REVIEW

The Quasigeostrophic Omega Equation: Reappraisal, Refinements, and Relevance

HUW C. DAVIES

Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

(Manuscript received 29 March 2014, in final form 1 September 2014)

ABSTRACT

A two-component study is undertaken of the classical quasigeostrophic (QG) omega equation. First, a reappraisal is undertaken of extant formulations of the equation’s so-called forcing function. It pinpoints shortcomings of various formulations and prompts consideration of alternative forms. Particular consideration is given to the contribution of flow deformation to the forcing function, and to the role of the advection of the geostrophic flow by the thermal wind (the \( R \) vector). The latter is closely related to the \( Q \) vector, the horizontal component of the ageostrophic vorticity, and the forcing function itself. The reexamination promotes further examination of the physical interpretation and diagnostic use of the omega equation particularly for assessing richly structured subsynoptic flow features.

Second, consideration is given to the dynamics associated with the equation and its more general utility. It is shown that the \( R \) vector is intrinsic to a quasigeostrophic cascade to finer-scaled flow, and that a fundamental feature of the QG omega equation—the in-phase relationship between cloud-diabatic heating and the attendant vertical velocity—has important potential ramifications for the assimilation of data in numerical weather prediction (NWP) models. Finally, it is shown that, in the context of considering NWP model output, mild generalizations of the quasigeostrophic \( R \) vector retain interpretative value for flow settings beyond geostrophy and warrant consideration when addressing some contemporary NWP challenges.

1. Introduction

In essence the quasigeostrophic (QG) omega equation enables a qualitative indication to be gained of the major synoptic-scale ascent regions from inspection of the contemporaneous synoptic charts. Furthermore, it permits a first-order quantitative estimation of the vertical velocity field to be computed given the specification of the three-dimensional geostrophic flow field. One indication of its diagnostic value follows from noting that strong low-level ascent can promote cyclogenesis and favor the occurrence of precipitation.

The omega equation’s derivation was the culmination of a half-century quest that had been plagued by the almost-geostrophic character of synoptic-scale flow. Ramifications of such a flow state are that 1) the vertical velocity field is comparatively weak and 2) its direct computation from the horizontal divergence field is beset both by the need to determine the comparatively small ageostrophic departure of the flow away from geostrophy (Exner 1902; Jeffreys 1919) and by the lack of adequate observational data (Margules 1904).

The seeming impasse posed by these ramifications was elegantly circumvented by the flowering [see the reviews of Eliassen (1984), Phillips (1990), and Bosart (1999)] of the concept of quasigeostrophy, which hinges upon the decomposition of the flow into a primary geostrophic component and a secondary ageostrophic component. Pivotal to its emergence and consolidation were two ingredients: first, the realization that there is an ageostrophic component that can be inferred from and “forced” directly by the geostrophic flow (Sutcliffe 1938, 1939, 1947) and second, the deduction that there is a quasigeostrophic flow whose evolution is determined by the geostrophic flow itself (Charney 1948; Eliassen 1949).

An approximate form of the QG omega equation was set out by Bushby (1952), more formal derivations were
supplied by Fjortoft (1955) and Thompson (1961), and subsequent reformulations were developed by, among others, Wiin-Nielsen (1959), Hoskins et al. (1978), and Hoskins et al. (2003).

The equation has become a staple aid to diagnosing the nature, structure, and dynamics of synoptic-scale flow phenomena. Early on it provided theoretically based succor for two empirically based nascent concepts of cyclogenesis (see, e.g., Petterssen et al. 1962) that were linked respectively to significant low-level thermal advection and strong upper-level vorticity advection, and later these concepts become the basis for the erstwhile two-type classification of cyclogenesis (Petterssen and Smeby 1971). Thereafter the equation has in its various formulations been used, for example, to diagnose the dynamics of individual events (see, e.g., Boyle and Bosart 1986; Rolfson and Smith 1996; Pedder 1997; Bracken and Bosart 2000; Strahl and Smith 2001); supplement NWP guidance (e.g., Barnes 1985; Barnes and Colman 1994); distinguish between the dynamics of baroclinic synoptic-scale development and subsynoptic across-frontal circulation patterns (e.g., Keyser et al. 1989; Martin 1999); explore the relative amplitude and nature of the forcing at different elevations (e.g., Trenberth 1978; Durran and Snellman 1987); infer qualitatively the major regions of ascent (e.g., Hoskins and Pedder 1980; Sanders and Hoskins 1990); examine the contribution of diabatic heating to cyclogenesis (e.g., Chang et al. 1984; Tsou et al. 1987; Strahl and Smith 2001); study jet flow and frontogenesis at upper levels (Cammas and Ramond 1989; Lang and Martin 2010); establish classification schemes for cyclogenesis based upon uppersynoptic versus lower-level forcing of vertical motion (Clough et al. 1996; Deveson et al. 2002; Gray and Dacre 2006); and associate distinctive potential vorticity features of the ambient flow with the vertical velocity field (Hoskins et al. 2003; Dixon et al. 2003; Plant et al. 2003). It has also often been adopted as a diagnostic tool in oceanographic studies.

Nowadays it could be argued that recourse to the OG omega equation is obsolete. Current high-resolution research and numerical weather prediction (NWP) models deliver the vertical velocity field as a model-output variable, but the resulting patterns are often somewhat noisy in appearance and not always readily interpretable. Again NWP models now operate at subsynoptic spatial scales and the attendant flow dynamics is beyond the strict validity of quasigeostrophy. Indeed a range of higher-order versions of the omega equation have been formulated to supplant its use in diagnosing the secondary ageostrophic flow. Also, many of today’s pressing dynamical and NWP challenges (such as predictability, data assimilation, and the representation/parameterization of near-grid-scale processes) appear to be substantially unrelated to diagnosis undertaken with the OG omega equation.

In light of the foregoing the objective of the present study is twofold. Following some preliminaries (section 2), a synthesis and review is provided of the various extant formulations of the OG omega equation along with formulating some alternative forms (section 3). Then consideration is given to the relationship and current relevance of OG dynamics to more general flow settings (section 4).

2. Preliminaries

a. Four formulations of the OG omega equation

The quasigeostrophic omega equation prescribes an ageostrophic vertical velocity field that is inferable from the geostrophic flow. To focus on the essence of this equation, consider the form it takes for a Boussinesq fluid in adiabatic flow above an f plane and possessing a uniform background mean-state stratification N. In this limit, and with height taken as the vertical coordinate, the equation for the vertical velocity w reduces to the form

\[ N^2 \nabla^2 w + f_0^2 \sigma^2 w/\partial z^2 = F, \]

where \( f_0 \) is the Coriolis parameter. For completeness and later use the full quasigeostrophic set of equations is provided in appendix A. From Eq. (1) it follows that in this limit-form \( w \) satisfies a second-order elliptic equation with constant coefficients, and as detailed below the forcing term \( F \) is only a function of the geostrophic flow.

The expression for \( F \) can be cast in various forms that, although mathematically equivalent, are open to markedly different interpretation. The four most common extant formulations of the term are

\[ F = f_0^b[\nabla_G \cdot \nabla_H]k_G/\partial z - \nabla_H^2[\nabla_G \cdot \nabla_H]b, \] (2a)

\[ = f_0^b[2(\partial w_G/\partial z \cdot \nabla_H)k_G + \text{DEF}], \] (2b)

\[ = 2(\nabla_H \cdot \nabla_G) \text{ and } \] (2c)

\[ = f_0^b[(\nabla_G \cdot \nabla_H)q]/\partial z - [\nabla_H^2 + (f_0/N)^2 \partial^2/\partial z^2](\nabla_G \cdot \nabla_H)b. \] (2d)

Here the deformation (DEF) term in the refined Sutcliffe formulation takes the form (appendix B; see Wiin-Nielsen 1959)

\[ \text{DEF} = [D_1(\partial D_2/\partial z) - D_2(\partial D_1/\partial z)], \] (3a)

where \((D_1, D_2)\) are the deformation components of the geostrophic flow in Cartesian coordinates, or it can be recast as

\[ \text{DEF} = -D^2 \partial(2\lambda)/\partial z, \] (3b)
where \( D \) is the total deformation of the geostrophic flow and \( \lambda \) is the orientation of the dilatation axis.

Also the vector \( Q \) in Eq. (2c) can be written (Sanders and Hoskins 1990) such that

\[
Q = -|v_H|\theta^*|k \times \partial G/\partial s|, \quad (3c)
\]

with \( s \) denoting a local coordinate aligned along an isentrope such that in the Northern Hemisphere the cold air located to the left.

Now we indicate how these various formulations can be derived most directly from the full quasigeostrophic set (appendix A). The conventional formulation (Fjortoft 1955; Thompson 1961) can be derived from taking the \( z \) derivative of Eq. (A5) and \( \nabla^2_H \) of Eq. (A6) followed by using Eq. (A3) to yield the two-term expression of Eq. (2a) for \( F \). The two terms are related respectively to the vertical gradient of the vorticity advection and the horizontal Laplacian of the thermal advection.

The refined Sutcliffe formulation (Fjortoft 1955; Wiin-Nielsen 1959) also follows directly from Eqs. (A5) and (A6), but now noting from Eq. (A3) that

\[
D_G(f_0h\partial G/\partial z)/Dt = D_G(\nabla^2_H b)/Dt
\]

and yields the two-term expression of Eq. (2b) for \( F \).

The first term refers to the advection of vorticity \( \xi_G \) by the “thermal wind” (\( \partial v_G/\partial z \)) and it is the erstwhile traditional Sutcliffe development term (Sutcliffe 1947), and the second term—the so-called deformation term—is as defined earlier [Eqs. (3a) and (3b)].

The \( Q \)-vector formulation (Hoskins et al. 1978) is obtained most directly by using the thermal wind relationship, \( f_0\partial v_G/\partial z = k \times \nabla_H b \), and noting following Eliassen (1962) that

\[
D_G(f_0h\partial G/\partial z)/Dt = D_G(k \times \nabla_H b)/Dt. \quad (4)
\]

Further manipulation yields a compact single-term representation of the forcing [Eq. (2c)] in terms of the vector \( \vec{Q} \) that is the product of the baroclinicity and the spatial variation in the strength and the direction of the geostrophic wind [Eq. (3c)].

Finally, the potential vorticity formulation (Hoskins et al. 2003) follows directly from recognizing that the quasigeostrophic potential vorticity \( q \) is linked to the buoyancy \( b \) such that

\[
f_0\partial q/\partial z = [\nabla^2_H + (f_0/N)^2 \partial^2/\partial z^2]b.
\]

This formulation was devised to link the QG omega equation with the so-called potential vorticity (PV) perspective (Hoskins et al. 1985). In the quasigeostrophic framework the PV perspective relates flow development to the evolution of the interior distribution of \( q \) and the surface distribution of \( b \). The link is emphasized by treating the two terms of the forcing in Eq. (2d) separately, and then obtaining the solution to the omega equation as the sum of three components, \( w = w_1 + w_2 + w_3 \), such that

\[
N^2\nabla^2_H w_1 + f_0^2 \partial^2 w_1/\partial z^2 = f_0 \partial[\langle v \rangle \cdot \nabla_H] q/\partial z
\]

subject to \( w_1 = 0 \) at \( z = 0 \),

\[
w_2 = -(1/N^2)[\langle v \rangle \cdot \nabla_H] b, \quad (5a)
\]

\[
N^2\nabla^2_H w_3 + f_0^2 \partial^2 w_3/\partial z^2 = 0
\]

subject to \( w_3 = -w_2 \) at \( z = 0 \). \quad (5c)

It follows that \( w_1 \) involves the interior \( q \) distribution and \( w_3 \) is a contribution arising from the advection of potential temperature \( b \) at the surface, so that \( w_1 \) and \( w_3 \) are associated directly with the two ingredient of the PV perspective. In contrast \( w_2 \) relates to advection on isentropic surfaces and acting alone would ensure a zero thermal tendency (i.e., \( \partial b/\partial t = 0 \)).

**b. \( F \) forcing and the vertical velocity**

Equation (1) indicates that \( w \) is a diagnostic variable whose spatial distribution at a stipulated time can be determined by inverting the equation. This requires the specification of the contemporaneous distribution of the “forcing function” in the interior and the application of the bottom \( w_{z=0} = 0 \) and upper \( w_{z=\infty} = 0 \) boundary conditions. Hence the formal solution of Eq. (1) prescribes the ageostrophic vertical velocity field \( w \) required by the prevailing geostrophic flow to maintain geostrophy, and the \( w \) field can be obtained by solving the equation numerically.

Equation (1) also provides an avenue for tackling the classical challenge of synoptic meteorology of inferring qualitatively the regions of major ascent and descent from inspection of standard synoptic charts, and it is also a tool for diagnosing the linkage between observed coherent features of the synoptic-scale flow and the accompanying vertical velocity field.

1) **SIMPLE QUALITATIVE AND QUANTITATIVE INFERENCE**

A customary qualitative aid to the inference procedure is to recognize that locally the following approximate relationship holds:

\[
[\nabla^2_H + (f_0/N)^2 \partial^2/\partial z^2]w \approx -w
\]
because $w$ varies about a zero mean. On this basis a local negative region of $\mathcal{F}$ forcing would be associated with a positive value of $w$, and likewise positive forcing connotes descent. An alternative is to invoke the geometric interpretation of a Laplacian as the difference of local positive extremum then the forcing would be associated with a positive value of $w$ to the vertical velocity field $w$. A quantitative prescription of the linkage of $\mathcal{F}$ forcing to the vertical velocity field $w$ can be obtained for the simple setting shown in Fig. 1a. It comprises an ellipsoidal region of uniform negative forcing $\mathcal{F} = -F$ located at a height $h$ above the surface and occupying the region $r < a$ with $r^2 = [x^2 + y^2 + (N/f_0)^2(z - h)^2]$. Note that the forcing region has a width-to-height aspect ratio of $\approx 100$ for typical extratropical values of $f_0$ and $N$.

The accompanying vertical velocity signature takes the form (cf. Clough et al. 1996)

$$w = (1/2)(F a^2/N^2)[1 - (1/3)(r_+/a)^2] - (1/3)(F a^2/N^2)(a/r_-) \quad \text{for } r \leq a, \quad (6a)$$

$$w = (1/3)(F a^2/N^2)[a/r_+ - a/r_-] \quad \text{for } r > a \quad (6b)$$

with $r_+ = [x^2 + y^2 + (N/f_0)^2(z - (\pm)h)^2]^{-1/2}$. This solution has been derived using the method of images with an opposite sign anomaly placed at $z = -h$ below the ground to ensure that $w = 0$ at $z = 0$. For consideration of the influence of a uniformly stratified stratosphere located above the ellipse, see Clough et al. (1996). To fulfill the continuity requirement of no net vertical mass flux across any given horizontal level the isolated $\mathcal{F}$ forcing needs to be supplemented by a similar ellipsoidal region of uniform positive forcing located at the same elevation in the far field but far enough away for its contribution to be negligible in the neighborhood of the negative forcing.

The form of the solution [Eqs. (6a) and (6b)] indicates that the ascent is a maximum within the core of the forcing, decreases like $1/r$ in the far field but with a more rapid decrease in the vicinity of the surface. This is illustrated in Fig. 1b, which depicts the pattern for a lens located in the lower troposphere at around 2.5 km with the nondimensional vertical velocity, $w^* = w/(Fa^2/N^2)$, shown as a function of $(x^*, z^*) = [x/a, (N/f_0)(z/a)]$. The vertical velocity retains a value of $w^* \approx (3/4)w_{\text{max}}^*$ within the forced region, but it reduces to $(1/10)w_{\text{max}}^*$ at a lateral distance of approximately $x^* = 3.0$.

To further calibrate the nature of the response Fig. 1c depicts two measures of the vertical velocity. The first measure is the dependence of $w^*$ at the center of the lens (i.e., $w_{x=-h}^*$) to the height, $h^* = (N/f_0)(h/a)$, of the lens.

![Figure 1](image-url)
Above $h^* = 4.0$, corresponding typically to a height of approximately 4 km and for a 2-km-deep forcing region, $w_{z \rightarrow h^*}^*$ is comparatively unmodified by the presence of the bottom boundary. For a forcing region located in the upper troposphere the maximum vertical velocity approximates to $w_{\text{max}}' \approx (1/2)(Fa^2/N^2)$.

The second measure depicted in Fig. 1c is the dependence of the vertical velocity at a fixed location in the lower troposphere ($\approx 1.5$ km) directly beneath the forcing for different elevations of the lens. The dotted line in Fig. 1c demonstrates that low-level ascent is favored by low-level forcing whereas the influence of mid- to upper-tropospheric forcing is substantially less effective. This result is germane to consideration of surface cyclogenesis because ascent in the lower troposphere connotes an increase in the low-level vorticity [see Eq. (A5)].

To relate the amplitude of the response to a prescribed forcing consider a 2000-km-wide and troposphere-spanning forcing region characterized by a thermal wind of 5 m s$^{-1}$ km$^{-1}$ and a vorticity gradient along the baroclinic zone of $2 \times 10^{-5}$ s$^{-1}$ (2000 km)$^{-1}$. For the first term in the refined Sutcliffe formulation [Eq. (2b)] this equates to a value of less than 0.05 m s$^{-1}$.

2) MORE REFINED CONFIGURATIONS

The contribution of a localized region of (say) negative $F$ forcing to the net forcing could be countered or indeed negated by a region of strong positive forcing located in the near vicinity. This is noteworthy because the atmosphere’s pattern of forcing often exhibits dipole-like structures.

Schematics of two such common configurations are shown in the left panels of Fig. 2. They comprise a vertically aligned forcing pattern astride the tropopause at $h^* = 9$, and a horizontally aligned pattern in the lower troposphere at $h^* = 2.5$. The main panel of Fig. 2 shows the pattern of $w^*$ at fixed nondimensional height of $z^* = 1.5$ for these two configurations.

The upper-level dipole reduces $w^*$ at $z = 1.5$ beneath the dipole by a factor of $\approx 2$ compared to the response for a single-signed region located aloft at $h^* = 9$ (note the different strengths assigned to the two configurations and compare with Fig. 1c). Likewise the horizontally aligned low-level dipole reduces $w^*$ at $z = 1.5$ beneath one of the forced region by some 30% compared to an isolated single forcing region. It follows that the characteristic dipole structure of synoptic-scale

Fig. 2. (left) Schematics of two dipole configurations characteristic of extratropical forcing and corresponding respectively (top) to a vertically aligned forcing pattern astride the tropopause at, say, $(x^*, y^*, z^*) = (0, 0, 9.0)$, and (bottom) to a horizontally aligned pattern in the lower troposphere centered at $(x^*, y^*, z^*) = (0, 0, 2.5)$. Each dipole element is assumed to be of uniform ellipsoidal forcing with strengths of $F = \pm 20F$ for the tropopause level and $\pm F$ for the lower troposphere configuration. (right) The resulting $w^*$ field at $z^* = 1.5$ in the $(x^*, z^*)$ plane with the dash-dotted and the dotted profiles pertaining respectively to the vertical and horizontally aligned dipoles, respectively.
forcing can significantly modulate the amplitude and structure of the low-level ascent, and that the response of the low-level dipole is \( \approx 20 \) times stronger per unit of forcing than that of the upper-level configuration.

In the foregoing we considered only ellipsoidal forcing elements with a large width-to-height aspect ratio. However, forcing regions with a smaller aspect ratio can be conceived as a vertical column of overlapping ellipsoidal elements, and the elements can be combined to represent richer forcing patterns.

3. Reappraisal and variants of the \( F \) formulation

a. The conventional formulation

Interpretation of the terms in the conventional formulation for \( F \) [see Eq. (2a)] is beset by issues of mathematical uniqueness, physical meaningfulness, and practical utility (Trenberth 1978; Hoskins et al. 1978; Durran and Snellman 1987; Walters 2001). Here we briefly underscore these shortcomings.

First, the expression is not Galilean-invariant in that it can be modified to the form

\[
F = f_0 \delta \left[ \left( (\mathbf{v}_G - c) \cdot \mathbf{V}_H \right) \xi_G \right] \partial z - \mathbf{V}_H \left( (\mathbf{v}_G - c) \cdot \mathbf{V}_H b \right)
\]

by the addition of an uniform translational velocity \( c \). This leaves the net forcing unchanged while modulating the relative amplitude of the two terms. For a given synoptic setting assigning a value for \( c \) corresponding to an estimate of the translational velocity of a prevailing synoptic system could shed light on the system-relative dynamics.

Second, the two terms in this formulation contain compensating contributions of the form

\[
\pm f_0 \left[ (\mathbf{v}_G \cdot \mathbf{V}_H) \partial \xi_G / \partial z \right],
\]

which can often be comparable in magnitude to that of the individual terms. This feature detracts from the physical meaningfulness of the latter terms.

Third, qualitative inference of the two forcing terms is cumbersome. For example, it is necessary to assess the vorticity advection, \( (\mathbf{v}_G \cdot \mathbf{V}_H) \xi_G \), at adjacent elevations, and moreover a typical synoptic flow setting with a localized maximum of vorticity advection at the tropopause equates to a zero contribution to the forcing at that level with an accompanying dipole structure above and below, and this in turn corresponds to a reduced effect in the far field (section 2).

By way of a counter to the foregoing shortcomings it is noteworthy that the two terms of this formulation appear to be linked to the two-category classification of surface cyclogenesis into frontal wave development accompanied by either significant low-level thermal advection or strong upper-level vorticity advection. However, it is noteworthy that it is the vertical derivative of the vorticity advection that appears in the QG forcing.

b. The refined Sutcliffe formulation

This formulation [Eqs. (2b), (3a), and (3b)] is Galilean invariant, and moreover from a purely kinematic standpoint is free from cancellation effects. The latter deduction follows because geostrophic flow at any level comprises the two independent components of vorticity and deformation and the formulation’s two terms relate separately to these components.

1) The Sutcliffe term

The Sutcliffe forcing term, \( 2f_0 [(\partial \mathbf{v}_G / \partial z) \cdot \mathbf{V}_H] \xi_G \), relates to the vorticity gradient along the thermal wind, and inspection of a single chart portraying both the thermal and the pressure patterns generally provides a ready indication of the regions of strong forcing. This is illustrated in Figs. 3a and 3b for the two archetypical configurations of the so-called thermal steering and jet left-exit settings. For both configurations negative forcing connoting ascent will be large along the baroclinic zone where it traverses away from the region of enhanced vorticity, and is consistent with respectively the displacement of the low along the baroclinic zone and of cyclogenesis in the left exit segment of the jet. Insight into the accompanying dynamics follows from considering the poleward slope of the isentropic surfaces within the baroclinic zone. The inferred patterns of ascent and descent are consistent with the flow associated with ambient vorticity moving on the isentropic surfaces.

For the synoptic setting of an incipient frontal cyclone (see Fig. 6a) the Sutcliffe term is indicative of ascent along the warm front ahead of the cyclone and descent on its rearward southwestern flank.

2) The deformation term

The form of the deformation term, \( \text{DEF} = -D^2 \partial (2\lambda) / \partial z \), appears to require assessing the difference in the angle of orientation of the deformation field’s dilatation axis at adjacent levels in the vertical, and the term is formally of the same order as the Sutcliffe term itself. In practice it tends to be small in the midtroposphere where the thermal gradient is often weak and almost coaligned with the geostrophic flow (Wiin-Nielsen 1959), but it can be significant in amplitude in frontal and jet-stream regions where it usually exhibits a rich localized spatial structure (see, e.g., Martin 1999).

An alternative form can be derived that enables a direct qualitative inference of the in situ DEF forcing.
by inspecting a single surface (see appendix B). The key to the derivation is to recognize that the thermal wind, $
abla T = \partial T / \partial z$, is such that $\nabla T = (1/f_0)(k \times \mathbf{V}_H)$, and hence the isentropes on a given surface can be viewed as the streamfunction of the $T$ field.

It then follows that the DEF term can be recast in the form (see appendix B)

$$\text{DEF} = -[(D)(S)] \sin(2\gamma),$$

where, as before, $(D, \lambda)$ correspond to the total deformation and the orientation of the dilatation axes of the geostrophic flow, and now $(S, \epsilon)$ correspond to the total deformation and dilatation angle of the “thermal wind” field, while $\gamma = (\epsilon - \lambda)$ defines the relative orientation of the two dilatation axes. Using Eq. (7) a synoptic chart depicting both the potential temperature field and the pressure field allows a qualitative inference to be made of the locations of strong DEF by assessing the regions of major $D$ and $S$ plus the local relative orientation of their dilatation axes.

This is illustrated in Fig. 4. First, the upper panels (Figs. 4a–d) depict prototypical deformation fields and the orientation of the attendant dilatation axes for respectively: a pure deformation field, a uniform horizontal shear and a jetlike profile, and a Rankine vortex comprising an inner core in solid body rotation and an outer domain in irrotational flow. The lower panels (Figs. 4e–h) show the DEF forcing resulting from various simple combinations of these deformation signatures for the flow and the thermal wind.

Figure 4e is for the canonical setting of pure deformation field acting on a baroclinic zone that is aligned along the dilation axis, and for this configuration the Sutcliffe term is zero. The deformation field $D$ is such that $\lambda = 0$ (cf. Fig. 4a) whereas the deformation of the thermal wind $S$ is strongest astride the zone of enhanced baroclinicity with $\epsilon = +45^\circ$ on the warm side and $\epsilon = -45^\circ$ on the cold side (cf. Fig. 4c). It follows that on the warm side of the zone $\gamma \approx 45^\circ$, $\sin(2\gamma)$ is maximized, and the DEF forcing favors ascent. Contrariwise on the cold side the forcing favors descent.

The setting in Fig. 4f of a baroclinic zone traversing across a Rankine vortex configuration serves to emphasize that, if the vorticity and its gradient are weak in the outer regions of a cyclone, the contribution of the Sutcliffe term is less effective, and that the prevailing deformation and the orientation of the dilatation axes favors descent to the east and ascent to the west of the cyclone.

The settings in Figs. 4g and 4h relate to the character of cold and warm fronts in different ambient flow environments. For these panels the flows comprise respectively of uniform negative $(\lambda = +45^\circ)$ and positive $(\lambda = -45^\circ)$ horizontal shear, while the thermal pattern corresponds in both settings to a highly idealized cold and warm frontal pattern. Again the Sutcliffe term itself is zero. For Fig. 4g the relative orientation of the dilatation axes favors a direct circulation at the cold front with warm air ascent and cold air descent, but with no DEF forcing along the warm front. Contrariwise the presence of positive shear promotes a direct circulation along the warm front.

This dichotomous behavior is pertinent to consideration of the relative behavior and movement of fronts within a single cyclonic system [see the discussion in Schultz and Vaughan (2011)] and is in harmony with the
distinctive cyclonic and anticyclonic shear (Davies et al. 1991) and LC1 and LC2 (Thorncroft et al. 1993) classifications of baroclinic development. (Likewise, inspection of Fig. 6b hints that, for a frontal-wave cyclone setting, the DEF term favors frontogenesis along a restricted segment of the cold front.)

c. The $Q$-vector formulation

The $Q$-vector formulation of the forcing [Eq. (2c)] provides a compact single term expression, and avoids both non-Galilean and cancellation effects. It offers a tractable approach to inferring the regions favorable for ascent; although it has some intrinsic limitations, it also sheds light on the subtle nature of quasigeostrophic flow. We comment in turn upon these aspects.

1) Inference Procedure

In this formulation ascent will be favored by a distinct maximum of $Q$ convergence (i.e., negative $\nabla_H \cdot Q$), and in the vicinity of such a region the $Q$ vector will tend to point toward the region of convergence and away from a divergent region. It follows from Eq. (3c) that on a synoptic chart a region of large $Q$ will be characterized by a combination of strong baroclinicity and a marked change in the strength and/or direction of $\mathbf{v}_G$ along the isentropes, and furthermore the resulting $Q$ vector will

Fig. 4. Illustrations of the $\text{DEF} = -[(D)(S)] \sin(2\gamma)$ forcing term of the refined Sutcliffe formulation. (a)–(d) The archetypical geostrophic flow or thermal wind fields corresponding, respectively, to pure deformation, an uniform shear, a jetlike profile, and a Rankine vortex comprising a region of solid rotation surrounded by irrotational flow, and the double-arrowed dotted line signifies the orientation of the dilatation axis for each setting. Also shown are various combinations of these configurations: (e) the canonical deformation-induced frontogenesis setting of a baroclinic zone (colored contours) embedded in a pure deformation flow field (light black contours), (f) a baroclinic zone embedded within a Rankine vortex configuration, and (g),(h) an arched baroclinic zone embedded within, respectively, a uniform negative and positive horizontal shear. Again the double-arrowed dotted lines signify the orientation of the dilatation axes (black for the geostrophic flow, and red and blue for, respectively, the warm and cold flanks of baroclinic features). The positive and negative symbols indicate the regions of DEF forcing favorable respectively for ascent and descent.
be aligned 90° to the right (in the Northern Hemisphere) of the vector wind change. This deductive chain is illustrated in Figs. 5a–c for three archetypical settings, and also Fig. 6c for the setting of an incipient frontal-wave cyclone.

2) SOME LIMITATIONS

In general the Q-vector formulation has met with widespread acceptance, but it has some potential shortcomings. First, the need to estimate the divergence term, $H/Q$, is demanding if the flow has a rich subsynoptic spatial structure.

Second, the two-dimensional Q vector decomposes into independent irrotational and rotational components (see, e.g., Keyser 1999), that is,

$$ Q = Q_{\text{irrot}} + Q_{\text{rot}}. $$

In effect the quasigeostrophic flow is accompanied by both an ageostrophic irrotational flow related to $V_H \cdot Q$ and diagnosed by the QG omega equation, and an ageostrophic rotational flow related to $V A \cdot Q$. Subject to the caveat noted in appendix A, the latter is diagnosed via Eq. (A8), that is,

$$ f_0 \zeta_{AG} = \frac{1}{2}(\zeta_G^2 - D^2) $$

and the relationship (Hoskins and Pedder 1980; Keyser 1999)

$$ 4(V A \cdot Q) = f_0 \theta(\zeta_G^2 - D^2)/\partial z. $$

A key point here is that it is only the irrotational component of Q that contributes to the $F$ forcing.

Thus qualitative inference of the forcing by locating a dominant Q vector is conditional upon the rotational component ($Q_{\text{rot}}$) not swamping or masking the signal of the irrotational component ($Q_{\text{irrot}}$). Furthermore, recourse to evaluating $V H \cdot Q$ more precisely as the net flux of Q from a given area might also be challenging if $|Q_{\text{irrot}}| \sim |Q_{\text{rot}}|$. Compare the computation of the horizontal divergence ($V H \cdot V$) of synoptic-scale flow when it is small compared to the vertical component of the vorticity.

3) QUASIGEOSTROPHIC DYNAMICS AND THE Q VECTOR

A starting point for the Q-vector formulation is two equations derived directly from Eqs. (A4) and (A6), namely

FIG. 5. Illustrations of the $Q$ vector for three archetypical flow settings: (a),(b) the thermal steering and jet-exit configurations respectively (cf. Fig. 3), and (c) the canonical deformation-induced frontogenesis setting. Solid black arrows are indicators of the flow’s direction and strength, the dashed line refers to the change in the flow along the baroclinic zone (i.e., $\partial V_G/\partial s$), and the bold open arrow denotes the $Q$ vector (obtained by rotating the vector of the wind change by 90° to the right). (d) The R vector (shaded arrow) for the frontogenesis setting [cf. (c)] and the implied ageostrophic circulation in a vertical plane across the front is as indicated.
\[
D_G(f_0\partial v_G/\partial z)/Dt = -f_0 R - f_0^2 [k \times (\partial v_{AG}/\partial z)], \tag{9a}
\]
\[
D_G(k \times \nabla_H b)/Dt = +f_0 R - N^2 (k \times \nabla_H w), \tag{9b}
\]

where \( R = (\partial v_G/\partial z \cdot \nabla_H) v_G \), so that it relates to the advection of the geostrophic flow by the thermal wind. [As noted in appendix A, an additional term would appear in Eq. (9a) unless the primacy of the pressure field is asserted over the wind field.]

The subtle nature of OG dynamics [see Hoskins et al. (1978)] is rendered transparent on examining the role of \( R \) in Eqs. (9a) and (9b). Its contribution to the evolution of the vertical shear \( \partial v_G/\partial z \) is equal and opposite to its influence upon the baroclinicity term \( k \times \nabla_H b \). Thus the intrinsically geostrophic effect of \( R \) acts to destroy the thermal wind balance, and thereby geostrophy itself. Therefore the remaining ageostrophic terms on the right-hand side of Eq. (9) are required to concomitantly restore that balance. In short, the maintenance of the primary geostrophic flow necessitates the existence of a secondary ageostrophic flow.

Here we explore further features of the \( Q \)-vector formulation but now do so using the \( R \) vector. Note, however, that these two vectors are formally closely linked by the relationship

\[
Q = -f_0 (k \times R),
\]

and this linkage is further emphasized by recognizing in harmony with the physical interpretation of \( R \) that, from Eq. (3c), \( Q \propto \{\nabla_H \theta^* \partial v_G/\partial s\} \).

Aspects of the role and significance of the \( R \) vector and other closely related vectors have been noted by Davies-Jones (1991) in his generalization of the \( Q \)-vector approach to flow states beyond strict quasigeostrophy, and by Steinacker (1992) in his application of the \( Q \)-vector approach to study frontal dynamics. Also, the \( R \) vector equates in essence to the horizontal component of the \( C \) vectors derived by Xu and Keyser (Xu, 1992; Xu and Keyser, 1993) in their extension of the \( Q \) vector to three dimensions, and likewise to the curl of the horizontal component of the \( P \) vector formulated
flow is evident in embryonic form in Hoskins et al. (1978), and it is rendered explicit in Steinacker (1992) and Xu (1992).

In applying Eq. (10b) directly to observed data again note that the estimate of a “divergence” and the other a “curl.”

The key feature of the Q- or R-vector formulation is that they condense the physics to consideration of the spatial structure of a single, easily diagnosed vector, and moreover the dynamical attributes of the R vector forms a platform for some of the considerations in section 4.

d. The potential vorticity formulation

The PV formulation [Eq. (2d)] is reminiscent to that of the conventional formulation (Eq. 2a), and it suffers from the same shortcoming of non-Galilean invariance and compensating contributions (CC). The latter now take the form

$$CC = \pm f_0 \left[ (\mathbf{v}_C \cdot \mathbf{V}_H) \left( \partial q / \partial z \right) \right].$$  

(12)

The formulation can nevertheless provide some insight into the dynamics of a single system by considering (Hoskins et al. 2003) a system-relative framework and replacing Eq. (2d) by the set of Eqs. (5). Another significant mitigating factor is that the potentially debilitating influence of the compensating contributions is nullified over much of the flow domain. This arises because the atmosphere’s major interior “q features,” and hence the regions with potentially strong contributions to the compensating contributions [Eq. (12)], are confined predominantly to highly localized regions near the jet stream and to the immediate vicinity of air parcels previously subject to strong diabatic heating.

However, from a PV perspective the flow and thermal patterns can be reconstructed from the interior q and surface b distributions, and the latter distributions carry a far-field signal. Thus an alternative to the foregoing PV formulation is to evaluate the net forcing at a given level as the sum of contributions arising from the interior q distribution alone, the surface b distribution alone, and their interaction. Such a procedure could be extended to

$$2f_0 \mathbf{R} = -N^2 \{ \mathbf{k} \times [(f_0/N)^2 (\partial \mathbf{v}_{AG} / \partial z) - (\mathbf{V}_H w)] \}$$

$$= -N^2 \{ \mathbf{k} \times [(\partial \mathbf{v}_{AG} / \partial z^*) - (\mathbf{V}_H w)] \}$$

with $Z^* = (N/f_0)^2 z$

$$\eta_1 = \mathbf{k} \times [(\partial \mathbf{v}_{AG} / \partial z^*) - (\mathbf{V}_H w)].$$  

(10b)

Hence, with a rescaled vertical coordinate $(Z^*)$, R is proportional to the negative of the horizontal vorticity component $(\eta_1)$ of the ageostrophic flow $(v_{AG}, w)$. The linkage of the Q and R vectors to the vorticity of the ageostrophic flow is evident in embryonic form in Hoskins et al. (1978), and is rendered explicit in Steinacker (1992) and Xu (1992).

In effect the amplitude and direction of the R vector is, subject to the foregoing caveat, an indicator of the strength and sign of the ageostrophic circulation, such that the ageostrophic circulation is anticlockwise about the R vector. This is illustrated in Figs. 5d and 6d for some previously considered settings. For example, for the canonical setting of deformation-induced frontogenesis (Fig. 5d) the R vector points along the thermal wind and implies a counterclockwise vorticity/circulation in the across-front plane with warm air ascent and cold air descent.

4) $F$ FORCING AND THE R VECTOR

The formulation of the $F$ forcing in terms of R follows directly from applying the operator [k $\cdot$ V] to Eq. (10a) to yield

$$N^2 \nabla^2 w + f_0^2 \partial^2 w / \partial z^2 = 2f_0 [\mathbf{k} \cdot (\mathbf{V} \times \mathbf{R})].$$  

(11a)

The utility of this formulation becomes more apparent on using Stokes’s equation to note that

$$\int_S [\mathbf{k} \cdot (\mathbf{V} \times \mathbf{R})] \, ds = \int_C \mathbf{R} \cdot \, dc,$$  

(11b)

where S represents a specified horizontal surface area enclosed by the closed contour C.

Thus an assessment of the net forcing over a selected horizontal area S can be obtained by examining the component of the R vector aligned along the contour C. A simple procedure to identify a region of strong forcing is to locate the zone of strong baroclinicity and then select a contour C in its neighborhood such that the contribution $\mathbf{R} \cdot \, dc$ to the line integral is only significant over a part of the contour. The challenge to infer the dominant Q or R vector is equivalent with one requiring the estimate of a “divergence” and the other a “curl.”

The key feature of the Q- or R-vector formulation is that they condense the physics to consideration of the spatial structure of a single, easily diagnosed vector, and moreover the dynamical attributes of the R vector forms a platform for some of the considerations in section 4.

A simple procedure to identify a region of strong forcing is to locate the zone of strong baroclinicity and then select a contour C in its neighborhood such that the contribution $\mathbf{R} \cdot \, dc$ to the line integral is only significant over a part of the contour. The challenge to infer the dominant Q or R vector is equivalent with one requiring the estimate of a “divergence” and the other a “curl.”
examine the contribution to the forcing of spatially coherent interior $q$ and surface $b$ features that are deemed to be of seminal dynamical significance.

For example, consider the incipient frontal-wave cyclone setting depicted in Fig. 6. The trough and attendant flow resemble the signature at this level of an isolated upper-level $q$ anomaly while the baroclinic zone is a manifestation of a surface front. The interaction of these two signals would then influence the forcing at the displayed level.

The far-field contributions of $q$ and $b$ features also carry ramifications for the interpretative value of cyclogenesis classification schemes based upon the relative magnitude of the upper- and lower-level $F$ forcing. Such schemes will not capture the interlevel forcing effects noted above.

4. Relevance

It was noted in the introduction that the contemporary relevance of the QG omega equation can be called into question on several accounts. Current research and NWP models deliver the vertical velocity field as a model output variable. High-resolution numerical models simulate subsynoptic flow features whose dynamics, although often balanced, is beyond the strict validity of quasigeostrophy. Moreover the dynamics of the QG omega equation appears unrelated to many of today’s pressing dynamical and NWP challenges.

In the following we address and at seek to at least partially alleviate some of the foregoing concerns. To this end consider sequentially pertinent aspects of the dynamics intrinsic to quasigeostrophic flow, diabatically induced flow, and balanced flow.

a. Further comments on quasigeostrophy

There are two fundamental features of quasigeostrophy that bear directly upon its relevance to subsynoptic flow.

1) CASCADE TO FINER-SCALED FLOW FEATURES

Quasigeostrophic dynamics can itself induce finer-scaled flow systems and thereby contribute to the development of, and possibly account for a part of, the panoply of the atmosphere’s subsynoptic flow features.

This can be demonstrated by introducing the concept of “thermal wind lines” (TWLs), defined as being everywhere tangential to the local orientation of the thermal wind, $T_G = \partial \mathbf{v}_G / \partial z$. Hence the TWLs at any given level are aligned with the isentropes, and their strength proportional to the local baroclinicity (inversely proportional to the isentrope spacing), and this linkage is only valid in the quasigeostrophic limit.

An evolution equation for TWLs follows from introducing the thermal wind relationship Eq. (A2) into Eq. (9b) to yield

$$D_G(T_G)/Dt = (T_G \cdot \mathbf{v}_H) \mathbf{v}_G - (N^2/f_0)\{\mathbf{k} \times \mathbf{v}_H(w)\}.$$

(13a)

Note that $\mathbf{R} = (T_G \cdot \mathbf{v}_H) \mathbf{v}_G$, and hence $\mathbf{R}$ can distort the TWLs either modifying their strength by along-flow stretching and/or reorienting them by deformation (see, e.g., Keyser et al. 1988). The remaining ageostrophic term modifies $T_G$ by the tilting effect of the vertical velocity acting in the presence of a background vertical stratification.

At the surface $w = 0$, and the equation reduces to

$$D_G(T_G)/Dt = + (T_G \cdot \mathbf{v}_H) \mathbf{v}_G.$$

(13b)

Hence at the surface the TWLs behave like “pseudo-material lines” moving with the two-dimensional quasi-geostrophic flow in the absence of frictional and diabatic effects. An apposite analogy is with the reorientation and deformation of vortex lines in a homogeneous three-dimensional flow (von Helmholtz 1858). From a similar standpoint and in the context of examining balanced flow DJ91 considered the evolution equation for the vertical shear of the horizontal wind.

Here the dynamics of TWLs is intrinsic to so-called two-dimensional surface-geostrophic dynamics (see Held et al. 1995; Smith and Bernard 2013) and can produce a cascade to finer-scaled flow. An illustration of this physical scale cascade is provided in Fig. 7. The initial state, in harmony with the so-called surface geostrophy, comprises a surface warm band of $b$ that decays with height accompanied by zero interior quasigeostrophic potential vorticity $q$. This configuration (see Fig. 7a) was studied by Schär and Davies (1990) and is unstable. They examined the growth rate of linear normal mode perturbations and also numerically simulated its initial nonlinear evolution. In the latter phase the pattern evolves to form a train of localized warm-cored vortices that are linked together by scale-reduced warm bands (see Figs. 7b,c). These new bands are themselves unstable and breakdown (see Fig. 7d, and Fehlmann 1997) to form finer-scaled warm-cored vortices.

The salient inference here is that quasigeostrophic dynamics can contribute to the development of subsynoptic flow features, and intrinsic to this scale reduction is the role of the $\mathbf{R}$ vector, which is itself central to the forcing of the QG omega equation.

2) AGEOSTROPHY AND SCALE CONTRACTION

Ageostrophy plays a key role in the development of some subsynoptic-scale flow phenomena such as fronts.
For example, ageostrophic effects substantially accelerate the lateral contraction and refine the vertical structure of a baroclinic zone embedded within a synoptic-scale deformation field (see, e.g., Hoskins 1971; Davies and Muller 1988; Reeder and Keyser 1988). However, this rapid frontal development is not captured adequately by the QG prediction Eq. (A7). Notwithstanding diagnostic application of the QG omega equation at any given time during the scale-contraction process will nevertheless, given pertinent information on the prevailing flow configuration, deliver an estimate of the key ageostrophic flow component. Indeed for the particular setting considered by Davies and Muller (1988) the surface ageostrophic flow component of the QG omega equation corresponds to that obtained with the semi-geostrophic equations.

In effect, application of the QG omega equation at a given instant can deliver useful information about and insight into the flow that might strictly be beyond the essence of quasigeostrophy itself.

b. Cloud-diabatic effects

The QG omega equation can serve as a test bed for examining issues related to balanced flow. Consideration of the quasigeostrophic response to cloud diabatic effects will serve to highlight a significant difference between balanced and unbalanced flows.

Diabatic effects are customarily incorporated into the QG omega equation by including an additional forcing term so that Eq. (1) becomes

$$N^2\nabla_H^2 w + f_0^2 \partial z^2 = \mathcal{F} + \nabla_H^2 \mathcal{H},$$

(14)

where $\mathcal{H}$ is a measure of the diabatic heating rate with a unit of $\mathcal{H}$ corresponding to approximately $3.0 \times 10^{-2}$ degrees heating per day. The equation remains

Fig. 7. The quasigeostrophic development of a dimensional surface warm band with zero interior potential vorticity. (a) This unstable configuration is perturbed by its most unstable normal mode. The subsequent panels capture (b), (c) the evolution of a train of localized warm-cored vortices linked together by scale-reduced warm bands, which themselves develop (d) finer-scaled warm-cored vortices.
self-contained provided \( \mathcal{H} \) can be specified externally or in terms of geostrophic flow variables.

The \( \nabla^2 \mathcal{H} \) form of the diabatic forcing implies that the \( w \) field might not to be necessarily positive within the diabatic region itself (see Krishnamurti 1968; Hoskins et al. 2003). Here we explore and quantify this effect by considering the vertical velocity field associated with a localized positive lenticular ellipsoidal distribution of \( \mathcal{H} \) of the form

\[
\mathcal{H} = H \{1 - [(r/a)^2 + (z/d)^2]\}^2
\]

within the ellipsoid \([r/a^2 + (z/d)^2] = 1\), \hspace{1cm} (15a)

\[\mathcal{H} = 0 \quad \text{elsewhere}. \] \hspace{1cm} (15b)

Here \( r \) is the radial distance in cylindrical coordinates, and as before the horizontal-to-vertical aspect ratio of this forcing element is defined to be such that \( a = (N/f_0)d \).

It follows from Eqs. (14) that the actual diabatic forcing pattern is given by

\[
\nabla^2 \mathcal{H} = -(8H/a^2)[1 - 2(r/a)^2 - (z/d)^2]
\]

and the region of positive forcing is confined to within a narrower ellipsoid,

\[2(r/a)^2 + (z^2/d^2) = 1,\]

that is encased within an even narrower sheath of negative forcing (cf. Hoskins et al. 2003).

Setting \( F = 0 \) in Eq. (14), adopting a rescaled vertical coordinate, \( z^* = (N/f_0)z \), and writing Eq. (14) in spherical polar coordinates \((R, \phi, \chi)\), we have

\[
\nabla^2 w = -H^* [1 - 2(R/a)^2 + (R/a)^2 \cos^2 \phi],
\]

(17)

where \( H^* = 8H/(Na)^2 \).

Recalling that the Laplacian operator with longitudinal symmetry takes the form

\[
\nabla^2 w = (1/R^2) \{ \partial / \partial R [ R^2 \partial w / \partial R ] \\
+ (1/\sin \phi) \partial / \partial \phi [\sin \phi \partial w / \partial \phi] \}
\]

and solving Eq. (17) subject to continuity conditions at \( R = a \) yields the following expression for the vertical velocity:

\[w/H^* = (2/3)[1 - (R/a)^2]^2 - (4/105)(R/a)^2[7 - 5(R/a)^2](1 - 3 \cos^2 \phi) \quad \text{for} \quad R \leq a, \] \hspace{1cm} (18a)

\[w/H^* = -(8/105)(a/R)^3(1 - 3 \cos^2 \phi) \quad \text{for} \quad R > a. \] \hspace{1cm} (18b)

This pattern of vertical velocity corresponds to ascent in a band narrower than the original diabatic region with descent elsewhere (Fig. 8). The vertical velocity attains its maximum value \( w_{\text{max}} \) at the core of the diabatic region (i.e., at \( R = 0 \)) with a dimensional value of \( 2.0 \times 10^{-2} \) m s\(^{-1}\) for \( N = 10^{-2} \) s\(^{-1}\) and a 10-K heating rate per day. On the perimeter of the heating domain the vertical velocity has decreased such that \( w = (1/4)w_{\text{max}} \) at the base \((r, z) = [0, -(f/N)a]\), and has reversed with \( w \approx -1/(8)w_{\text{max}} \) on the lateral edge at \((r, z) = [a, 0]\).

Note that the pattern for a cloud-diabatic region with a smaller width-to-height ratio would be equivalent to a vertical superposition of elements, and this would strength the entire central shaft of ascending air.

In the present context a key feature is that, in the quasigeostrophic limit, the vertical velocity accompanying diabatic heating is essentially in phase with the heating.
This has dichotomous ramifications. For quasigeostrophic flow the phase match can be viewed as a mix of a positive feedback between the diabatic forcing and the response, or a constraint upon the amplitude of the diabatic heating imposed by the moisture flux at cloud base attributable to the ascent forced by an ambient flow. An assessment of the relative efficacy of these two effects could be obtained by applying the QG omega equation for a specified flow along with an estimate of the contemporaneous diabatic heating.

Another ramification is that for unbalanced flow a phase match does not necessarily prevail between heating and ascent. In the linear limit the vertical velocity response to stationary (Lüthi et al. 1989), steady (Davies 1987), or pulsed (Lin and Smith 1986) imposed diabatic heating is in general out of phase with the heating. Likewise, the nonlinear adjustment following a local instantaneous injection of heat in a resting ambient atmosphere yields a net descent in the inner core of the initially warmed region with ascent above and below (Fanelli and Bannon 2005).

This contrast in the heating-ascent relationship is relevant in an NWP context. Longstanding challenges in NWP are to deliver reliable quantitative forecasting of precipitation, and to develop robust and effective techniques to assimilate/adjust model fields to account for misrepresented or unrepresented cloud diabatic heating. The latter requires the amalgam of available information acquired from diverse data sources on the intensity, movement, vertical profile, and location of cloud diabatic activity. One inference from the foregoing phase mismatch is that an inadequate assimilation technique might produce descent rather than the desired ascent in a region experiencing cloud-diabatic effects.

More generally these ramifications highlight that a higher form of balance must prevail between the dynamics and thermodynamics in an active cloud-diabatic region. To the extent that this form of hybrid balance also pertains for diabatic processes at the grid/subgrid scale, then it should also be taken into account in formulating stochastic representation schemes.

c. Balanced flow and the R-like vectors

Some large-amplitude synoptic- and subsynoptic-scale phenomena, although not strictly geostrophic, do exhibit a balanced character with the horizontal flow linked directly but nonlinearly to the prevailing pressure field. Such phenomena satisfy a state of balance beyond quasigeostrophy and a range of predictive equations intermediate between the QG and the primitive equations have been developed for theoretical [e.g., geostrophic momentum equations (Hoskins 1975), higher-order balance, and Hamiltonian balance (Delahaies and Hydon 2011)] and NWP (e.g., Bolin 1955) purposes.

Likewise generalizations of the quasigeostrophic omega equation have been formulated to diagnose the secondary ageostrophic flow component accompanying a primary balanced flow. These generalizations fall into three classes. One class relates explicitly to refined balanced systems such as semigeostrophic and geostrophic momentum-like systems (Keyser et al. 1989; Keyser 1999; Pedder and Thorpe 1999), and a second class relates to the balance formulations favored in NWP (Ziemianski and Thorpe 2003). In the third class are equations derived directly from the primitive equations by extending either the procedure followed to obtain the conventional formulation (see, e.g., Krishnamurti 1968; Tarbell et al. 1981; Hirschberg and Fritsch 1991; Pauley and Nieman 1992; Räisänen 1995) or the procedure for obtaining a generalized Q-vector formulation (e.g., DJ91; Viúdez et al. 1996).

This third class of equations contain time-tendency terms and are therefore pseudo, rather than strictly diagnostic, omega equations. Nevertheless comparison of $w$ fields derived neglecting some terms, including the time-tendency terms, with those delivered by the QG omega equation often show a resemblance beyond that expected by simple scaling arguments (see e.g., Pauley and Nieman 1992; Räisänen 1995). This success tallies with the experience of qualitative application of $Q/R$-vector approach to synoptic charts.

Here the objectives are to highlight the basis for this effectiveness over and above the factors already noted in section 4a. To this end mild generalizations of the quasigeostrophic R vector will be formulated, applied to in the customary manner to examining NWP model output, and to examine their potential value in assessing the balanced versus unbalanced state of the output. The need for such an indicator has previously been highlighted (e.g., Zhang et al. 2000).

1) R AND THE PSEUDO-OMEGA EQUATIONS

We again consider the Boussinesq form of the primitive equations on an $f$ plane, and follow the procedure of DJ91 and Viúdez et al. (1996) to obtain a pseudo-omega equation.

The NWP model output fields for the horizontal wind $v_f$ and the potential temperature $\theta$ are nominally independent, and neither need necessarily be representative of a balanced model state. However, these fields do provide independent measures for the vertical shear of the horizontal flow, namely a shear calculated directly, $T_f = \partial(v_f)/\partial z$, and a “virtual” shear based upon the potential temperature and defined by the relationship, $T_\theta = [k \times V_f b]/f_0$. The difference between these two measures,
\[ T_{\text{AG}} = T_F - T_\theta, \]  
(19)
defines and quantifies a departure of the model state away from the thermal wind balance.

Evolution equations for \( T_F \) and \( T_\theta \) can be derived that are the analog of the quasigeostrophic Eqs. (9a) and (9b):

\[
D[f_0 T_F]/Dt = -f_0 \mathcal{R}_F f_2^2 (k \times T_{\text{AG}}) - f_0 (\partial w/\partial z) T_F, \tag{20a}
\]

\[
D[f_0 T_\theta]/Dt = -f_0 \mathcal{R}_\theta - N^2 (k \times V_H w) + f_0 (\partial w/\partial z) T_\theta. \tag{20b}
\]

Here \( D/Dt \) is the Lagrangian derivative following the model’s three-dimensional flow field, and \( N \) is now the horizontal and vertically varying Brunt–Väisälä frequency. Also \( \mathcal{R}_F = [T_F \cdot V_H] V_F \) and \( \mathcal{R}_\theta = [T_\theta \cdot V_H] V_F \), and these two horizontally aligned “\( R \)-like” vectors are such that \( \mathcal{R}_F \) is a function of the model’s horizontal wind field (\( V_F \)), and \( \mathcal{R}_\theta \) a function of the model’s thermal wind (\( T_\theta \)) and horizontal wind field (\( V_F \)).

Subtracting Eq. (20b) from Eq. (20a) and using Eq. (19) yields an evolution equation for \( T_{\text{AG}} \):

\[
D[f_0 T_{\text{AG}}]/Dt = -f_0 (\mathcal{R}_F + \mathcal{R}_\theta) - N^2 \eta_2 - f_0 (\partial w/\partial z)(T_F + T_\theta). \tag{21}
\]

Here \( \eta_2 = k \times [(f_0/N)^2 (\partial V_{\text{AG}}/\partial z) - V_H w] \) is akin to the horizontal ageostrophic vorticity [cf. \( \eta_1 \) defined in Eq. (10)]. Equation (21) is equivalent to that derived in DJ91, albeit using a different notation, and in that study the first and the last terms on the right-hand side were taken together to define generalized \( R \) vectors.

Equation (21) indicates that the Lagrangian rate of change (amplitude and orientation) of a fluid parcel’s departure from thermal wind balance is determined by the net of three terms: the sum of the two \( R \)-like vectors involving only the horizontal velocity and thermal fields, a horizontal vorticity vector involving the vertical velocity field, and a third term involving both the horizontal flow variables and the vertical velocity. In principle each contribution could include strong geostrophic and/or ageostrophic contributions because at any given instant Eq. (21) encompasses both the “slow” evolution of a balanced flow component, and a possible rapid emission of inertia–gravity waves during a down-scale flow cascade plus the subsequent readjustment to a state of balance.

One manifestation of the slow evolution is captured by quasigeostrophy. In this limit Eq. (21) reduces to the diagnostic relationship Eq. (10a) because \( T_{\text{AG}} = 0 \), \( R_F = \mathcal{R}_\theta = R \), the \( \partial w/\partial z \) term is not present, and \( N^2 \) is prescribed as a function of height only. A further illuminating example is a circularly symmetric baroclinic vortex characterized by strong cyclostrophic balance so that \( T_{\text{AG}} \) is appreciable. For this system Eq. (21) reduces to \( D[f_0 T_{\text{AG}}]/Dt = 0 \) because \( R_F = \mathcal{R}_\theta = 0 \) and \( w = 0 \). In effect the prevailing magnitude of a flow’s departure from geostrophy, \( T_{\text{AG}} \), need not be an indicator of imbalance.

In DJ91 the slow evolution is assumed to be consistent with fluid parcels remaining in a state of thermal wind equilibrium—that is, \( D[T_{\text{AG}}]/Dt = 0 \)—and an omega equation derived by applying the \([k \cdot V_\Lambda]\) operator to Eq. (21). Solution of the latter equation, given the horizontal velocity field and the thermal field, delivers a vertical velocity field forced by the prescribed (possibly unbalanced) fields. The influence of \( D[T_{\text{AG}}]/Dt \) has been shown to be weak in some simple settings for which \( w = 0 \) (Davies-Jones 2009), and in that study the accompanying dynamical adjustment is characterized by inertial waves. The restriction to horizontal motion singles out inertial waves as a mechanism for adjustment, whereas in realized atmospheric flow the contribution of inertia buoyancy could be, and conceivably is, more significant.

Here the objective is to explore the value of a conventional \( R \) vector in examining the space–time output of NWP models. In particular emphasis is directed to the basis for its apparent effectiveness in diagnosing the major regions of ascent and to the insight it can shed on the balanced versus unbalanced state of the model state.

To this end consider two further formulations of Eq. (21),

\[
D_H[f_0 T_{\text{AG}}]/Dt + f_0 [T_{\text{AG}} \cdot V_H] V_F = -2f_0 \mathcal{R}_\theta - N^2 \eta_2 + \Pi \tag{22a}
\]

and

\[
D_H[f_0 T_{\text{AG}}]/Dt - f_0 [T_{\text{AG}} \cdot V_H] V_F = -2f_0 \mathcal{R}_F - N^2 \eta_2 + \Pi. \tag{22b}
\]

Here \( D_H/Dt \) refers to the Lagrangian derivative following the model’s horizontal flow (\( V_F \)), \([T_{\text{AG}} \cdot V_H] V_F \) refers to the horizontal stretching and deformation of \( T_{\text{AG}} \), and

\[ \Pi = -f_0 (\partial w/\partial z)(T_F - T_\theta) - f_0 \partial w/\partial z T_{\text{AG}}/\partial z. \]

Thus Eqs. (22) involve only a single \( R \)-like vector and all the terms that explicitly involve the vertical velocity are assembled on the right-hand side. Again applying the \([k \cdot V_\Lambda]\) operator to Eqs. (22) would yield pseudo-omega equations. Here we merely note that Eqs. (22) approximate to diagnostic relationships
\[ 2f_0 \mathcal{R}_\theta \approx -N^2 \eta_2 + \mathbf{II}, \]  
\[ \text{if } D_H[f_0 \mathcal{T}_{AG}] / Dt \approx -f_0[\mathcal{T}_{AG} \cdot \mathbf{V}_H] \mathbf{v}_F, \]  
\[ \text{and } 2f_0 \mathcal{R}_F \approx -N^2 \eta_2 + \mathbf{II}, \]  
\[ \text{if } D_H[f_0 \mathcal{T}_{AG}] / Dt \approx +f_0[\mathcal{T}_{AG} \cdot \mathbf{V}_H] \mathbf{v}_F. \]  

Direct computations using space–time model output will indicate the measure of applicability of Eq. (23b) or (24b). If either Eq. (23b) or (24b) is applicable then we recover a diagnostic omega equation, whereas inapplicability could be a helpful indicator of regions of unbalanced model flow. Note that if \( \mathcal{T}_{AG} \) is predominantly merely advected with and distorted by the horizontal flow, then

\[ D_H[f_0 \mathcal{T}_{AG}] / Dt \approx f_0[\mathcal{T}_{AG} \cdot \mathbf{V}_H] \mathbf{v}_F \]

(cf. the analogy with material lines), and Eq. (24b) would be almost valid. This might prevail if the unbalanced flow comprises high-frequency, small-amplitude inertia–buoyancy and buoyancy waves that pass through the balanced flow pattern without stretching or deforming it significantly. In a particular circumstance the selection of Eq. (23a) or (24a) will depend upon the relative appropriateness of Eq. (23b) or (24b).

2) MODEL OUTPUT AND \( \mathcal{R} \)

It is possible to infer \( \mathcal{R}_\theta \) directly from the model thermal and wind fields as set out in section 3. An alternative is to consider a corresponding \( \mathbf{Q} \) vector (\( \mathcal{Q}_\theta \)) also based upon the \( \mathcal{T}_\theta \) and \( \mathbf{v}_F \) fields.

Figures 9 and 10 pertain to an event of rapid cyclogenesis over the western Atlantic, and were derived from the assimilation-cum-prediction cycle of the European Centre for Medium-Range Weather Forecasts (ECMWF) operational forecast model. They allow an assessment to be made of the \( \mathcal{R} \) or \( \mathbf{Q} \) patterns of the model fields.

For example, Fig. 9 shows both the computed omega-field and corresponding \( \mathcal{Q}_\theta \)-vector pattern. The consistency of the two fields over much of the domain points to the not unsurprisingly balanced nature of the model state, and concomitantly and more trenchantly to the dominance of this form of forcing. There are discrepancies with some distinctive \( \mathcal{Q}_\theta \)-vector signals not accompanied by significant ascent, and these features can be interpreted in terms of the remarks made in section 3 regarding the properties of \( \mathbf{Q} \) vectors. For example, the band of large \( \mathcal{Q}_\theta \) vectors that arches around the inner core of the incipient cyclone (\( L \)) is symptomatic of a strong rotational but nondivergent component to the \( \mathbf{Q} \) field, and serves to emphasize the desirability of estimating (\( \nabla \cdot \mathbf{Q} \)) or (\( \nabla \times \mathcal{R} \)) to infer the forcing.

Figure 10 pertains to a time 12 h ahead of the features depicted in Fig. 9. The satellite image in Fig. 10a demonstrates that the event under consideration has the classical structure of a rapidly developing cyclone. Figure 10b shows the corresponding analyzed ECMWF thermal pattern and horizontal wind vectors on the 700-hPa level at a given instant and Fig. 10c shows the \( \mathcal{Q}_\theta \) vectors and omega fields. Qualitative inference of the regions of ascent using Fig 10b correspond relatively well with the major features of the realized vertical velocity field (Fig. 10c) including the strong across-front circulations. The inferences also capture the movement of the systems as is evident on comparing with the corresponding \( \mathcal{Q}_\theta \) vectors and omega fields of Fig. 9 that pertain for a time instant 12 h later.
In effect adoption of the $\mathbf{R}$-vector approach, based upon the generalization set out in Eq. (23a), appears to retain some interpretative value in a more general flow setting, and thereby lends credence to the theoretical considerations underpinning the approach to the equation’s derivation.

Again the relative amplitude and spatial structure of the $\mathcal{T}_{AG}$ terms on the left-hand side of Eqs. (22a) and (22b) versus the $\mathbf{R}$ term can be readily evaluated and assessed from model output. A significantly larger amplitude and spatially coherent region of the $\mathbf{R}$ term would be indicative of a locally balanced flow state.

FIG. 10. Features of a rapidly developing cyclone over the western North Atlantic. (a) A Geostationary Operational Environmental Satellite (GOES) East satellite image (channel 1: visible-green to near IR) for 1500 UTC 31 Dec 2006 (i.e., 9 h ahead of the plots in Fig. 9). (b),(c) The 700-hPa surface, computed from the analysis fields of the ECMWF operational forecast model for 1200 UTC 31 Dec 2006: (b) depicts the analyzed isentropic pattern (colored) and the horizontal wind vector field, and (c) shows the $\mathbf{Q}_u$ vectors and omega fields. The satellite image is from the archives of NERC Satellite Receiving Station, Dundee University, Scotland (see http://www.sat.dundee.ac.uk/).
whereas a larger amplitude and richer or less coherent spatial structure to the sum of the $T_{AG}$ terms on the left-hand side would be indicative of a nonbalanced flow contribution.

Thus comparison of the terms can be instructive in deciding whether it is useful to employ the corresponding pseudo-omega equation as a diagnostic tool and help diagnose the balanced or unbalanced nature of the local dynamics, and thereby concomitantly shed light on the effectiveness of assimilating data in specific regions.

5. Final comments

The reappraisal of the classical quasigeostrophic omega equation undertaken in this study has served several purposes. It has pinpointed possible shortcomings of using various extant formulations of the forcing to infer regions of ascent by inspection of standard synoptic charts, and prompted consideration of various alternative formulations. In particular it is shown that the deformation term in the refined Sutcliffe formulation can be cast in a form that is illuminating from an interpretative standpoint. Likewise the $\nabla_H \cdot Q$ formulation of the forcing can be recast in terms of the line integral of the $R$ vector. Variants of the $R$ vector have been considered previously, and it is closely related to the $Q$ vector and defines, subject to a caveat regarding the scale of QG flow, the horizontal component of the ageostrophic vorticity accompanying a quasigeostrophic flow.

A series of further considerations highlights the nature of the QG omega equation. It is shown that the dynamics of the $R$ vector can be intrinsic to a quasigeostrophic cascade to finer scaled physical flow systems. It is demonstrated that in the QG setting there is an in-phase relationship between forcing due to cloud-diabatic heating and the associated vertical velocity, and it is argued that this has implications for the assimilation of cloud-diabatic features in NWP models. Finally it is shown theoretically and demonstrated empirically that a variant of the quasigeostrophic $R$ vector retains some interpretative value in a more general flow setting. In particular it provides a rationale for using an extended form of the $R$ vector to interpret model output of richly structured flow features, and a framework for assessing the balanced versus unbalanced state of NWP model output. Thus it is argued that the use of the QG omega equation albeit in a refined form remains meaningful for interpretative purposes, and that the accompanying mild generalization of the $R$ vector is helpful when addressing some contemporary modeling challenges.

Acknowledgments. It is a pleasure to express my appreciation to Dan Keyser, Jonathan Martin, and Reinhold Steinacker for insightful discussions on aspects of the QG omega equation, and to the reviewers for their fleet of helpful comments. My sincere thanks also to Rene Fehlmann for the material in Fig. 7 and to Heini Wernli for preparing Figs. 9 and 10.

APPENDIX A

Basic Relationships

The flow variables of the quasigeostrophic system are defined in terms of the departure ($p^*, \rho^*, \theta^*$) of the pressure, density, and potential temperature away from their horizontal mean values ($p_0, \rho_0, \theta_0$). For the adiabatic flow of a Boussinesq fluid on an $f$ plane, the geostrophic velocity, $v_G$, and the hydrostatic equation take the respective forms

$$v_G = k \times \nabla_H \psi, \quad \text{and} \quad b = f_0 \partial \psi/\partial z, \quad (A1)$$

where $\psi = (p^*/f_0\rho_0)$ is the geostrophic streamfunction, and $b = (g\theta^*/\Theta)$ is the buoyancy with $\Theta$ referring to a constant background value.

It follows that the thermal wind equation and the equation for the vertical component of the relative vorticity ($\zeta_G$) take the form

$$f_0 \partial v_G/\partial z = (k \times \nabla_H b), \quad (A2)$$

and

$$f_0 \partial \zeta_G/\partial z = \nabla^2_H b. \quad (A3)$$

A power series expansion in terms of a Rossby number provides (see Pedlosky 1964) the corresponding evolution equations for the geostrophic velocity, relative vorticity, and potential temperature:

$$D_G v_G/\partial t = -f_0(k \times v_{AG}) - (\nabla_H p_{AG}), \quad (A4)$$

$$D_G \zeta_G/\partial t = -f_0(\nabla_H \cdot v_{AG}), \quad (A5)$$

and

$$D_G b/\partial t = -N^2 w, \quad (A6)$$

with $(\nabla_H \cdot v_{AG}) + \partial w/\partial z = 0$.

Here $N$ denotes a constant mean-state value of the Brunt–Väisälä frequency, $f_0$ is the Coriolis parameter on an $f$ plane, $D_G/\partial t$ refers to the rate of change following the horizontal geostrophic flow, $(v_{AG}, w)$ refer to the “second-order” ageostrophic horizontal and vertical velocities, and $p_{AG}$ denotes the leading-order departure of the pressure field away from a base geostrophic state.
From the above it also follows that the evolution equation for the quasigeostrophic potential vorticity \( q \) is given by
\[
D_G q / Dt = 0 \quad \text{with} \quad q = \left[ \nabla_H^2 \psi + \left( f_0 / N \right)^2 \nabla^2 \psi \right], \tag{A7}
\]
the omega equation takes its conventional form, and
\[
f_0 \zeta_{AG} = -\frac{1}{2} \left[ D_G^2 - D^2 \right] - \nabla_H^2 \psi_{AG}, \tag{A8}
\]
where \( \zeta_{AG} \) is the ageostrophic vorticity, and \( D \) is the total deformation of the geostrophic flow. A further comment on the ageostrophic pressure field, \( p_{AG} \), appearing in Eq. (A4) and in Eq. (A8) is given below.

A compressed form of the QG equations is encapsulated by the prognostic equation for \( q \) and the diagnostic equation for \( w \). The former provides a self-consistent equation for the evolution of “a” geostrophic flow state. In this study the latter equation is to be used to infer the \( w \) field for a realized flow state.

The appearance of the \( \nabla_H p_{AG} \) term in Eq. (A4) indicates that to a leading order the selection of a given pressure rather than a given velocity field to specify the geostrophic flow remains open. In effect, given observational estimates of (perhaps incompatible) pressure and horizontal velocity fields, an ambiguity arises in defining a geostrophic flow, and further dynamical grounds are need to assert the primacy of one field relative to the other. For example, if the spatial scale of the flow is such that to the first order the flow adjusts to the pressure field, then it would be appropriate to define the state using the pressure. In this limit \( p_{AG} \) could then be neglected and the vertical component of the ageostrophic field, \( \zeta_{AG} \), is determined by Eq. (A8).

In particular the primacy of the pressure field is not assured for subsynoptic-scale flow features, and the observed pressure, temperature, and horizontal velocity fields might include ageostrophic signals. Thus there is a dynamical caveat to neglecting the \( \nabla_H p_{AG} \), and this is noted where necessary in the main body of the text.

**APPENDIX B**

**Formulas for the Deformation Contribution to the \( \mathcal{F} \) Forcing**

The contribution of the DEF term in the refined Sutcliffe formulation of the \( \mathcal{F} \) forcing [see Eq. (2b)] takes the form
\[
\text{DEF} = \left[ D_L (\partial D_L / \partial z) - D_S (\partial D_S / \partial z) \right], \tag{B1}
\]
where \( D_L = (\partial u_G / \partial x + \partial u_G / \partial y) \) and \( D_S = (\partial u_G / \partial x - \partial u_G / \partial y) \) are the deformation components of the geostrophic flow in Cartesian coordinates.

Equation (B1) can be reformulated in two suggestive ways. First, it can be cast in terms of the total deformation of the geostrophic flow, \( D = (D_L^2 + D_S^2)^{1/2} \) and the orientation (\( \lambda \)) of its dilatation axis with respect to the \( x \) axis such that \( \tan 2\lambda = (D_L / D_S) \).

Noting that
\[
(D_L^2 + D_S^2)^{1/2} \partial (D_L / D_S) / \partial z = -\left[ 1 + (D_L / D_S)^2 \right] \text{DEF},
\]
it follows that
\[
\text{DEF} = -D_L^2 \left[ 1 + (\tan 2\lambda)^2 \right] \partial (\tan 2\lambda) / \partial z = -D_L^2 \partial (2\lambda) / \partial z. \tag{B2}
\]

Second, it can be cast in a form that incorporates the deformation of the thermal wind. The “thermal wind” \( \mathcal{T} = \partial v_G / \partial z \) is such that \( \mathcal{T} = (1/f_0) (k \times \nabla_H h) \), and hence the two components of the thermal wind, \( \mathcal{T} = (u_G, v_G) \), have deformation signatures given by
\[
either S_1 = (\partial u_G / \partial x + \partial u_G / \partial y) \quad \text{and} \quad S_2 = (\partial u_G / \partial x - \partial u_G / \partial y)
\]
or equivalently a total deformation, \( S^2 = (S_1^2 + S_2^2) \), and an orientation (\( \epsilon \)) of its dilatation axis with respect to the \( x \) axis such that \( \tan 2\epsilon = (S_1 / S_2) \).

It then follows that DEF can be written successively as
\[
\text{DEF} = [D_L S_2 - D_S S_1] \tag{B3}
\]
\[
= [D \sin(2\lambda) S \cos(2\epsilon) - D \cos(2\lambda) S \sin(2\epsilon)]
\]
\[
= -\{D S\} \sin(2\gamma) \tag{B4}
\]
such that \( \gamma = (\epsilon - \lambda) \) is the relative orientation of the dilatation axes.

**REFERENCES**


