} = \frac{U}{L}\varepsilon\frac{L^2}{D^2} \leqslant \frac{U}{L}. \tag{12}$$

Expanding the variables on the right hand side of (5) in a series in Rossby number yields:

$$\begin{aligned} HQ - f_0 = \{\varepsilon\}q = \{\varepsilon\}[\zeta_0 + \beta y - \{F_{\mathrm{R}}\}f_0\tilde{h}_0] + \{\varepsilon^2\}[\zeta_1 - \{F_{\mathrm{R}}\}f_0\tilde{h}_1 + \\ +(\{F_{\mathrm{R}}\}f_0\tilde{h}_0 - \zeta_0 - \beta y)\tilde{h}_0] + O(\varepsilon^3) \end{aligned} \tag{13}$$

where the Coriolis parameter is represented as $f_0 + \{\varepsilon\}\beta y$. Here, and in the rest of the paper, expressions are written in dimensional form. The relevant non-dimensional parameters are also given, enclosed in curly brackets, { }. The pure dimensional form may be recovered simply by setting to unity the contents of the curly brackets. The non-dimensional form is recovered by setting all the physical parameters (such as g and f) to unity. This notation enables both the asymptotic ordering and physical appreciation of the terms to be facilitated. The terms of a particular order will normally be contained in square brackets, [ ].

At lowest order it is clear that quasi-geostrophy arises. The potential vorticity is given by $q = [\zeta_0 + \beta y - \{F_{\mathrm{R}}\}f_0\tilde{h}_0]$. The divergence equation yields $f_0\nabla^2\psi_0 = gH\nabla^2\tilde{h}_0$, and the imbalance equation yields, at lowest order, zero divergence. These results are equivalent to quasi-geostrophy, for explicitly they give the inversion for the streamfunction,

$$\nabla^2\psi_0 - \{F_{\mathrm{R}}\}\frac{f_0^2}{gH}\psi_0 = q - \beta y, \tag{14}$$

and potential vorticity is advected by the inverted velocity. We assume, here and throughout, that an appropriate non-dimensionalization (i.e. a scaling) for time is L/U; that is, an advective, rather than a gravity wave, time-scale.

We stated above that the potential vorticity q should be regarded purely as a zeroth order quantity. Then the constraint arises that the coefficient of ε^2 in the expansion (13) must vanish. (Equivalently, note that the lowest order vorticity and height fields satisfy $q = \zeta_0 + \beta y - f\tilde{h}_0$. Therefore the coefficients of ε, ε^2 and so on must vanish.) Thus:

$$\zeta_1 - \{F_{\mathrm{R}}\}f_0\tilde{h}_1 = -(\{F_{\mathrm{R}}\}f\tilde{h}_0 - \zeta_0 - \beta y)\tilde{h}_0. \tag{15}$$

The divergence equation gives:

$$f_0\nabla^2\psi_1 - gH\nabla^2\tilde{h}_1 = \beta u_0 - 2J(\psi_{0x},\ \psi_{0y}) \tag{16}$$

and the imbalance equation gives the shallow-water 'omega' equation:

$$\mathcal{H}\delta = \frac{1}{D^2 f_0}(J(\psi_0,\ \nabla^2\psi_0) + \beta v_0), \tag{17}$$

where $\mathcal{H} = \nabla^2 - 1/D^2$.

Equations (15) to (17) form a closed system, with the lowest-order (quasi-geostrophic) streamfunction calculated from $(\nabla^2 - 1/D^2)\psi_0 = q - \beta y$. Explicit expressions for the first-order streamfunction and height are:

$$\mathcal{H}(\nabla^2\psi_1) = \frac{1}{D^2 f_0}(\nabla^2(\psi_0 q) + 2J(\psi_{0x},\ \psi_{0y}) - \beta u_0), \tag{18}$$

and

$$h_1 = \frac{H}{f_0}\nabla^2\psi_1 - \frac{1}{g}\psi_0 q. \tag{19}$$

The stream function and height are then reconstituted from:

$$\psi = \psi_0 + \{\varepsilon\}\psi_1 \tag{20}$$

and

$$h = \frac{f_0}{g}\psi_0 + \{\varepsilon\}h_1. \tag{21}$$

Note that gradient-wind balance is incorporated at this order. It arises in conjunction with the nonlinear terms in the definition of potential vorticity, although the inversion requires solution only of linear elliptic equations. Note that all velocities obtained via an inversion are 'balanced', in the sense that they can be diagnosed from other fields. They are not unbalanced, for this implies independent gravity-wave motion, although they may be ageostrophic, if they satisfy a different or higher-order balance.

(*b*) *Small Froude-number flow*

Balanced flow in the shallow-water equations also arises for small Froude number, even for non-rotating flow. Taking the Rossby number to be O(1) or larger, and assuming a balance between the nonlinear advective terms and the pressure, the scaling becomes:

$$\frac{h}{H} \sim \frac{U^2}{gH} = F^2. \tag{22}$$

Using this to non-dimensionalize the height field, the height equation (2) can be written

$$\frac{\partial h}{\partial t} + J(\psi,\ h) + \nabla\cdot(h\nabla\chi) + \left\{\frac{1}{F^2}\right\} H\delta = 0. \tag{23}$$

This reveals that the divergence is zero at lowest order, and gives the estimate

$$\delta = \nabla\cdot\mathbf{u} \sim \frac{U}{K}F^2. \tag{24}$$

The potential-vorticity expansion in Froude number is:

$$HQ = (f + \{\varepsilon\}\zeta_0) + \{F^2\}[\{\varepsilon\}\zeta_1 - \tilde{h}_0(f + \{\varepsilon\}\zeta_0)] + O(F^4). \tag{25}$$

The lowest order approximation is then given by neglecting terms of $O(F^2)$ and higher, giving:

$$\nabla^2\psi_0 = QH - f. \tag{26}$$

The appropriate evolution equation is then,

$$\frac{\partial Q}{\partial t} + J(\psi_0,\ Q) = 0 \tag{27}$$

where streamfunction is obtained from (26). The divergent part of the velocity field does not advect the potential vorticity at this order. Equation (27) is just the equation of motion for two-dimensional flow. The model, however, differs from the perhaps more familiar barotropic or two-dimensional flow that characterizes, say, the large-scale atmosphere in that the stream function is no longer the pressure. The height field is obtained from the divergence equation which now is

$$\frac{\partial \delta}{\partial t} + J(\chi,\ \nabla^2\psi) + \frac{1}{2}\nabla^2[(\nabla\chi)^2] - \nabla^2 J(\chi,\ \psi) - 2J(\psi_x,\ \psi_y) + \left\{\frac{1}{\varepsilon}\right\}\beta u$$
$$= \left\{\frac{1}{\varepsilon}\right\} f\nabla^2\psi - g\nabla^2 h. \tag{28}$$

Since the velocity potential at lowest order is zero, the pressure (i.e. height) field is given to lowest order by a solution of the gradient-wind balance equation:

$$g\nabla^2 h_0 = \left\{\frac{1}{\varepsilon}\right\} \nabla\cdot f\nabla\psi_0 + 2J(\psi_{0x},\ \psi_{0y}) \tag{29}$$

where, because divergence is small and neither f nor β are large, we can include the βu term in the first term on the right hand side. This is a linear elliptic equation, as the stream function is already determined from the potential vorticity using (26). Note too that the height field is not required to evolve the model.

This model is not valid for small Rossby number (more precisely if $F^2/\varepsilon \sim O(1)$), because in the potential-vorticity expansion the neglected term $\tilde{h}_0 f$ may become as important as the retained term ζ_0. However, within the confines of this constraint the Coriolis parameter f is in fact allowed to vary fully: the horizontal divergence is small because of the small Froude number, and does not require f to be constant. The two constraints of this model (corresponding to the two constraints of geostrophic balance in rapidly rotating flow) are that divergence is zero (to lowest order) and gradient-wind balance (29).

(c) *Small Froude-number, rapidly rotating flow*

It is possible to proceed to higher order in Froude number to obtain inversions of formally increasingly high accuracy. However, it is probably more useful to note that a single model can be obtained that is accurate to lowest order in Froude number and valid for all Rossby numbers. This is done by retaining the term in $h_0 f$ in the expansion of potential vorticity. This is formally a higher order term in a Froude number expansion, but one that appears at lowest order in a Rossby number expansion. Potential vorticity is given by:

$$HQ = f_0 + \{\varepsilon\}\zeta_0 - \left\{\frac{F^2}{\varepsilon_{\min}}\right\}\tilde{h}_0 f_0 \tag{30}$$

where $\varepsilon_{\min} = \min(1,\ \varepsilon)$. This retains all the terms in the lowest order expansion of both the Rossby number expansion (13) and Froude-number expansion (25). The height field and streamfunction are then related by the gradient-wind balance, which similarly retains all terms of both models:

$$\{\varepsilon_{\max}\} g\nabla^2 h_0 = \nabla \cdot f_0 \nabla \psi_0 + \{\varepsilon\} 2J(\psi_{0x},\ \psi_{0y}) \tag{31}$$

where $\varepsilon_{\max} = \max(1,\ \varepsilon)$. The evolution equation is simply:

$$\frac{\partial Q}{\partial t} + \mathbf{u}_0 \cdot \nabla Q = 0 \tag{32}$$

where $\mathbf{u}_0 = \mathbf{k} \times \nabla\psi_0$, and ψ_0 is obtained using (30) and (31). The divergent field is not needed to advect the flow, as it is zero at lowest order. This model has one additional restriction over the low Froude-number model: the Coriolis parameter is not allowed to vary fully, although as in quasi-geostrophy a weak beta-effect is allowed. This is because if f varies by an O(1) amount and height variations are not small, at small Rossby number the leading term in the potential-vorticity expansion is f/h. The evolution equation at lowest order is then $\mathrm{D}(f/h)/\mathrm{D}t = 0$, or planetary geostrophy (see section 2(d)). This case is not included in the above model because h/H is assumed small and $1/(1 + h/H)$ is expanded in a series. Also, for planetary scales ($L^2 \gg D^2$) the divergence is given by the 'Sverdrup' relationship:

$$f\nabla \cdot \mathbf{u} \sim \beta v \tag{33}$$

and is only small if β is small. If h/H does remain small, then at small Rossby number the leading term in the potential-vorticity expansion is f/H and the evolution equation yields $\beta v = 0$ at lowest order. Although this does illustrate the constraining effects that the β-effect has on meridional motion, it is not an especially useful evolution equation.

Even though the model incorporates gradient-wind balance, it is not first order accurate in Rossby number, even for small Rossby number. This is because it neglects the nonlinear terms in the potential vorticity expansion, which arise at the same order as gradient-wind balance. The terms could be included in (30) but the inversion would be correspondingly more complex. Also, for any given class of motion (namely low Froude number or low Rossby number) the model will formally be only as accurate as the appropriate one of the two simpler models of the previous two subsections. However, in regimes under which both rotation and stratification are important, or in which there is a transition from one regime to the other, it may be a more useful model of the shallow-water equations than either the pure Rossby-number or pure Froude-number expansions, carried to any order. Similarly, the gradient-wind correction may well be the most important correction to quasi-geostrophy for most geophysical flows, because it arises at zeroth order in a low Froude-number expansion. An apparent disadvantage is that in solving (30) and (31) for streamfunction a non-linear Monge–Ampère equation arises. However, its solution is not in fact not difficult here, because the contribution of the Jacobian term is always small if the scaling is satisfied. From (30) and (31) we obtain, for any Rossby number,

$$q = \zeta_0 - \left\{\frac{F^2}{\varepsilon^2}\right\} \frac{f_0^2}{gH}[\psi_0 + \{\varepsilon\} 2\nabla^{-2} J(\psi_{0x},\ \psi_{0y})]. \tag{34}$$

Thus (formally) the Jacobian term may always be omitted from inversion for the stream function, since $F^2/\varepsilon \ll 1$: either $F^2/\varepsilon^2 \leqslant O(1)$ and $\varepsilon \ll 1$ (the quasi-geostrophic case) or $F^2 \ll 1$ and $\varepsilon \sim 1$ (the highly stratified case, finite Rossby number case). If desired its effects could easily be included in an iterative fashion, and the term is needed at lowest order to solve for the pressure. See section 4(c) for a discussion of a related point.

(*d*) *Planetary-scale flow*

The characteristic of planetary-scale motion is that the Coriolis parameter should be allowed to vary by an O(1) amount, and the variations in the height field should likewise not be restricted. However, except for truly global flow it may still be demanded that Rossby number remain small, and it thus forms a natural expansion parameter. Thus,

$$\zeta(x,\ y,\ t) = \zeta_0 + \{\varepsilon\}\zeta_1 + \{\varepsilon^2\}\zeta_2 + O(\varepsilon^3) \tag{35}$$

and

$$h(x,\ y,\ t) = h_0 + \{\varepsilon\}h_1 + \{\varepsilon^2\}h_2 + O(\varepsilon^3). \tag{36}$$

In the latter expression, h_0 may be thought of as consisting of two parts, $H + h'_0$. The parameter H determines the mean thickness of the fluid, and is thus given by the total amount of fluid, whereas the variable h'_0 varies dynamically. However, no restriction that $|h'_0| \ll H$ is implied, although h_0 must remain positive.

Expanding the constituents of potential vorticity in powers of ε gives:

$$Q = \frac{f}{h_0} + \{\varepsilon\}\left[\frac{\zeta_0}{h_0} - \frac{fh_1}{h_0^2}\right] + O(\varepsilon^2). \tag{37}$$

The lowest-order field is given by:

$$h_0 = \frac{f}{Q}. \tag{38}$$

The appropriate velocity fields are given from the lowest-order balances in the imbalance and divergence equations of section 2(a), using $\beta L/f = O(1)$ and $L^2 \gg D^2$. These respectively yield

$$\beta v_0 = -f\nabla \cdot \mathbf{u}_0 \tag{39}$$

and

$$f\zeta_0 = \beta u_0 + g\nabla^2 h_0. \tag{40}$$

These are exactly equivalent to geostrophic balance:

$$u_0 = -\frac{g}{f}\frac{\partial h_0}{\partial y}, \qquad v_0 = \frac{g}{f}\frac{\partial h_0}{\partial x}. \tag{41}$$

The evolutionary model then comprises the advection of potential vorticity by velocities determined by the inversion (38) and (41). That is, we have recovered the planetary geostrophic or Phillips type II equations (Phillips 1963) at lowest order. Proceeding to next order in the expansion, from (37) the first-order field is determined from,

$$\frac{\zeta_0}{h_0} = \frac{fh_1}{h_0^2} \tag{42}$$

or

$$h_1 = \frac{\zeta_0 h_0}{f} = \frac{\zeta_0}{Q}. \tag{43}$$

Hence, to first order the height field is given by

$$h = h_0 + \{\varepsilon\}h_1 = \frac{f + \varepsilon\zeta_0}{Q}. \tag{44}$$

In this expression, ζ_0 is the lowest order representation of the vorticity, obtained from geostrophic balance. Thus, (44) becomes:

$$h = \frac{1}{Q}\left[f + \frac{g}{f}\nabla^2 h_0 + \frac{\beta u_0}{f}\right] \tag{45}$$

where $u_0 = -(g/f)\partial h_0/\partial y$ and $h_0 = f/Q$. Plainly, the inversion fails if $Q = 0$; but this is not worrying, since the scheme is limited to small Rossby number and we have no right to consider zero Q cases. It is instructive to express the potential vorticity in terms of the height field. This gives

$$Q = \frac{f + (g/f)\nabla^2 h_0 + (\beta u_0/f)}{h\ (= h_0 + h_1)}. \tag{46}$$

This model therefore includes in some sense the advection of relative vorticity, although we still have $Q = f/h_0$.

To obtain a model for time integration a velocity is needed to advect the potential vorticity. A completely consistent asymptotic approach to this is involved and uninformative. A more serious difficulty with the model lies in the differential order of h_1. We have $h_1 \sim \nabla^2 h_0 \sim \nabla^2(f/Q)$. Thus, h_1 will amplify details in Q, rather than relaxing them through an integral operator, and thus the inversion is completely inappropriate if there is small-scale detail in Q. As the presence of such detail violates the scaling on which the inversion is based, formally it is not a problem. However, numerical tests indicate that the inversion is in fact not very robust, and the model of the next section is superior.

(*e*) *A model for planetary and synoptic-scale flow*

We now derive a model valid for both planetary and synoptic-scale flow. It is hard to do this by formal asymptotic methods, since these normally lead to either the planetary geostrophic or quasi-geostrophic equations, neither of which are valid for both synoptic-scale and planetary-scale flow. Rather, we obtain an inversion by neglect of certain small terms in the divergence and imbalance equations. If the model is evolved by advecting potential vorticity, *and* the chosen scaling implies that the time derivatives in the other equations are properly small and thus can be neglected, then the model is rational and balanced. This is the approach of 'truncation'. At lowest order the method can, and will often, give the same results as an asymptotic technique. It is harder to go to arbitrarily high order, because some terms (e.g. certain time derivatives) are simply neglected, whereas in an asymptotic method they may be folded back in at higher order. However, some higher-order terms, or more importantly terms which appear at lowest order in more than one asymptotic expansion, may be included immediately, thereby extending the range of validity of the model. Thus, although less formal than asymptotics, the method is more flexible for deriving equations of motion valid in more than one flow regime. The model of section 2(c) could be considered an example. It is also a particularly appropriate approach for planetary/synoptic-scale flow, where we are able to derive a single model appropriate for both scales of motion.

For the shallow-water equations, we begin with the definition of potential vorticity

$$Q = \frac{\zeta + f}{h}. \tag{47}$$

The lowest-order approximant at small Rossby number is just f/h. A better approximation is obtained by using the geostrophically balanced vorticity: the lowest-order small Rossby-number balance in the divergence equation (6) is

$$\beta u = f\nabla^2\psi - g\nabla^2 h, \tag{48}$$

and using this in (47) yields,

$$Q = \frac{f + (g/f)\nabla^2 h - \beta g(\partial h/\partial y)f^2}{h} \tag{49}$$

where we have also substituted $u = -(g/f)(\partial h/\partial y)$. Hence we obtain the following single equation for the height field,

$$\frac{g}{f}\left(\nabla^2 h - \frac{\beta}{f}\frac{\partial h}{\partial y}\right) - Qh = -f. \tag{50}$$

The equation is linear, but has non-constant coefficients. Since it is truly higher order than the planetary geostrophic equations, it allows boundary conditions to be satisfied by inertial rather than by the addition of frictional effects. The velocity field is then obtained from the height field using geostrophic balance,

$$u = -\frac{g}{f}\frac{\partial h}{\partial y}, \qquad v = \frac{g}{f}\frac{\partial h}{\partial x} \tag{51}$$

valid for both synoptic and planetary-scales. The validity of the scaling can be made transparent using the divergence and imbalance equations. The dominant terms in these are:

$$g\nabla^2 h + \left\{\frac{L}{L_p}\right\}\beta u = f\nabla^2\psi + O(\varepsilon), \tag{52}$$

and

$$\left\{\frac{L}{L_p}\right\}\beta v = -f\left(1 - \left\{\frac{D^2}{L^2}\right\}D^2\nabla^2\right)\delta + O(\varepsilon), \tag{53}$$

where L is the length scale of the motion and L_p the scale over which f varies by $O(1)$. The quasi-geostrophic approximation assumes $L/L_p \ll 1$, and neglects the β term and sets $f = f_0$ in both equations, so that (53) gives $\delta = 0$. Planetary geostrophic scaling assumes $L_p \sim L \gg D$ and neglects the second term in brackets on the right-hand-side of (53). Both cases are allowed for in (51), and in both cases the time derivatives are small in (6) and in (8), so the motion is indeed properly balanced.

The model is defined by the advection of potential vorticity, with an inversion consisting of (50) and (51). It may be useful for modelling planètary and synoptic-flows, since it is valid at both scales provided either quasi-geostrophy or planetary geostrophy is valid. For very large-scale motions the dominant terms are the same as those of the planetary geostrophic equations, whereas for synoptic-scale flow, with small variations in height field and Coriolis parameter, the equations asymptotically reduce to the quasi-geostrophic set. In either limit the motion is properly balanced, and the model neither neglects terms which are potentially important nor retains terms which are never important. The equations conserve potential vorticity, by construction, and global mass—this by choice of boundary condition in the solution of (50). The model allows large variations in the height field only if they occur on a large scale. It is not a demonstrably appropriate balanced model if rapid large-scale changes in topography (dynamic or bottom) were to occur, such as in certain coastal oceanographic regimes.

An improvement in the model's description of synoptic scales can be obtained by using a higher order relationship between vorticity and height. Instead of pure geostrophic balance we can incorporate gradient-wind balance:

$$g\nabla^2 h + \beta u - f\nabla^2\psi = 2J(\psi_x,\ \psi_y) \tag{54}$$

where $u = -\psi_y + \chi_x$ in the beta term. For small Rossby-number flow we may simplify this to $u = -(g/f)\partial h/\partial y$; this is valid because at synoptic scales the contribution of β term is small and the approximation is unimportant, and at large scales the approximation is valid by geostrophy. The divergent flow field can in turn be more accurately obtained using the shallow-water omega equation,

$$\mathcal{H}\delta = \frac{1}{D^2 f}(J(\psi,\ \nabla^2\psi) + \beta v). \tag{55}$$

Approximation of $\mathcal{H}$ by $(-1/D^2)$ and neglect of the first term on the right hand side constitute the conventional planetary geostrophic approximation, which is all that is technically required at the order of accuracy of the model: for both quasi-geostrophic scales and highly stratified (low Froude-number) flow advection by the divergent field is unimportant. The advection of potential vorticity (47) by the inverted velocity obtained from (54) and (55) constitutes a closed model.

The model is globally valid, provided either the Froude number or Rossby number is small everywhere, since the inversion now includes all the terms retained in the model described in section 2(c), and the balanced dynamics can be characterized by potential-vorticity advection in both small Froude-number and small Rossby-number regimes. The inversion is not singular if $f = 0$: gradient-wind balance is still well-posed, whereas (55) gives $\delta = 0$. However, it is unclear how useful this model, or really its stratified generalization, presented in section 5(b), would be for large-scale cross-equatorial flow. The disadvantages of this model over any of planetary geostrophy, quasi-geostrophy or the small Froude-number model, used separately in their appropriate regimes, are that the inversion problem is more difficult and energy and mass conservation are not assured.

The advantage of the family of models described above is not that it goes to higher order in Rossby number for planetary or synoptic-scale flow. Rather, it is that the models incorporate the important lowest-order physics at *both* planetary and synoptic scales in a conceptually simple and readily implementable form, and thereby provides a single framework for both synoptic and planetary-scale flow. Pedlosky (1984), Salmon (1985) and Shutts (1989), take different approaches to this issue. Pedlosky requires a scale *separation* between planetary and synoptic scales, and derives a coupled set of planetary synoptic equations; in contrast the model presented here is a single set of equations, most appropriate if the flow transitions relatively smoothly between planetary and synoptic regimes. Salmon and Shutts generalize the semi-geostrophic equations to allow a variable Coriolis parameter, and thus retain good conservation properties.

3. Some numerical results

Here we briefly report some numerical results in shallow-water systems. Our goal is to demonstrate the basic utility (or otherwise) of the procedures rather than to completely explore the parameter regime or provide especially realistic simulations. Extensive numerical experiments in more realistic settings (both atmospheric and oceanic) will be reported subsequently.

(a) *Inversions for synoptic-scale flow*

The basic approach is this. A shallow-water model (with prognostic variables u, v and h), is numerically integrated in various parameter regimes, until it is in a statistically steady state. Its potential vorticity is calculated. Balanced height and velocity fields are then calculated from the potential vorticity using some of the balance relationships derived above. These are compared with the true fields.

This set of experiments was designed to test the synoptic-scale inversion of section 2(a). This is valid for small Rossby number, with small variations of Coriolis parameter and small variations in the height field. This in turn necessitates that the scale of motion be not much larger than the deformation radius. Of the experiments performed, we report on two, one 'successful' and one not, as these are illustrative of the main points. Both have small Rossby number. We use a doubly-periodic model on an f-plane, to allow homogeneous flow and the cleanest test of the inversion, although the parameters were then chosen to have some geophysical relevance. The numerical code used is a gridded model, using an energy-enstrophy conserving scheme (Arakawa and Lamb 1981), with a resolution typically of 256×256 grid points. Dimensionally, the basin size is 5116 km in both north and east directions, giving a grid size of about 0.18°. The reduced gravity g' is 98.06 cm s^{-2}. The reference height H is 10 km, the fastest gravity wave has a speed about 100 m s^{-1}, and the Rossby deformation radius is about 961 km, or about or wavenumber 5. Initially, the model was spun up with a random forcing of vorticity (the forcing divergence was enforced to be zero) of isotropic wavenumbers 2–4; that is, somewhat larger than the deformation radius. (Experiments with moderately varying deformation radius did not lead to qualitative differences.) The inversions were tested when the model is in a forced-dissipative statistical steady state. The forcing was then turned off, and the inversions again applied to the slowly decaying fields.

Figure 1 shows various instantaneous fields obtained in the forced-dissipative regime. The Rossby number is estimated to be between 2×10^{-3}, using $U_{\rm rms}/fL$, and 3×10^{-2} using $\zeta_{\rm rms}/f$. The Reynolds number is approximately 10^4. The potential vorticity field in particular has a familial resemblance to fields produced in pure two-dimensional turbulence studies. The height field and the streamfunction are plainly close to geostrophic balance, although by no means exactly. Spectra of the fields indicate that the rotational energy roughly obeys a k^{-3} law at intermediate scales, whereas the gravitational spectrum is whiter at small scales.

Fields obtained from zeroth-order (i.e. quasi-geostrophic) and first-order inversion are shown in Fig. 2. In both cases the inverted height and streamfunction fields match fairly well with their true counterparts. However, there is no improvement in going to first order and the inverted divergence field is rather poor. That is, even though the Rossby number is small a quasi-geostrophic inversion is as good as the next order iteration; the rms difference between inverted and true field is not changed. A third inversion (not shown) scheme follows the scheme of section 2(c). This also shows no improvement over quasi-geostrophy.

The forcing was then turned off and the system allowed to decay under the action of a weak dissipation. Generally, the results obtained in such decaying motion were a little better than those of the forced-dissipative experiment. Significant improvement was obtained by artificially suppressing the divergence. In a stratified fluid, there is much more freedom for a fluid to radiate gravitational instabilities away. To parametrize such effects, a 'divergence damping' was added to the momentum equations. This is designed to preferentially damp divergence, having no direct effect on vorticity. After 70 days of such decay, the instantaneous fields are shown in Fig. 3. The inverted fields are shown in Fig. 4. In this case the higher-order inversion is noticeably better than a quasi-geostrophic inversion. The streamfunction and height fields of the first-order inversion are indistinguishable from the real field. The normalized r.m.s. error in the streamfunction (between the inverted and actual fields), which is 1.4×10^{-3} for the quasi-geostrophic inversion, falls to 5.3×10^{-5} using the first-order inversion. Using the inversion of section 2(c) also improves the estimate of streamfunction. As we have pointed out, this inversion is properly zeroth order because it essentially linearizes the potential vorticity. However, the additional use of the gradient-wind balance evidently improves the estimate of the streamfunction. (The nonlinear Monge–Ampère equation is solved iteratively.)

Two restricted conclusions may be drawn from these experiments. First it is clear that proceeding to asymptotically higher order in the manner described can in some circumstances give an improved calculation of the inverted fields. However, a higher-order calculation clearly does not always improve matters. In the forced-dissipative experiment,

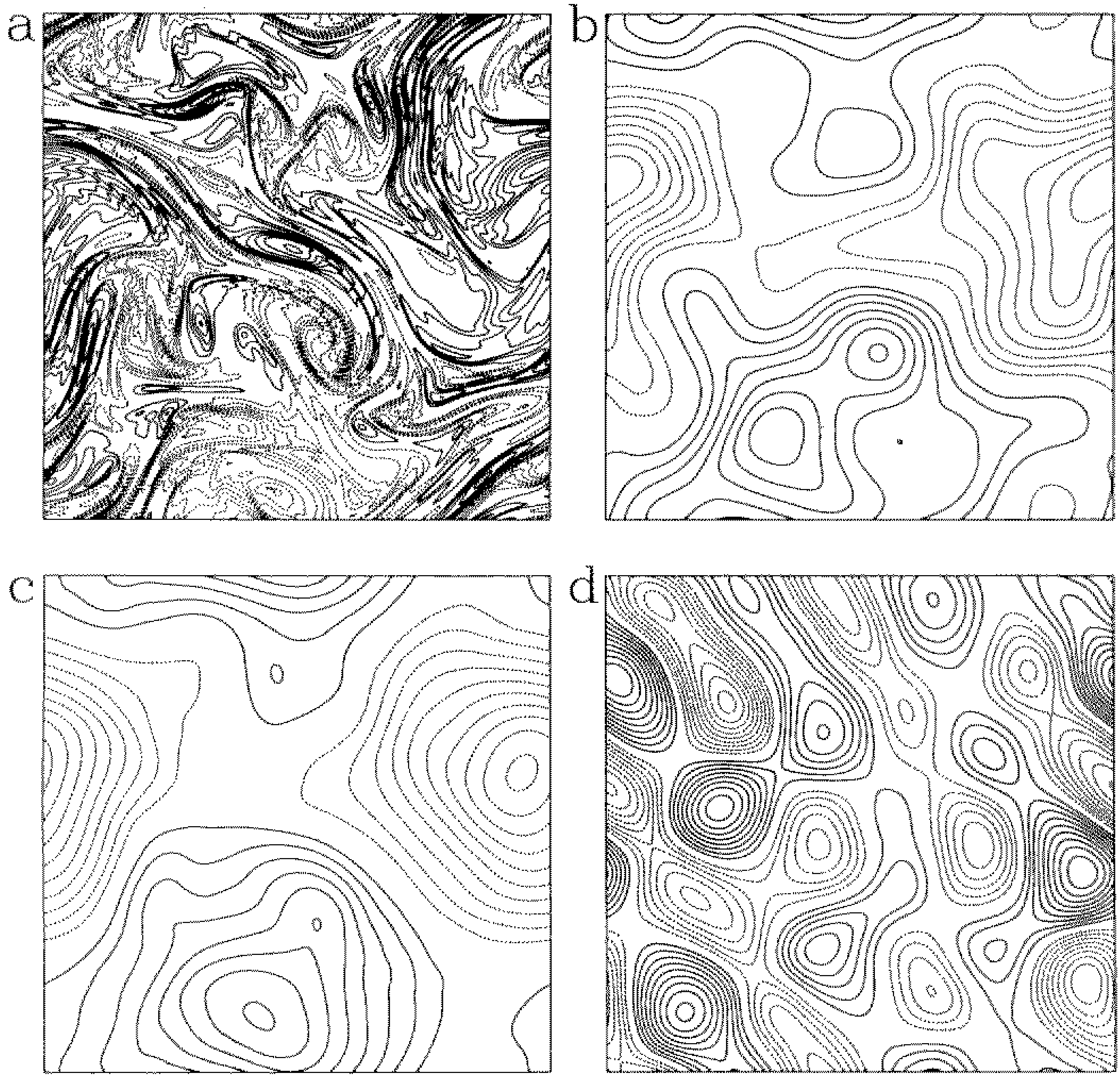

Figure 1. Typical instantaneous fields from forced-dissipative shallow-water integration. The geometry is f-plane, doubly-periodic, deformation radius about wavenumber 5. (a) Potential vorticity; (b) height; (c) streamfunction; (d) divergence.

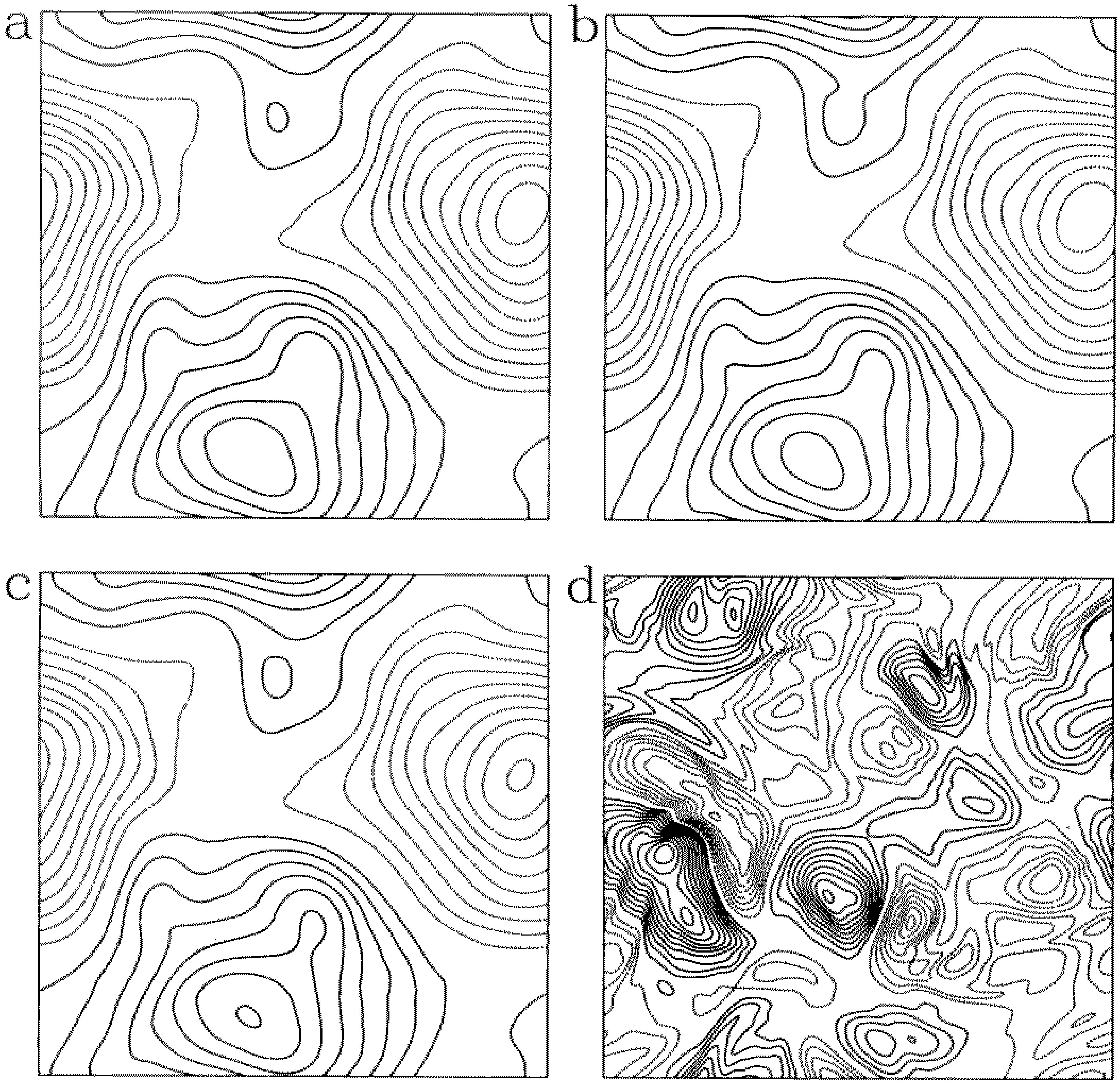

Figure 2. (a) Streamfunction (and therefore height) obtained from a zeroth order (quasi-geostrophic) inversion of the potential vorticity of Fig. 1. Panels (b), (c) and (d) are height, streamfunction and divergence fields obtained using the first order inversion of section 2(a).

the forcing maintained the divergence at a rather high level, although it is still much smaller than vorticity, and the inverted divergence field was over an order of magnitude smaller than the true field. In this case a higher-order inversion was no improvement over quasi-geostrophy.

(*b*) *Large-scale flows*

In these experiments a balanced flow with scales larger than the deformation radius, and with significant variations in both the height field and the Coriolis parameter, is created. For simplicity we retain Cartesian geometry. The *geometric* accuracy of a β-plane representation of spherical geometry is of course dependent on small variations in latitude, and thus small variations in f. This restriction should be differentiated from the *dynamical* restriction of small variations in the Coriolis parameter required in quasi-geostrophic

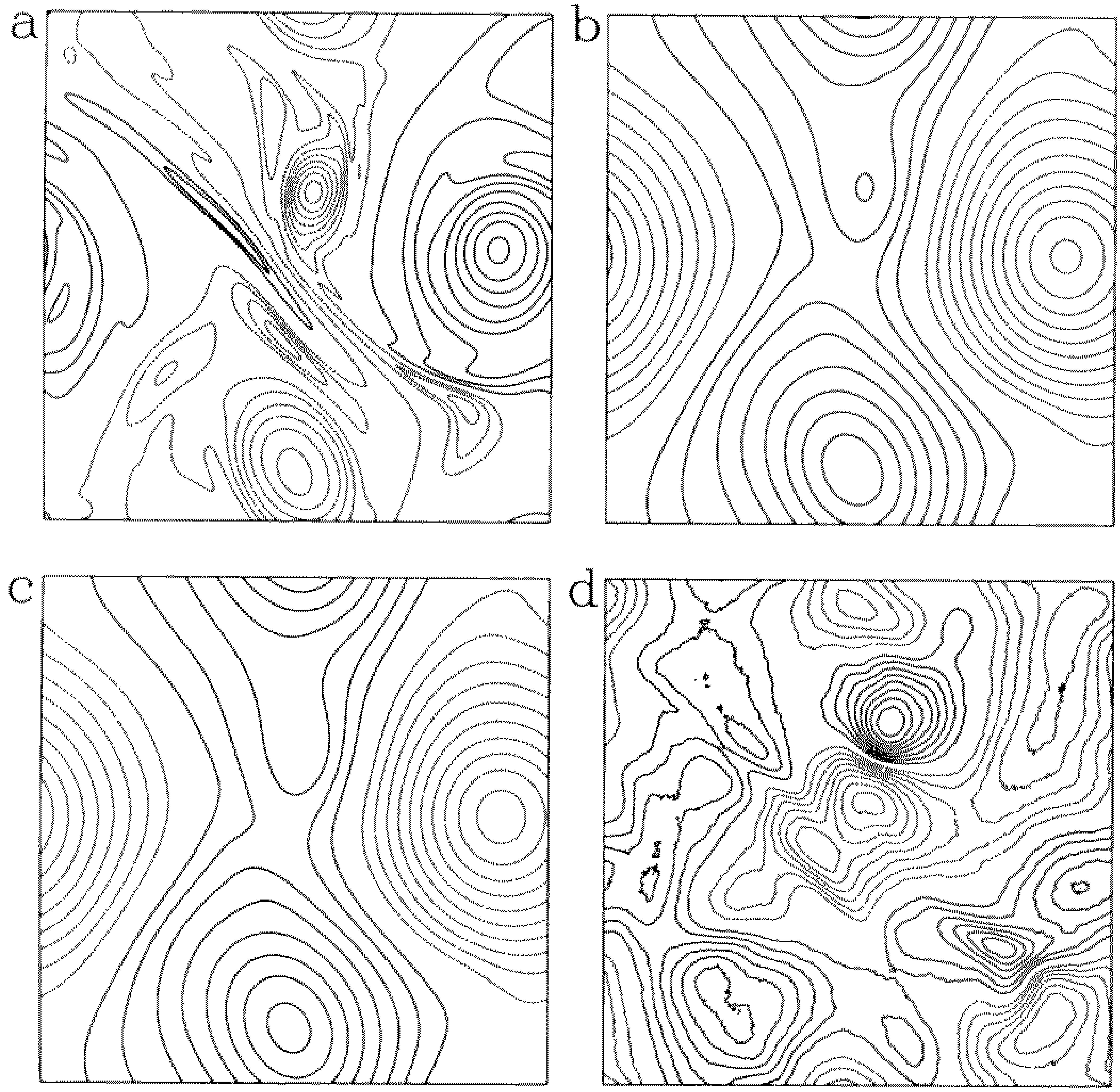

Figure 3. Instantaneous fields obtained from a decaying shallow-water integration. (a) Potential vorticity; (b) height; (c) streamfunction; (d) divergence (smoothed slightly for presentation purposes).

theory. If the geometry is Cartesian *ab initio* the dynamical effects of a large variation in Coriolis parameter may be studied independently of the geometric constraints.

First we perform an oceanically relevant experiment, and configure the shallow-water model in an closed square basin, of dimensional size $L = 4000$ km. The Coriolis parameter varies linearly with y. With $\beta = 2 \times 10^{-11}$ the variation of f is approximately equal to its mean value, 9×10^{-5}. The deformation radius is then about 42 km. The flow is forced with a 'two-gyre' wind-stress, proportional to $\cos 2\pi y/L$, and dissipated with a bottom (Ekman) friction and a weak lateral dissipation. The dynamical field generated has large height variations, the depth varying by about a factor of ten, from a thickness. of less than 100 m in the sub-polar gyre to over 800 m in the sub-tropical gyre. Figure 5 illustrates the height field when the model is in a near-steady state. A planetary geostrophic inversion of the height reproduces a qualitatively good streamfunction, except in the Western boundary layer where small-scale features are present. A quasi-geostrophic inversion is qualitatively poor, although it captures some small-scale features of the relative vorticity in the Western

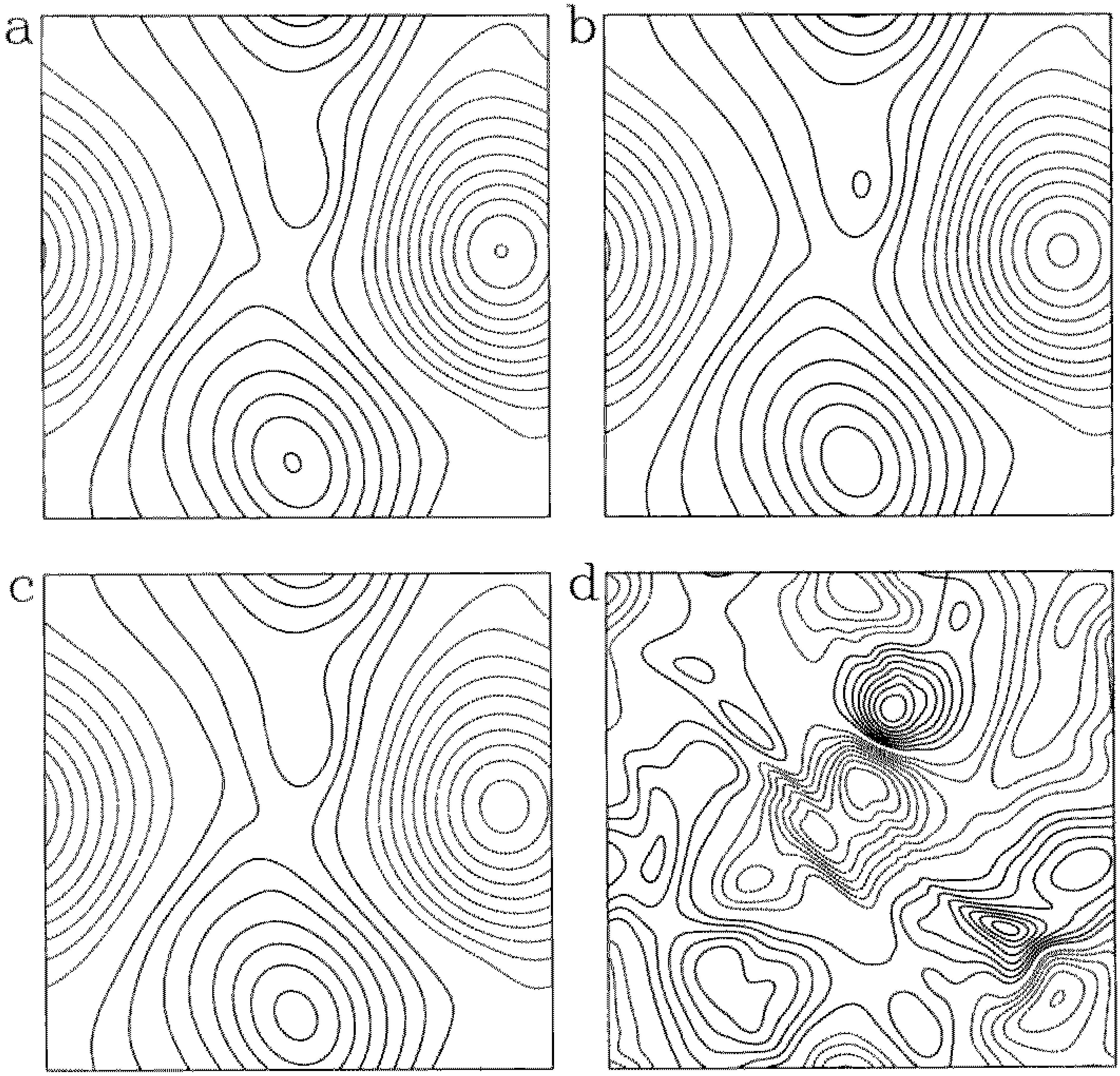

Figure 4. (a) Streamfunction (and therefore height) obtained from a zeroth-order (quasi-geostrophic) inversion of the potential vorticity of Fig. 3. Panels (b), (c) and (d) are height, streamfunction and divergence fields obtained using the first-order inversion of section 2(a).

boundary current missed by planetary geostrophy (Fig. 6). An inversion using the truncated planetary/synoptic model of section 2(e) gives a qualitative (and quantitative) improvement over both planetary- and quasi-geostrophy. It is able to capture the small-scale features at least as accurately as quasi-geostrophy, while giving an excellent reproduction of the large-scale height field (Figs. 5 and 6). The integrating effects of the elliptic operator in the inversion provides a smooth transition between the planetary and synoptic regimes. Inversions using the asymptotic first-order inversion (45) are poor (not shown) because the expression for the first-order height field magnifies the detail, rather than integrating over it.

We have also integrated the truncated model in time to ascertain whether it is a viable model for flow evolution. Preliminary results show that it is well-behaved and robust, and at least two or three times more efficient than a shallow-water model. For these experiments, we configured a shallow-water model, the truncated planetary/synoptic

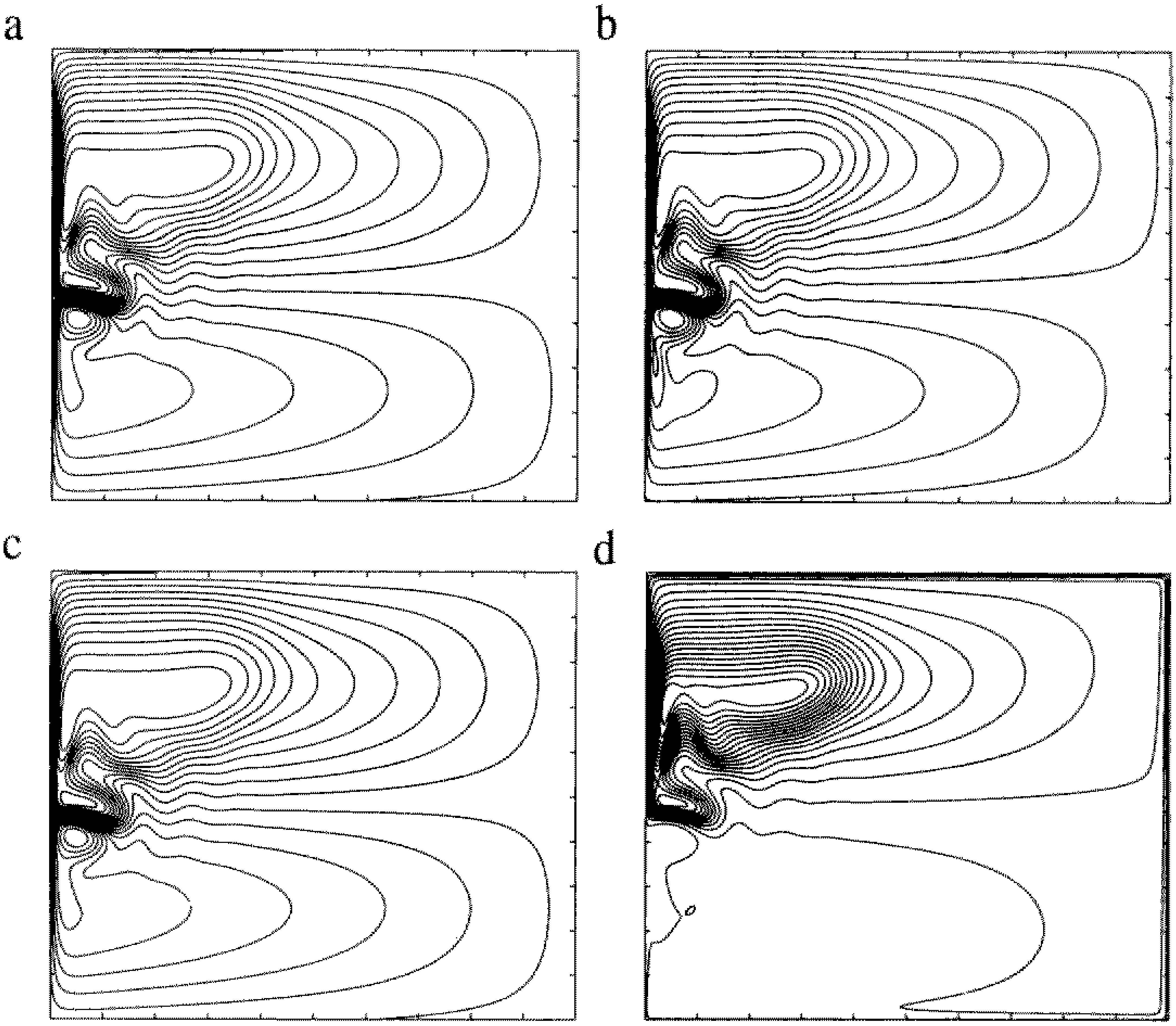

Figure 5. (a) The height field obtained by an integration of the shallow-water equations in a 'two-gyre' configuration. (b) The height field obtained from an inversion of the potential vorticity field of the shallow-water model using the planetary geostrophic equations (i.e. $h = f/Q$). Normalized rms difference between this and field (a) is 1.02×10^{-2}. (c) As for (b), but the inversion uses the planetary/synoptic truncated model (i.e. equation (50)). Rms difference is 1.59×10^{-3}. (d) As for (b), but with a quasi-geostrophic inversion. Rms difference is 0.51.

model, and a quasi-geostrophic model in a periodic channel on a β-plane. Model resolution and other physical parameters are identical. For small height variations and weak β-effect, the behaviour of all models is very similar, both in the physical characteristics of the flow field and in the decay of energy and enstrophy. For larger variations of height and Coriolis parameter the planetary/synoptic model performs noticeably better than quasi-geostrophy, and its evolution is in good agreement with that of a shallow-water model (Fig. 7). The dimensional parameters of this experiment are as follows: The domain size 6000×6000 km, and resolution 130×130 grid-points; $F_0 = 2\Omega \sin 45^\circ = 1.03 \times 10^{-4} \mathrm{s}^{-1}$; $\beta = (2\Omega/a) \cos 45^\circ = 1.6 \times 10^{-11}\ \mathrm{m}^{-1}\mathrm{s}^{-1}$; $H = 10$ km; $g' = 2.5\ \mathrm{m\,s}^{-2}$, $D = 1581$ km. With a velocity scale of $10\ \mathrm{m\,s}^{-1}$ (the initial velocity magnitude) this gives the following non-dimensional parameters: $\beta L/f_0 = 0.94$, $\varepsilon = 0.016$, $F_R = 15.3$, $F = 0.063$. The advective eddy turn-over time-scale is approximately seven days, and the Ekman spin-down time approximately 70 days. Since $\varepsilon F_R = 0.25 < 1$, the parameters violate quasi-geostrophy primarily through the large β-effect.

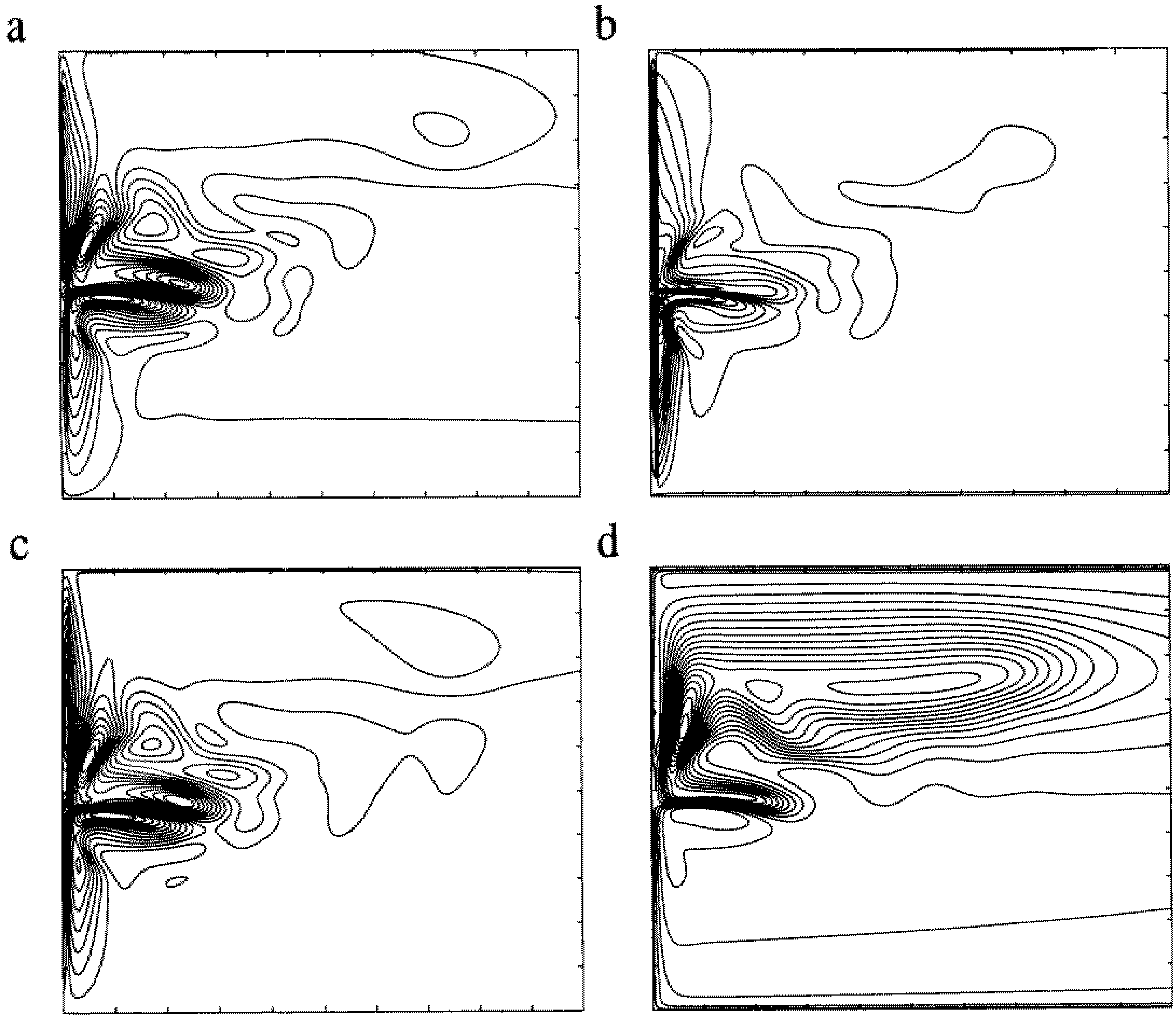

Figure 6. As for Fig. 5, but the panels show the relative vorticity in western half of model ocean. (a) The shallow-water integration. (b) The planetary geostrophic inversion; rms difference = 0.72. (c) The planetary/synoptic (truncated) model inversion; rms difference = 0.12. (d) The quasi-geostrophic inversion; rms difference = 0.99.

The quasi-geostrophic model shows quantitative and qualitative deficiencies. For example, compared to either of the two other models, the tendency toward zonality produced by the beta-effect appears unrealistically strong. Figure 8 illustrates the loss of correlation in the quasi-geostrophic simulation; in time, all the fields become de-correlated because of the loss of predictability in the flow. Overall, the preliminary results presented in Figs. 5 through 8 show a very good performance by the planetary/synoptic model. However, many further experiments are called for to investigate, for example, the formulation and influence of boundary conditions, and the model's actual utility in real physical problems.

4. The stratified primitive equations

We shall carry through some of the analysis previously done for the shallow-water equations for the stratified primitive equations. Where the analysis and scaling is similar, the presentation is brief. For simplicity we shall assume hydrostatic balance at all orders,

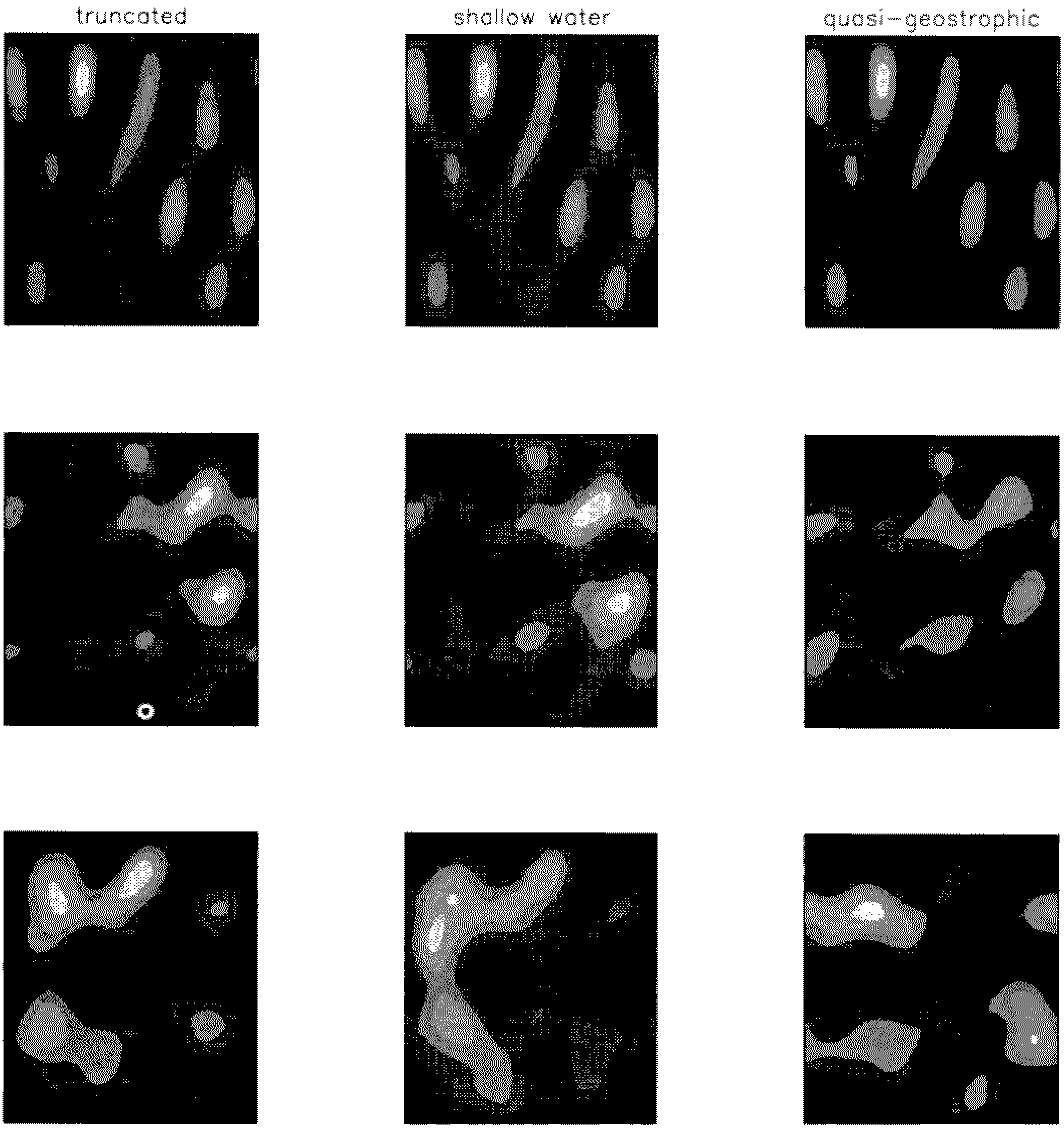

Figure 7. Evolution of the height field of the truncated planetary/synoptic model (left column), a shallow-water model (centre), and quasi-geostrophic model (right). The truncated model evolves potential vorticity, with velocities obtained using (50) and (51). The initial height field of all models (upper row) is identical; the lower two rows show height field at eddy-turnover times $t = 0.6$ and $t = 1.2$.

and use a Boussinesq approximation. Height and pressure coordinates are then essentially equivalent. The analysis holds with little difference for the slightly more general anelastic equations which allow density to vary in the vertical derivative in (58), needed to treat atmospheric motion of vertical extent similar to a density-scale height. We examine balanced systems at small Rossby number and small Froude number, derive a balanced system valid for small Froude number at all Rossby numbers, and then consider synoptic and planetary-scale flow in isentropic coordinates.

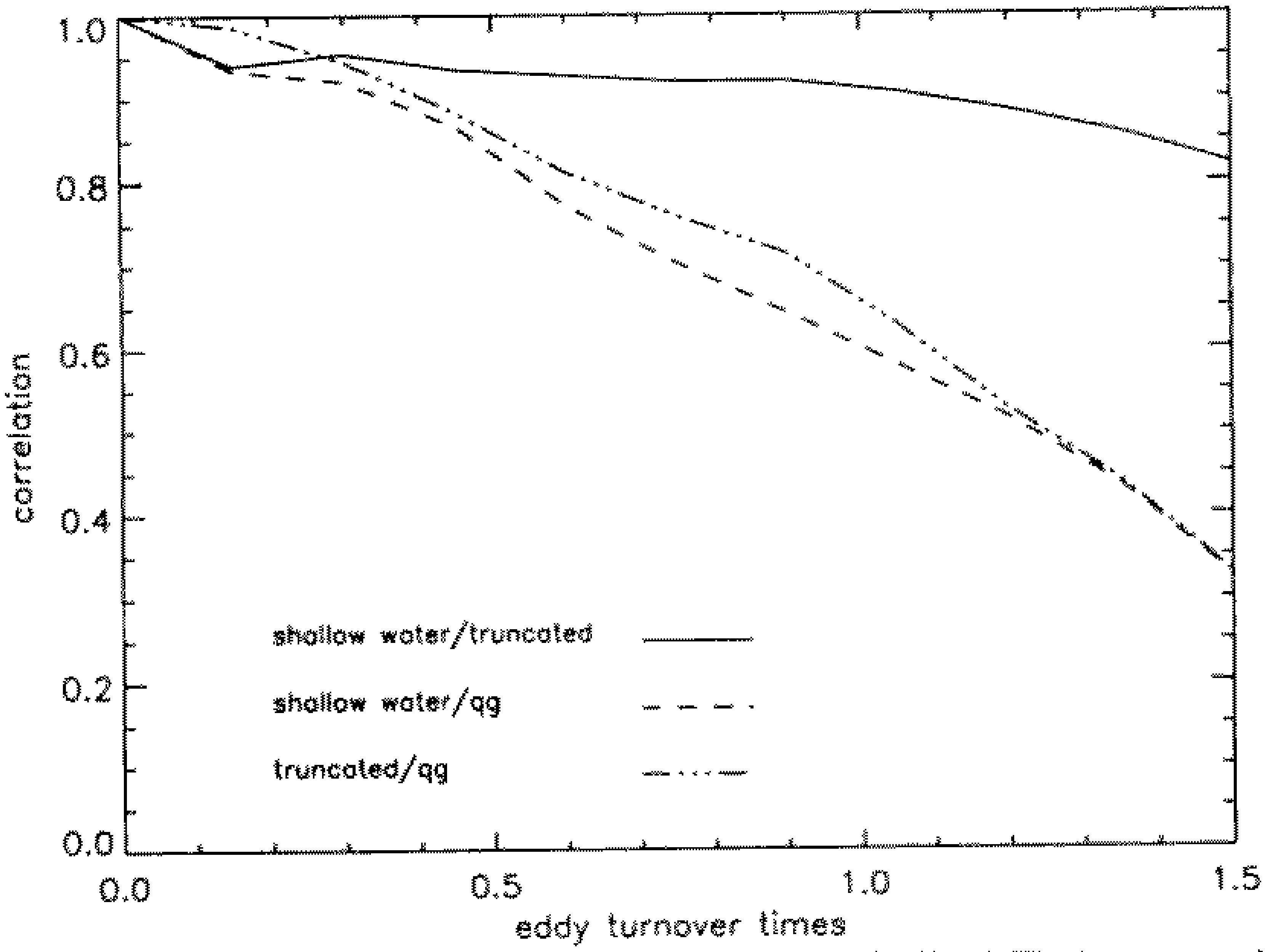

Figure 8. Evolution of the correlation between the three fields shown in Fig. 7. The largest correlation is maintained between the shallow-water model and the truncated planetary/synoptic model (upper curve). The two lower curves show the correlation between the quasi-geostrophic model and the shallow-water model, and between the quasi-geostrophic model and the planetary-synoptic (truncated) model.

The fundamental equations of motion are written, for unforced, inviscid flow:

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + f\mathbf{k} \times \mathbf{u} = -\nabla\phi \tag{56}$$

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0 \tag{57}$$

$$\nabla \cdot \mathbf{u} + \frac{\partial \omega}{\partial z} = 0 \tag{58}$$

$$g\frac{\theta}{\hat{\theta}} = \frac{\partial \phi}{\partial z}. \tag{59}$$

The notation is fairly standard. In a height co-ordinate system, ϕ is pressure (divided by a reference density), θ is deviation potential temperature and $\hat{\theta}$ is the reference profile. Alternately, z may be considered to be pressure and ϕ a geopotential. For oceanic applications θ is replaced by $-\rho$, but otherwise the equations are similar.

(*a*) *Rapidly rotating flow*

First, we carry out an asymptotic expansion valid for rapidly rotating flow, that is with small Rossby number. We recover the usual quasi-geostrophic approximation, and then carry through the analysis to higher order. The Coriolis parameter is kept constant, although it is not difficult to add a β-effect. Planetary-scale flow is considered in section 4(d).

Standard scaling (e.g. Pedlosky 1987) at small Rossby number leads to the following estimates:

$$\phi \sim fUL \tag{60}$$

$$\frac{\theta}{\hat{\theta}} \sim \frac{fUL}{gH} = \varepsilon F_{\mathrm{R}}, \qquad \frac{\theta_z}{\hat{\theta}_z} \sim \frac{F^2}{\varepsilon}. \tag{61}$$

Here, $F_{\mathrm{R}} = f^2L^2/gH$, and F is the Froude number now given by $F = U/NH$, where $N^2 = (g/\hat{\theta})\hat{\theta}_z$ and H is a characteristic height scale. The Rossby number, ε is, as usual, $\varepsilon = U/fL$. The basic assumption of the analysis is that $\varepsilon \ll 1$ and $F^2/\varepsilon \ll 1$, so that $F^2/\varepsilon^2 \leqslant O(1)$. The latter is more restrictive than $\varepsilon F_{\mathrm{R}} \ll 1$ since $|\hat{\theta}_z| < |\theta/H|$.

Potential vorticity is a Lagrangian invariant. That is,

$$\frac{\mathrm{D}Q}{\mathrm{D}t} = 0 \tag{62}$$

where

$$Q = (\omega + f_0\mathbf{k}) \cdot \nabla\theta \tag{63}$$

or

$$Q = (f_0 + \nabla^2\psi)(\hat{\theta}_z + \theta_z) - (v_z\theta_x - u_z\theta_y). \tag{64}$$

In height (or pressure) co-ordinates we have

$$\frac{\mathrm{D}Q}{\mathrm{D}t} = \frac{\mathrm{D}q}{\mathrm{D}t} + w\frac{d\hat{q}}{dz} = 0 \tag{65}$$

where $Q = q + \hat{q}$ and $\hat{q} = f_0\hat{\theta}_z$. One may regard q as the fundamental evolution variable.

At small Rossby number, the terms comprising Q scale as:

$$Q = f_0\hat{\theta}_z + \{\varepsilon\}\left[\hat{\theta}_z\nabla^2\psi + \left\{\frac{F^2}{\varepsilon^2}\right\} f_0\theta_z\right] - \{\varepsilon^2\}\left\{\frac{F^2}{\varepsilon^2}\right\}[(v_z\theta_x - u_z\theta_y) - \nabla^2\psi\theta_z]. \tag{66}$$

Expanding the dependent variables ψ, θ and so on in a series in ε, and substituting in (66), gives at lowest order,

$$Q = f_0\hat{\theta}_z + \{\varepsilon\}\left[\hat{\theta}_z\nabla^2\psi_0 + \left\{\frac{F^2}{\varepsilon^2}\right\} f_0\theta_{0z}\right]. \tag{67}$$

The system is closed by using hydrostasy, which presumptively is valid at all orders, and geostrophy, the lowest order diagnostic relationship between pressure and velocity. That is,

$$g\frac{\theta_0}{\hat{\theta}} = \frac{\partial\phi_0}{\partial z} \tag{68}$$

and

$$f_0\psi_0 = \phi_0 \tag{69}$$

which together give the thermal-wind relationship

$$f_0\frac{\partial\psi_0}{\partial z} = g\frac{\theta_0}{\hat{\theta}}. \tag{70}$$

Equations (67) and (70) essentially comprise the quasi-geostrophic approximation. They combine to give

$$q = \left[\hat{\theta}_z\nabla^2\psi_0 + \left\{\frac{F^2}{\varepsilon^2}\right\}\frac{f_0^2}{g}\frac{\partial}{\partial z}(\psi_{0z}\hat{\theta})\right]. \tag{71}$$

The lowest-order evolution equation is then

$$\frac{\partial q}{\partial t} + \mathbf{u}_0 \cdot q + w_1 \frac{\partial \hat{q}}{\partial z} = 0. \tag{72}$$

The vertical velocity w_1 ($w_0 = 0$) can be obtained from the 'omega' equation, obtained in a similar manner to (17), namely:

$$\mathcal{L} w_1 = f_0 \frac{\partial}{\partial z}(J(\psi_0,\ \nabla^2 \psi_0) + \beta v_0) - f_0 \nabla^2 \left(J \left(\psi_0,\ \frac{\partial \psi_0}{\partial z} \right) \right) \tag{73}$$

where $\mathcal{L} = f_0^2(\partial^2/\partial z^2) + N^2 \nabla^2$. However, w_1 may be completely eliminated with the aid of the thermodynamic equation:

$$\frac{\partial \theta_0}{\partial t} + \mathbf{u}_0 \cdot \nabla \theta_0 + w_1 \hat{\theta}_z = 0 \tag{74}$$

and the normal quasi-geostrophic result emerges. This derivation, direct from potential vorticity, serves to emphasize that the quasi-geostrophic potential vorticity is not an approximation to the Ertel potential vorticity, and would perhaps be more properly called a 'pseudo potential vorticity' (Charney and Stern 1962). The awkwardness involved in eliminating the vertical advection of the mean stratification is eliminated by a treatment in isentropic coordinates, although this has certain difficulties of its own.

(i) *A step beyond quasi-geostrophy*. We will now proceed asymptotically one order beyond quasi-geostrophy. Continuing the expansion of potential vorticity gives

$$\begin{aligned} Q = f_0 \hat{\theta}_z &+ \{\varepsilon\}[\hat{\theta}_z \nabla^2 \psi_0 + \{F^2/\varepsilon\} f \theta_{0z}] \\ &+ \{\varepsilon^2\} \left[\left(\hat{\theta}_z \nabla^2 \psi_1 + \left\{ \frac{F^2}{\varepsilon^2} \right\} f_0 \theta_{1z} \right) - \left\{ \frac{F^2}{\varepsilon^2} \right\} (v_{0z} \theta_{0x} - u_{0z} \theta_{0y} - \theta_{0z} \nabla^2 \psi_0) \right] \end{aligned} \tag{75}$$

Since the lowest-order fields are defined by (67), the coefficient of the ε^2 term must vanish, giving an expression for the first-order terms in terms of the zeroth-order terms:

$$\hat{\theta}_z \nabla^2 \psi_1 + \left\{ \frac{F^2}{\varepsilon^2} \right\} f_0 \theta_{1z} = \left\{ \frac{F^2}{\varepsilon^2} \right\} (v_{0z} \theta_{0x} - u_{0z} \theta_{0y} - \theta_{0z} \nabla^2 \psi_0). \tag{76}$$

The system is completed by hydrostasy:

$$g \frac{\theta_1}{\hat{\theta}} = \frac{\partial \phi_1}{\partial z} \tag{77}$$

plus a relationship between the pressure and streamfunction. Just as in the shallow-water case, this comes from the divergence equation

$$\frac{\mathrm{D}\delta}{\mathrm{D}t} + \delta^2 - 2J(u,\ v) + (w_x u_z + w_y v_z) + \nabla^2(\phi - f_0 \psi) = 0. \tag{78}$$

At zeroth order in Rossby number this gives:

$$\nabla^2(\phi_0 - f_0 \psi_0) = 0 \tag{79}$$

and at next order we get

$$\nabla^2(\phi_1 - f_0\psi_1) = 2J(\psi_{0x},\ \psi_{0y}). \tag{80}$$

Using (76), (77) and (80) we obtain,

$$\frac{\hat{\theta}_z}{f_0}\nabla^2\phi_1 + \left\{\frac{F^2}{\varepsilon}\right\}\frac{f_0}{g}\frac{\partial}{\partial z}(\hat{\theta}\phi_{1z}) = 2J(\psi_{0x},\ \psi_{0y}) + \left\{\frac{F^2}{\varepsilon}\right\}[v_{0z}\theta_{0x} - u_{0z}\theta_{0y} - \theta_{0z}\nabla^2\psi_0] \tag{81}$$

where the zeroth-order fields are given by quasi-geostrophy. Note that the elliptic operator on the left-hand-side is just the 'quasi-geostrophic' linear elliptic operator. Thus, going to higher order involves solution of two linear elliptic inversions, (80) and (81), in addition to the one involved in quasi-geostrophy. Note also that gradient-wind balance, and the additional nonlinear terms in the definition of potential vorticity, both enter at this order. The vertical velocity is zero at lowest order (i.e. $w_0 = 0$) and at first order it and the divergence are given by an omega equation. The advection of the potential vorticity by these should be included in any time-dependent model. Finally, we note that a higher-order asymptotic expansion does not, of course, extend the regime of validity of quasi-geostrophy in any way, although it may provide more accurate solutions where $\varepsilon \ll 1$ and $F^2/\varepsilon \ll 1$.

(*b*) *Highly stratified flow*

The balanced scaling for highly stratified flow is essentially that of Lilly (1983). The small parameter is now the Froude number. At moderate to large Rossby number the pressure and the buoyancy scale as:

$$\phi \sim U^2, \qquad \frac{\theta}{\hat{\theta}} \sim F_{\rm N}, \qquad \frac{\theta_z}{\hat{\theta}_z} \sim F^2 \tag{82}$$

where $F_{\rm N} = U^2/gH$. The divergence and the vertical velocity are small because of the stratification, being given by the estimates

$$\frac{W}{H} \sim \nabla\cdot\mathbf{u} \sim F^2\frac{U}{L}. \tag{83}$$

The terms in the expanded potential vorticity are then:

$$Q = f\hat{\theta}_z + \{\varepsilon\}\hat{\theta}_z\nabla^2\psi_0 + \{F^2\}f\theta_{0z} + \{F^2\varepsilon\}[\nabla^2\psi_0\theta_{0z} - (v_{0z}\theta_{0x} - u_{0z}\theta_{0y})]. \tag{84}$$

The lowest-order approximation in a Froude number (squared) expansion is:

$$Q = f\hat{\theta}_z + \{\varepsilon\}\hat{\theta}_z\nabla^2\psi_0 \tag{85}$$

which gives the simple inversion:

$$\nabla^2\psi_0 = \frac{Q}{\hat{\theta}_z} - f. \tag{86}$$

Note that ψ_0 is a function of z. Each level in the vertical then evolves according to the two-dimensional equation:

$$\frac{\partial Q}{\partial t} + \mathbf{u}_0\cdot\nabla Q = 0 \tag{87}$$

where $\mathbf{u}_0 = \mathbf{k} \times \nabla\psi_0$. That is to say we have

$$\frac{\partial}{\partial t}\nabla^2\psi_0 + +J(\psi_0,\ \nabla^2\psi_0 + f) = 0 \tag{88}$$

where f may vary fully, although it may not be so large that the Rossby number is small. There is no need to calculate the vertical velocity: at moderate or large Rossby number the term $wf\,\partial\hat{\theta}_z/\partial z$ is small compared to the horizontal advection of relative vorticity.

As in the shallow-water case, the pressure field can be diagnosed using the divergence equation. For small Froude-number scaling, and O (1) or greater Rossby number, (78) gives:

$$\nabla^2\phi_0 = \left\{\frac{1}{\varepsilon}\right\}\nabla\cdot(f\nabla\psi_0) + 2J(\psi_{0x},\ \psi_{0y}). \tag{89}$$

The buoyancy is then obtained from hydrostasy:

$$g\frac{\theta_0}{\hat{\theta}} = \frac{\partial\phi_0}{\partial z}. \tag{90}$$

The nature of the 'two-dimensional' turbulence is a little different from that which arises in rapidly rotating flow in a number of ways. In the latter, the streamfunction is pressure, whereas in the former pressure is determined from gradient-wind balance (89). Second, in stratified flow the flow at each level is uncoupled, whereas in the rotating case quasi-two-dimensional ('barotropic') flow normally comes about through neglect of the vortex stretching term $f\theta_z$ in the quasi-geostrophic expression for potential vorticity. This is valid when there is less vertical structure in the perturbation buoyancy than scaling suggests (i.e. $\theta_z \ll \theta/H$), or the domain depth is much less than the characteristic thickness H. There is then little structure in the vertical at all, a result characteristic of the barotropization effect of quasi-geostrophic turbulence. Finally, quasi-geostrophic barotropic dynamics are self-consistent, in that they do not necessarily lead to their own breakdown. On the other hand, in the stratified case, we can expect that large vertical shears of the horizontal flow may develop on the timescale of the motion itself. These will lead to a reduction of the vertical scale of the motion, which increases the Froude number and may cause the breakdown of the very scaling that leads to the flow, or via Kelvin–Helmholtz instability, a minimum vertical shear scale. Stratified two-dimensional turbulence may nonetheless be responsible for the observed shallowing of the atmospheric energy spectrum in the mesoscale (Lilly 1989; Maltrud and Vallis 1991).

(*c*) *Stratified, rotating flow*

Just as for the shallow-water equations, it is possible to combine the models for highly stratified and for rapidly rotating flow, and produce a model valid at low Froude number for all Rossby numbers in which the control of the interior fields by potential vorticity is made explicit. (See also McWilliams 1985, and Raymond 1992).

An approximation to the potential vorticity that manifestly includes both stratified and rapidly rotating limits is:

$$Q = f_0\hat{\theta}_z + \{\varepsilon\}\hat{\theta}_z\nabla^2\psi_0 + \left\{\frac{F^2}{\varepsilon_{\min}}\right\} f_0\theta_{0z}. \tag{91}$$

To close the set we need relationships between the dependent variables θ and ψ. The divergence equation gives:

$$\{\varepsilon_{\max}\}\nabla^2\phi_0 - \nabla\cdot f_0\nabla\psi_0 = 2\{\varepsilon\}J(\psi_{0x},\ \psi_{0y}). \tag{92}$$

In (91) and (92), $[\varepsilon_{\min}, \varepsilon_{\max}] = [\min(1, \varepsilon), \max(1, \varepsilon)]$. Hydrostatic balance (90) completes the inversion. For an evolutionary model, the vertical velocity is required to advect the mean stratification. This may in fact be obtained, to the accuracy of the model, from the conventional omega equation of quasi-geostropic theory (73), even in a high Rossby number or mixed regime. At low Rossby number the omega equation *is* first-order accurate (the lowest-order vertical velocity is zero), and would yield the correct vertical velocity to use. At low Froude number and moderate or high Rossby number, the omega equation may be inaccurate, but advection by the vertical velocity is unimportant compared to the horizontal advection of relative vorticity by the non-divergent flow. Knowing the vertical velocity gives the divergent velocity which may be used in the advective terms, although the correction will be small.

These equations have validity in regimes of low Rossby number and low Froude number. In either of those regimes simpler inversions (namely quasi-geostrophy and two-dimensional stratified flow) are in fact available which, in the respective regimes, are formally as accurate as (91) and (92). The utility of the model lies in that it spans the regimes, potentially an important advantage in constructing a model or diagnosing balanced fields for regimes where one or both of stratification or rotation may be important. Strictly, even though a gradient-wind balance equation is solved between pressure and streamfunction the model is not first-order accurate (where quasi-geostrophy is 'zeroth-order' accurate) even at small Rossby number, because of the neglect of the nonlinear terms in the potential-vorticity expansion. Also, f is not allowed a full variation, although a weak β-effect is easily allowed. Finally, the model appears to require the solution of a nonlinear Monge–Ampère equation rather than a simple linear elliptic equation. In any regime in which the equations are valid the equation is readily solvable by straightforward methods because the quasi-geostrophic inversion (71) always yields a good first approximation for the streamfunction. Forming a single expression for the potential vorticity in terms of streamfunction using (90), (91) and (92) gives, at any Rossby number

$$q = \hat{\theta}_z \nabla^2 \psi_0 + \left\{ \frac{F^2}{\varepsilon^2} \right\} \frac{f_0^2}{g} \frac{\partial}{\partial z} [\psi_{0z} \hat{\theta}) + \{\varepsilon\} 2\nabla^{-2} J(\psi_{0x}, \psi_{0y})]. \qquad (93)$$

(c.f. (34)). It is clear that the contribution of the Jacobian term is always small. At low Rossby number the Jacobian term on the right hand side of (92) is small. At low Froude number and moderate or large Rossby number the Jacobian term in (92) is not necessarily small, but its contribution to (93) is small because of the high stratification. In both cases a good estimate for the streamfunction is thus obtained from the solution of (71). The contribution of the Jacobian term may then be computed iteratively if required, although it should be small if the scaling is valid. The pressure, if required, is then obtained directly from (92). In no case where the scaling is respected should sophisticated numerical methods beyond Poisson-like solvers be needed to invert the potential vorticity.

5. Isentropic coordinates

(a) *Asymptotics and quasi-geostrophy*

For adiabatic flow, isentropic coordinates leads to a simplification in the potential vorticity evolution equation that can be exploited in understanding balanced dynamics (Charney and Stern 1962; Bleck 1973; Gent and McWilliams 1984; Hoskins *et al.* 1985).

The potential vorticity equation is then

$$\frac{DQ}{Dt} = 0 \qquad (94)$$

where

$$Q = \frac{f + \zeta}{-\partial p/\partial \theta}. \tag{95}$$

The procedures used above can be applied to this equation in a straightforward way. In particular, the derivation of quasi-geostrophic motion is particularly easy (c.f. Charney and Stern 1962). Their derivation was of quasi-geostrophy in height coordinates, beginning with (94) and (95). For adiabatic isentropic flow, vertical velocity is implicit. One need not go through the exercise of eliminating vertical velocity using the thermodynamic equation, or, rather, the exercise is trivial since $w_{(\theta)} = \mathrm{D}\theta/\mathrm{D}t = 0$. Instead, one must project the result onto flat (height) co-ordinates. If the isentropes are inclined only slightly from the horizontal, they directly project onto flat surfaces at lowest order. The quasi-geostrophic potential vorticity *is* then an approximation to the Ertel potential vorticity (95), and the advection of the former by the horizontal flow is an approximation to the advection of the latter along isentropes.

To obtain a first order system in 'physical space' (i.e. height or pressure) coordinates it does not particularly simplify matters to begin in isentropic coordinates, since one must then project the results onto the physical space coordinates with corresponding precision. If, however, a balanced model or inversion in isentropic coordinates is needed, the task is much simpler; derivations to any order are much the same as the task of going to the same order in the shallow-water equations. Thus, our exposition is brief; we give it for constant Coriolis parameter, although it is easy to allow f to have small variations.

The pressure field is written as the sum of a reference state, a function of θ only, plus a perturbation: $p = \hat{p}(\theta) + p'(x, y, \theta, t)$. Expanding the perturbation fields on the right-hand-side of (95) (i.e. $p' = p_0 + \varepsilon p_1 \ldots$ etc.) gives

$$Q = \frac{f_0}{\hat{\varrho}} + \{\varepsilon\}\left[\frac{\zeta_0}{\hat{\varrho}} - \left\{\frac{F^2}{\varepsilon^2}\right\} f_0 \frac{\varrho_0}{\hat{\varrho}^2}\right] + O(\varepsilon^2) \tag{96}$$

where $\varrho = -p_\theta$ and $\hat{\varrho} = -\hat{p}_\theta$. The scaling or non-dimensionalization uses $p'_\theta/\hat{p}_\theta \sim F^2/\varepsilon$, which is assumed small (Berrisford *et al.* 1993). The system is closed through the use of hydrostasy and geostrophy. Hydrostatic balance is

$$\frac{\partial M}{\partial \theta} = \Pi, \tag{97}$$

where $M = c_p T + gz$ is the Montgomery potential and $\Pi = c_p(p/p_s)^{R/c_p}$ is the Exner function. Given small deviations from the reference profile of pressure this becomes

$$\frac{\partial M'}{\partial \theta} = \hat{\Pi}_p p', \tag{98}$$

where $\hat{\Pi}_p = d\Pi(\hat{p})/dp = 1/(\theta\hat{\rho})$. We will assume this holds at all orders, and in particular that $\partial M_0/\partial\theta = \hat{\Pi}_p p_0$. Geostrophic balance gives

$$f_0\zeta_0 \equiv f_0 \nabla^2_{(\theta)} \psi_0 = \nabla^2_\theta M_0 \tag{99}$$

where ψ is the velocity streamfunction. Horizontal derivatives are taken at constant θ, rather than constant z or p, and thus there is no advection of the term $f_0/\hat{\varrho}$. These expressions constitute a quasi-geostrophic model in isentropic coordinates. Using (96), (98) and (99) gives the quasi-geostrophic inversion:

$$q = \frac{1}{f_0}\nabla^2 M_0 + \frac{f_0}{\hat{\varrho}}\frac{\partial}{\partial\theta}\left(\frac{1}{\hat{\Pi}_p}\frac{\partial M_0}{\partial\theta}\right) \tag{100}$$

where here $\{\varepsilon\}q = \hat{\varrho}Q - f_0$. The formulation requires $\varepsilon \ll 1$ and $F^2/\varepsilon \ll 1$, as for quasi-geostrophy in physical-space coordinates.

At next order in the potential vorticity expansion, the constraint that $O(\varepsilon^2)$ terms vanish leads to,

$$\zeta_1 - f_0\varrho_1 = q\varrho_0. \tag{101}$$

The first order streamfunction and Montgomery potential are given through the use of hydrostasy, $\partial M_1/\partial\theta = \hat{\Pi}_p p_1$, and gradient wind balance,

$$\nabla^2 M_1 = f_0\nabla^2\psi_1 + 2J(\psi_{0x},\ \psi_{0y}). \tag{102}$$

The divergent field is obtained from an omega equation. This obtained by combining the vorticity and thermodynamic equation to form an imbalance equation similar to (8)—i.e. essentially one eliminates the time-derivatives between the vorticity equation and thermodynamic equations using geostrophy and hydrostasy. After a little algebra (and with the β-effect included) this gives

$$\mathcal{G}\delta = f_0\frac{\partial^2}{\partial\theta^2}\left(J(\psi_0,\ \nabla^2\psi_0) + \beta\frac{\partial\psi_0}{\partial x}\right) - f_0\nabla^2\left(J\left(\psi_0,\ \frac{\partial^2\psi_0}{\partial\theta^2}\right)\right) \tag{103}$$

where $\mathcal{G} = -(f_0^2(\partial^2/\partial\theta^2) + \hat{\varrho}\hat{\Pi}_p\nabla^2)$. Note the similarity with (17). Taken together with the lowest-order system, these provide a closed first order system.

The use of an isentropic model does not in itself alleviate the quasi-geostrophic need for a 'reference profile' of stratification, or now a reference profile of height or pressure. Although the implied reference profiles need no longer be near the physical horizontal, they, and f, should be time and horizontal coordinate independent, or at most slowly varying, on the isentropic surfaces. If not, $f/\hat{\varrho}$ will be the leading order term in the potential-vorticity expansion, and equations similar to planetary geostrophy will perforce arise at lowest order.

(*b*) *A model for large-scale flow*

A truncated approach provides an attractive alternative to strict asymptotics, especially for large-scale motion, where, as in section 2(e), we are able to derive a model valid for both planetary-scale and synoptic-scale flow. The potential vorticity is written

$$Q = \frac{f + \zeta_g}{-\partial p/\partial\theta} \tag{104}$$

where $f\zeta_g - \beta u = \nabla^2_{(\theta)}M$. Then, using hydrostasy, we obtain the following elliptic equation,

$$\frac{1}{f}\left(\nabla^2_{(\theta)}M - \beta\frac{\partial M}{\partial y}\right) + \hat{Q}\frac{\partial^2 M}{\partial\theta^2} = -f. \tag{105}$$

where $\hat{Q} = Q/(\mathrm{d}\Pi/\mathrm{d}p)$ and

$$\frac{\mathrm{d}\Pi}{\mathrm{d}p} = \frac{R}{p_s}\frac{c_p^{1/(\gamma-1)}}{(\partial M_0/\partial\theta)^{1/(\gamma-1)}}, \tag{106}$$

where $\gamma = c_p/c_v$, bringing a weak nonlinearity to the inversion. There is a similarity with (50) (but which is a strictly linear equation), and with Bleck's (1973) model. The advecting velocities are obtained from geostrophic balance,

$$u = -\frac{1}{f}\left(\frac{\partial M}{\partial y}\right)_\theta, \qquad v = \frac{1}{f}\left(\frac{\partial M}{\partial x}\right)_\theta. \tag{107}$$

This model is valid (i.e. consistent and properly balanced) for both planetary and synoptic scales, provided either planetary geostrophic or quasi-geostrophic scaling holds. At large scales the dominant balance is that of the planetary geostrophic equations, whereas at the synoptic scales the model may be asymptotically reduced to the quasi-geostrophic equations. For flow with both synoptic and large-scale variations, common in both ocean and atmosphere, the equations can thus be expected to be an improvement over both quasi-geostrophy, since no reference profile is used and f is not restricted to small variations, and over planetary geostrophy, since relative vorticity is included. The model is not demonstrably valid if large amplitude variations occur on a synoptic scale.

As in the shallow-water case, an improvement in the accuracy of the velocity inversion at synoptic scales can be achieved by using gradient-wind balance instead of geostrophy. Thus, the model consists of advection of potential vorticity (95), with a velocity inversion achieved using gradient-wind balance,

$$\nabla^2 M - f\nabla^2\psi + \beta u - 2J(\psi_x,\ \psi_y) = 0, \tag{108}$$

along with hydrostasy and an omega equation. The nonlinearity now arising in the Monge–Ampère equation (108) is superficially non-trivial, but at small Rossby number the contribution of the Jacobian term is small and the equation can be solved iteratively. These equations are then valid globally, provided Froude number is small if the Rossby number becomes large, and accurate to lowest order in Rossby or Froude number. See also section 2(e).

Although it is clear that a number of advantages and simplifications of physical interpretation accrue through the use of isentropic coordinates, well-known problems with the horizontal boundary conditions have so far prevented their general adoption, balanced or not, as a forecast tool in a global domain, although Bleck (1974) achieved not inconsiderable success.

6. Discussion

A goal of this paper has been to explicitly utilize the central role of potential vorticity in the evolution of balanced dynamics, and to demonstrate the relative ease and consistency whereby appropriate models can then be rationally constructed. Thus, we consider a balanced model to be one determined by the evolution of potential vorticity, with various diagnostic relationships determining the pressure and velocity fields necessary for the evolution. A proper state of balance is assured if the time derivatives in the remaining prognostic equations (e. g. height or imbalance and divergence equations) are indeed small under the appropriate scaling. From this point of view, a state of balance is *equivalent* to the statement that potential vorticity can be 'inverted', and the degree of balance is related to the achievable accuracy of an inversion. Although other variables, such as pressure, could in principle be used as the primary inversion variable, potential vorticity is preferred because it is purely an advected field, and does not partake (in the linear approximation) of gravity-wave oscillations. Thus, in partially unbalanced situations, the instantaneous pressure field may contain a large component due to gravity waves, whereas the potential vorticity will not. Inversion of potential vorticity, and potential vorticity based initializations of primitive equation models, are thus more natural and likely more accurate than schemes based on other variables. The approximate evolution of a balanced system can likewise be effectively parametrized by potential vorticity advection and inversion, coupled as necessary with boundary advection of potential temperature.

It is often assumed that balanced dynamics will normally only be meaningful in the existence of a time or space-scale separation, which itself generally arises from the presence of a small parameter, such as Rossby number or Froude number. Utilizing such a small parameter in an asymptotic expansion is then one way to construct the inversion schemes. The advantages of an explicit asymptotic approach are two-fold. First, it allows approximations to be rationally carried to any order required. Second, in an asymptotic approach the defining dynamical balances and the range of applicability of the particular balance model are clearly delineated. The first of these points contrast with less formal truncated approaches. However, the latter allow more freedom in model construction, and a truncated model based on potential vorticity conservation (as in sections 2(e) and 4(d)) is arguably as rational as an asymptotic approach. A properly balanced model, containing one evolution variable plus constraints, can be assured either way. Certain other approaches, namely the bounded derivative method (e.g. Daley 1991) and the iterative methods employed by Allen (1993) are similar to the asymptotic methods. In particular, Allen's iterative method allows extension to high order, and the inversion formulae themselves similarly require multiple solutions of linear elliptic equations.

The derivations presented in this paper are straightforward. We considered both the shallow-water and stratified primitive equations, both for low Froude number (highly stratified) and low Rossby number (rapidly rotating). At low Rossby number the analysis was carried to first order for both, extending the shallow-water derivations of Warn *et al.* (1994). At the order beyond quasi-geostrophy, gradient-wind balance and additional terms in the expression for potential vorticity arise. Numerical implementation would not be difficult, primarily requiring the solution of additional linear elliptic equations at each timestep, although here we have given rather short-shrift to the correct formulation of boundary conditions. At low Froude number the stratified equations straightforwardly yield two-dimensional flow as the lowest-order approximation in Froude number, a result previously noted by Lilly (1983). It is also possible to derive a balanced model valid for highly stratified flow, whether rapidly rotating or not, for which the inversion involves only the solution of linear elliptic equations with constant coefficients. The model is appropriate for regimes, such as the atmospheric mesoscale or the oceanic sub-mesoscale, where either or both stratification and rotation can be important.

A family of models was also presented for the shallow water equations, and for the primitive equations in isentropic coordinates, which is valid for both planetary and synoptic scales, in that it combines the planetary geostrophic and quasi-geostrophic equations. Its dominant balance is that of the former at very large scales, and that of the latter at synoptic scales. It allows large variations of the height field and of the Coriolis parameter, yet also properly retains the effects of relative vorticity advection. The potential vorticity inversion is a linear (in the shallow-water case) or near linear (in the stratified, isentropic case) elliptic equation, although with non-constant coefficients, and numerical solution is quick.

Our preliminary numerical results indicate that, although proceeding to higher order can reduce errors, in some circumstances it in fact provides little gain. A Rossby number expansion is asymptotic, and not necessarily convergent, and thus provides no assurance that adding more terms to the expansion will improve accuracy, even in cases where the original problem has a solution—meaning, in this case, when the dynamics are essentially balanced. Even when Rossby number is quite small, ~ 0.1, the higher-order schemes examined here did not always perform significantly better than a quasi-geostrophic inversion. We may infer that for regions that are not highly balanced (perhaps the atmospheric mesoscale, or oceanic submesoscale) there will be little to be gained in going to higher order; more important will be to correctly capture the correct lowest-order physics of the balance flow. For large and synoptic-scale flow, the truncated model presented in section 2(e), equation (49), and the stratified version of section 4(d), are attractive. They contain all the dynamics of the planetary geostrophic and quasi-geostrophic equations as special cases, conserve mass and potential vorticity, and are relatively simple to implement, requiring the solution of but one elliptic equation, (50) or (105). Preliminary numerical solutions are encouraging, showing notable improvements over both quasi-geostrophy and planetary geostrophy; more extensive numerical simulations will be subsequently reported.

So-called intermediate models, whether derived asymptotically, iteratively, or by truncation, should, plainly, still respect the scaling which enabled the expansion and which produces the lowest-order system. In particular, proceeding to higher order requires that the lowest-order model be at least qualitatively if not quantitatively valid. A high-order asymptotic procedure cannot correct an inaccuracy arising from a scaling violation at lower order. For example, proceeding to higher order than quasi-geostrophy does not, *per se*, allow one to relax the constraint on the small variation of the Coriolis parameter or the height field, for then the lowest-order balance is incorrect. Other limitations of asymptotic schemes are the absence of automatic integral conservation laws, and possible difficulties in implementing boundary conditions. Various integrations (e.g. Allen and Newberger 1993) have indicated that integral conservation properties are not always to be preferred over higher accuracy, certainly for initial value problems. For long-term simulations it is not clear whether it is important to exactly maintain approximate conservation properties. Perhaps it suffices to place bounds on the degree of departure from exact conservation. This problem is related to whether an accurate inversion necessarily leads to accurate dynamics. Other problems include rationally extending the validity of balance beyond that enabled, as in this paper, by small Rossby number and/or small Froude number: balanced flow may still exist when no small parameter is obvious and an asymptotic approach difficult.

The atmosphere and ocean are plainly balanced to some degree, but not perfectly. Furthermore, on differing scales and locations the dynamics are balanced for varying reasons, with different terms contributing to the dominant balance. In my opinion, the most practically useful balanced models will be those which maintain their applicability over a wide parameter range, rather than those which just go to a higher order of accuracy in a particular parameter regime. Of the models presented in this paper, the model valid at small Froude number for all Rossby numbers, and the model valid at both planetary and synoptic scales, fall into this category. The overall utility of a model will also depend on the analytic tractability of its equations. Quasi-geostrophy and semi-geostrophy evidently perform well on this count.

Acknowledgements

I am very grateful to Jian Wang and Michael Mundt for help with the numerical integrations, and to an anonymous referee for a thorough review. This work was funded by the NSF Division of Atmospheric Sciences, and by ONR, Ocean Science Division.

References

Allen, J. S. | 1993 | Iterated geostrophic intermediate models. *J. Phys. Oceanog.*, **23**, 2447–2461

Allen, J. S. and Newberger, P. | 1993 | On intermediate models for stratified flow. *J. Phys. Oceanog.*, **23**, 2462–2486

Arakawa, A. and Lamb, V. R. | 1977 | A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon. Weather Rev.*, **109**, 18–36

Berrisford, P., Marshall, J. C. and White, A. A. | 1993 | Quasi-geostrophic potential vorticity in isentropic co-ordinates. *J. Atmos. Sci.*, **50**, 778–782

Bleck, R. 1973 Numerical forecasting experiments based on the conservation of potential vorticity on isentropic surfaces. *J. Appl. Meteorol.*, **12,** 737–752

1974 Short-range prediction is isentropic coordinates with filtered and unfiltered numerical models. *Mon. Weather Rev.*, **102,** 813–829

Charney, J. G. 1948 On the scale of atmospheric motion. *Geof. Publik.*, **17,** (2) 3–17

Charney, J. G. and Stern, M. E. 1962 On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19,** 159–162

Craig, G. C. 1993 A scaling for the three-dimensional semi-geostrophic approximation. *J. Atmos. Sci.*, **50,** 3350–3355

Daley, R. 1991 *Atmospheric Data Analysis*, Cambridge University Press. 457 pp.

Gent, P. R. and McWilliams, J. C. 1983 Consistent balanced models in bounded and Periodic domains. *Dyn. Atmos. Oceans*, **7,** 67–93

1984 Balanced models in isentropic co-ordinates and the shallow water equations. *Tellus*, **36A,** 166–171

Green, J. S. A. 1970 Transfer properties of the large-scale eddies, and the general circulation of the atmosphere. *Q. J. R. Meteorol. Soc.*, **96,** 157–185

Hoskins, B. J. 1975 The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, **32,** 233–242

Hoskins, B. J., McIntyre, M. E. and Robertson, A. 1985 On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.*, **111,** 887–946

Lilly, D. K. 1983 Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.*, **40,** 749–761

1989 Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.*, **46,** 2026–2030

Maltrud, M. E. and Vallis, G. K. 1991 Energy spectra and coherent structures in forced two-dimensional and beta-plane turbulence. *J. Fluid. Mech.*, **228,** 321–342

McWilliams, J. C. 1985 A uniformly valid model spanning the regimes of geostrophic and isotropic, stratified turbulence: balanced turbulence. *J. Atmos. Sci.*, **42,** 1773–1774

Norton, W. 1988 'Balance and potential vorticity inversion in atmospheric dynamics'. Ph.D. thesis. University of Cambridge, UK

Pedlosky, J. 1984 The equations for geostrophic motion in the ocean. *J. Phys. Oceanog.*, **14,** 448–455

1987 *Geophysical Fluid Dynamics*, 2nd ed. Springer-Verlag, 710 pp.

Phillips, N. A. 1963 Geostrophic motion. *Rev. Geophysics*, **1,** 123–176

Raymond, D. J. 1992 Nonlinear balance and potential-vorticity thinking at large Rossby number. *Q. J. R. Meteorol. Soc.*, **118,** 987–1015

Rhines, P. B. and Young, W. R. 1982 Homogenization of potential vorticity in planetary gyres. *J. Fluid. Mech.*, **122,** 347–367

Salmon, R. 1985 New equations for nearly geostrophic flow. *J. Fluid Mech.* **153,** 461–477

Shutts, G. J. 1989 Planetary geostrophic equations derived from Hamilton's principle. *J. Fluid. Mech.*, **208,** 545–573

Snyder, C., Skamarock, W. C. and Rotunno, R. 1991 A comparison of primitive-equation and semi-geostrophic simulations of baroclinic waves. *J. Atmos. Sci.*, **48,** 2179–2194

Warn, T., Bokhove, O., Shepherd, T. G. and Vallis, G. K. 1994 Rossby number expansions, slaving, and balance dynamics. *Q. J. R. Meteorol. Soc.*, **121,** 723–739

Xu, Q. 1994 Semibalance model—connection between geostrophic type and balanced-type intermediate models *J. Atmos. Sci.*, **51,** 953–970