Piecewise Potential Vorticity Inversion

CHRISTOPHER A. DAVIS

National Center for Atmospheric Research, * Boulder, Colorado

(Manuscript received 24 May 1991, in final form 27 September 1991)

ABSTRACT

The treatment of the potential vorticity (PV) distribution as a composite of individual perturbations is central to the diagnostic and conceptual utility of PV. Nonlinearity in the inversion operator for Ertel's potential vorticity renders quantitative piecewise inversion (inversion of individual portions of the potential vorticity field) ambiguous. Several methods of piecewise inversion are compared for idealized and observed potential vorticity anomalies of varying strengths. Even as the Rossby number of the balanced solutions increases well past unity, relative differences among the more plausible methods do not increase significantly near the anomaly. These relative differences are also found to be smaller than those obtained by comparing any of the methods to quasigeostrophic inversion. However, differences above and below anomalies increase with increasing Rossby number, suggesting that one cannot uniquely diagnose the interaction of large amplitude PV anomalies.

1. Introduction

Much of the usefulness of potential vorticity (PV) as a diagnostic of quasi-balanced atmospheric dynamics implicitly involves treating portions of the PV distribution in isolation. The kinematic description of systems that comprises “PV thinking” stems from determining the flows associated with these portions (piecewise inversion), thereby assessing how PV perturbations interact. For instance, the conceptual picture of baroclinic instability presented in Hoskins et al. (1985, hereafter HMR) consists of two counterpropagating, interacting Rossby waves that mutually amplify each other. In the two-layer and Eady models, the waves are cleanly separated from each other, and the nature of the mutual interaction of these waves can be seen by inverting the PV associated with each wave individually. Given the proper phase relationship between the two waves, the velocity field of one wave acts to amplify the PV perturbation of the other. Robinson (1989) has also performed similar piecewise inversions for the Charney and Green modes of baroclinic instability. In these cases the PV perturbations are not cleanly separated in space; however, similar kinematic interpretations result.

Piecewise PV inversion has also been applied to observations recently (Robinson 1988; Davis and Emanuel 1991, hereafter DE) and may be especially useful in diagnosing the importance of PV anomalies produced by nonconservative processes. Robinson’s approach to observational diagnosis was through the quasigeostrophic (QG) equations. The QG system consists of advecting both a pseudo-potential vorticity (PPV; Charney and Stern 1962) in the interior and potential temperature along horizontal boundaries with the geostrophic wind. For a β-plane geometry, PPV may be expressed

\[ q_p = \nabla^2 \Psi + f_0 + \beta(y - y_0) + \frac{f_0}{r} \frac{\partial}{\partial \pi} \left( \frac{r}{S} \frac{\partial \Psi}{\partial \pi} \right), \]

where \( \Psi \) is the geostrophic streamfunction,

\[ \pi = C_p \left( \frac{1}{\varrho_0} \right)^{1/2}, \quad r = \left( \frac{\pi}{\varrho_0} \right)^{1/2}, \quad \text{and} \quad S = -\frac{d\Theta(\pi)}{d\pi}, \]

\( \Theta \) being the potential temperature of a static reference state. The gas constants \( C_p, C_v, \) and \( R \) have their standard meteorological values. The relation of \( q_p \) to \( \Psi \), along with boundary conditions on \( \Psi \), specifies the invertibility principle for PPV (HMR). Because this equation is linear, the inversion of PPV perturbations is straightforward once the perturbation field is given. Thus, one can decompose the \( q_p \) field into an arbitrary number of parts and obtain a unique \( \Psi' \) associated with each part. The strength of QG is its simplicity; the weakness is its inaccuracy as the Rossby number becomes \( O(1) \), although it may be argued that QG still gives useful qualitative information in such situations (Kuo et al. 1991).

Ertel’s PV (hereafter EPV), defined as

\[ q = \frac{1}{\rho} \mathbf{n} \cdot \nabla \theta \]

© 1992 American Meteorological Society
(Rossby 1940; Ertel 1942), has the advantage over PPV of being conserved by the full primitive equations. Here \( \rho \) is the density, \( \mathbf{n} \) the absolute vorticity vector, and \( \theta \) the potential temperature; the gradient operator is three-dimensional. This property is especially useful in the real atmosphere because nonconservation of EPV is not related to scaling assumptions as is the nonconservation of PPV. In addition, highly accurate balance approximations exist for the inversion of EPV (Charney 1962; Gent and McWilliams 1983; McIntyre and Norton 1992). One commonly used balance definition is attributed to Charney (1955):

\[
\nabla^2 \Phi = \nabla \cdot \nabla \Psi + 2 \left[ \frac{\partial^2 \Psi}{\partial x^2} \frac{\partial^2 \Psi}{\partial y^2} - \left( \frac{\partial^2 \Psi}{\partial x \partial y} \right)^2 \right].
\]

(1.3)

Here \( \Psi \) is a streamfunction for the nondivergent wind, \( \Phi \) is the geopotential, and the divergence and gradient operators are two dimensional (horizontal) only. This balance assumption is quite accurate as long as the Froude number is small, even when the Rossby number is \( \gg 1 \) (McWilliams 1985). If the horizontal velocity in (1.2) is replaced by the nondivergent wind, one can rewrite (1.2) in terms of \( \Psi \) and \( \Phi \):

\[
q = \frac{g \kappa \pi}{\rho} \left[ (f + \nabla^2 \Psi) \frac{\partial^2 \Phi}{\partial \pi^2} - \frac{\partial^2 \Psi}{\partial \pi \partial x} \frac{\partial^2 \Phi}{\partial \pi \partial x} \right. \\
\left. - \frac{\partial^2 \Psi}{\partial \pi \partial y} \frac{\partial^2 \Phi}{\partial \pi \partial y} \right].
\]

(1.4)

The boundary conditions for this system are Dirichlet conditions for \( \Psi \) and \( \Phi \) on the lateral boundaries and Neumann conditions on horizontal boundaries:

\[
\frac{\partial \Phi}{\partial \pi} = f_0 \frac{\partial \Psi}{\partial \pi} = -\theta, \quad (\pi = \pi_0; \pi = \pi_T),
\]

\( \pi_T \) being the upper boundary. We will use the coupled, nonlinear system consisting of (1.3) and (1.4) for recovering the balanced mass and nondivergent wind fields given Ertel's PV. Details of lateral boundary conditions and numerical techniques appear in DE.

While solution of (1.3) and (1.4) is itself nontrivial, the real difficulty of using EPV lies in piecewise inversion: given the nonlinearities present in the equations. Simply put, there is no unique way to invert perturbations of EPV as there is with PPV. The price of greater accuracy is apparently the introduction of ambiguity. Quantitative application of piecewise inversion to high Rossby number flows is thus hampered by the presence of nonlinearity. This issue may become especially important as PPV diagnostics are applied to mesoscale phenomena. In this paper we will examine how imprecise the notion of piecewise EPV inversion becomes as the Rossby number increases past unity.

One method for inverting EPV perturbations, proposed by DE, uses a linearized version of (1.3) and (1.4). In this paper, a variety of methods is presented for inverting perturbations of EPV, and we examine how they differ from each other and from QG inversion. Semigeostrophic inversion will not be considered because it is not well suited to highly curved flows, yet its associated invertibility statement is still nonlinear when the interior PV is nonuniform.

There are two items to consider when comparing inversion methods. One point concerns differences obtained by inverting different forms of PV or by applying different balance constraints on the same form of PV. The second item is that given a nonlinear invertibility statement for the total flow, there is no unique method of piecewise inversion. We will consider both issues but focus more on the latter. In general, as perturbations become larger in magnitude, different methods of piecewise inversion will yield increasingly disparate results. Our ultimate goal is to assess how rapidly the different solutions depart from each other and at what point the discrepancies are so large as to render solutions only qualitatively useful. It is stressed that there is no unambiguously correct method of piecewise EPV inversion; the method of choice will depend upon the precise problem at hand.

The remainder of the paper is divided into three sections. In section 2, a simple analysis of a nonlinear algebraic equation will be performed to discuss the various methods of piecewise inversion to be studied. This gives us the ability to view the relationship among inversion methods over a large region of parameter space. However, such an analysis can only provide qualitative information; hence, in section 3, idealized inversions of perturbations are considered using the full balance equations. The ultimate comparison of methods is in the context of inverting observed atmospheric EPV anomalies, two examples of which are presented in section 4. Our primary result is that the relative differences between most of the inversion methods are small, generally less than 25% even at large Rossby number. In fact, normalized differences in the nondivergent wind are nearly independent of Rossby number, at least near the EPV anomaly. However, relative errors away from the anomaly appear to grow with Rossby number. This suggests that one may be relatively certain about the local strength of an EPV anomaly but not about the dynamical evolution resulting from its interaction with the remainder of the EPV field.

2. Simple analysis

This section will introduce the various means of inverting EPV perturbations discussed in this paper. Because the actual differential equations are quite complicated, a simple set of nonlinear algebraic equations will be considered to make the analysis tractable. As was done in DE, consider the system

\[
q = AB
\]

(2.1)
along with some relation between \( A \) and \( B \). Equation
(2.1) plays the role of the EPV definition (1.4), and the relation between $A$ and $B$ can be thought of as a balance relation. An important nonlinearity in the system is contained in (2.1); hence, to simplify matters further, we set $A = B$, which is akin to assuming linear balance. Then $\bar{B} = q^{1/2}$ represents the total solution. The positive root is chosen arbitrarily; the result of the analysis below is unaltered if the negative root is consistently chosen. We also define a mean value of $q (\bar{q})$ such that $\bar{B} = \bar{q}^{1/2}$. For now, $\bar{q}$ may be interpreted as some reference value of $q$. The equation for the perturbation is then

$$q' = 2\bar{B}B' + B'^2,$$  

(2.2)

obtained by subtracting the mean equation from the total. Although $B' = B - \bar{B}$ is known, the portion of $B'(B_n)$ associated with a given portion of $q'$ ($q_n$) is unknown and is the subject of our investigation. In all methods discussed here, $q_n$ is fixed, and it satisfies $\sum q_n = q'$. The method proposed by DE also requires $\sum B_n = B'$. Substituting these summations into (2.2) gives

$$\sum_{n=1}^{N} q_n = 2\bar{B} \left( \sum_{n=1}^{N} B_n \right) + \left( \sum_{n=1}^{N} B_n \right)^2.$$  

(2.3)

Davis and Emanuel linearize (2.3) by incorporating the nonlinear term into the linear term and writing

$$q_n = 2 \left( \bar{B} + \frac{1}{2} \sum_{n=1}^{N} B_n \right) B_n = 2 \left( \bar{B} + \frac{1}{2} B' \right) B_n$$  

(2.4)

for the $n$th perturbation. This method of perturbation inversion will be referred to as “full linear” inversion (FL); “full” because all terms are retained. As DE point out, if $A \neq B$, (2.3) contains three terms, two of which are linear. This adds an extra arbitrariness to the partitioning concerning the fraction of the nonlinear term to be grouped with each linear term. For symmetry reasons, DE choose one-half, that is, equal partitioning of the nonlinear term among the linear terms. An alternative to (2.4) follows from neglecting the nonlinear term in (2.2), an approach termed “truncated linear” inversion (TL). With TL inversion, $\sum B_n \neq B'$ in general.

A third method for inverting perturbations of $q$ in the nonlinear system $q = B^2$ is to simply subtract $q_n$ from the total $q$ and solve the full nonlinear problem again. This approach, labeled the ST (subtraction from the total) method, can be expressed

$$q - q_n = B_n^2.$$  

(2.5)

The perturbation associated with $q_n$ is then defined as $B_n^* = B - B_n$. The ST method has been used by Nielsen (1990) in the inversion of localized PV anomalies. The equation satisfied by $B_n^*$ can be written

$$q_n = 2BB_n^* - B_n^*^2.$$  

(2.6)

To relate $B_n^*$ to $B_n$, let $q_n = \alpha q'$, where $\alpha$ is a constant and specifies the fraction of the total $q'$ being inverted. In addition, let $\beta = (q/q')^{1/2}$ such that $\beta$ is near unity when the total perturbation field is very small and near zero when the perturbation field is very large. The parameters $\alpha$ and $\beta$ are sufficient to specify the difference between $B_n^*$ and $B_n$, which can be written (in normalized form)

$$R_n^* = \frac{B_n - B_n^*}{B_n} = 1 + \frac{[1 - \alpha(1 - \epsilon^2)]^{1/2} - 1}{\alpha(1 - \epsilon)}. $$  

(2.7)

The contours of $R_n^*$ as a function of $\epsilon$ and $\alpha$ appear in Fig. 1a. The choice of $B_n$ as the normalizing factor.

![Fig. 1. Plots of normalized differences between the full linear (FL) method and (a) ST ($R^*$) and (b) AM ($R$) nonlinear methods. The parameters $\alpha (~q_n/q')$ and $\epsilon (~q'/q)^{1/2}$, described in the text, define the coordinate axes. Contour interval is 0.1, and contours range from -1 to +1. Contours are only drawn where both fields have values in this range and have no imaginary part.](image-url)
is convenient because it has no singularities and is well defined throughout the parameter space. As \( \alpha \) and \( \epsilon \) approach unity, the difference between the solutions vanishes. This simply says that if the total perturbation is small, or if most of the total perturbation is inverted at once, then (2.4) and (2.6) become identical.

A fourth approach to nonlinear inversion views the \( q \) perturbation as an addition to the mean state (the AM method):

\[
\tilde{q} + q_n = \tilde{B}^2.
\]

(2.8)

Defining \( \tilde{B}_n = \tilde{B} - \tilde{B} \), the relation between \( q_n \) and \( \tilde{B}_n \) is

\[
q_n = 2\tilde{B}\tilde{B}_n + \tilde{B}^2.
\]

(2.9)

Using the same parameters, \( \epsilon \) and \( \alpha \), the normalized relationship between \( B_n \) and \( \tilde{B}_n \) can also be derived:

\[
\tilde{R} = \frac{B_n - \tilde{B}_n}{B_n} = 1 + \frac{\epsilon - \left[ \epsilon^{2} + \alpha(1 - \epsilon^{2}) \right]^{1/2}}{\alpha(1 - \epsilon)}.
\]

(2.10)

The dependence of \( \tilde{R} \) on \( \epsilon \) and \( \alpha \) is displayed in Fig. 1b. As \( \alpha \) approaches zero, (2.10) becomes \( \tilde{R} = (1 - \epsilon^{-1})/2 \), which has a singularity at \( \epsilon = 0 \). The fields in Figs. 1a and 1b differ in sign everywhere but have comparable magnitudes away from singularities. The two normalized differences are precisely equal and opposite for \( \alpha = 0.5 \). The key point is that over a wide range of parameter space, the average of the two nonlinear inversion methods is approximately the result of the full linear inversion method. There are also singularities evident in both fields. However, it is unclear how these singularities relate to the continuous differential equations because \( \alpha \) and \( \epsilon \) are difficult to define from realistic EPV distributions. For values of \( \epsilon \) near unity, both \( R^* \) and \( \tilde{R} \) are small. Even for \( \epsilon \approx 1 \), perturbations can be rather large because of the square-root dependence of \( \epsilon \) on \( \tilde{q}/q \).

The normalized departure of the TL method from the FL method turns out to be independent of \( \alpha \) and can be expressed

\[
R^\text{TL} = \frac{B_n - B^\text{TL}}{B_n} = \frac{1}{2}(1 - \epsilon^{-1}).
\]

(2.11)

This expression is the same as \( \tilde{R} \) in the limit of vanishing \( \alpha \). One possible interpretation of the singularity at \( \epsilon = 0 \) is that the perturbations are so large as to render the reference state unimportant.

### 3. Idealized inversions

The simple analysis of the proceeding section served only to give rough qualitative insight into the nature of the various plausible methods for inverting perturbations of Ertel’s PV. To go further, the full balance equations (1.3) and (1.4) are considered, and we compare different perturbation inversion methods for given idealized distributions of EPV. In Section 4, we will turn to inversions of observed EPV anomalies, but idealized cases are useful because one can consider a simple reference state and apply appropriate boundary conditions for the perturbation.

The FL inversion method used by DE, schematically represented by (2.2), consists of solving the following linear system for \( \Psi_n \) and \( \Phi_n \) given \( q_n \):

\[
\nabla^2 \Phi_n = \nabla \cdot f(\nabla \Psi_n) + L_1(\Psi_n, \Phi_n)
\]

(3.1)

\[
q_n = \frac{g\kappa p}{\rho} \left[ (f + \nabla^2 \Psi) \frac{\partial^2 \Phi_n}{\partial \xi^2} + \frac{\partial^2 \Phi}{\partial \xi^2} \nabla^2 \Psi_n + L_2(\Psi_n, \Phi_n) \right],
\]

(3.2)

where \( \nabla \) = \( \nabla \) \( \times \) \( \nabla \) \( \times \) \( \nabla \),

\[ L_1(A, B) = 2 \left[ \frac{\partial^2 A}{\partial \xi^2} \frac{\partial^2 B}{\partial \eta^2} + \frac{\partial^2 A}{\partial \eta^2} \frac{\partial^2 B}{\partial \xi^2} - 2 \frac{\partial^2 A}{\partial \xi \partial \eta} \frac{\partial^2 B}{\partial \xi \partial \eta} \right], \]

\[ L_2(A, B) = -\frac{\partial^2 A}{\partial \xi \partial \eta} \frac{\partial^2 B}{\partial \xi \partial \eta} + \frac{\partial^2 A}{\partial \eta \partial \eta} \frac{\partial^2 B}{\partial \eta \partial \eta}, \]

and

\[
\frac{\partial \Psi_n}{\partial \pi} = f_0 \frac{\partial \Psi_n}{\partial \pi} = -\theta_n
\]

on upper and lower boundaries. For inverting idealized EPV perturbations, commonly used lateral boundary conditions are periodicity (for wave-like fields) or vanishing normal flow (localized anomalies). For limited-domain inversions from real data, there is no obvious choice for \( \Psi_n \) and \( \Phi_n \) on the lateral boundaries. In these cases, the boundaries are placed more than a Rossby radius away from the perturbations of interest, and homogeneous Dirichlet conditions applied. The system for the truncated linear (TL) inversion is obtained by simply replacing \( \Psi \) and \( \Phi \) by \( \tilde{\Psi} \) and \( \tilde{\Phi} \) in (3.1) and (3.2); the boundary conditions remain the same.

The two linear methods and the AM nonlinear method discussed in section 2 require a balanced mean reference state. This is obtained from (1.3) and (1.4) with the mean EPV and all dependent variables replaced by their mean values. In the present idealized inversions, the mean state will be at rest. However, in section 4, we will be dealing with mean mass and wind fields that are averaged in time but vary in all three spatial dimensions.

#### a. Localized, positive EPV anomaly

Consider a localized anomaly of EPV embedded within a mean state at rest on an \( f \) plane (Fig. 2a). This anomaly, with maximum value of 9 PVU (1 PVU = \( 10^6 \) m\(^2\) K kg\(^{-1}\) s\(^{-1}\)) centered near 400 mb, constitutes the total EPV perturbation. The EPV field falls off from the central maximum out to 300 km as 1/r\(^2\), where \( r \) is the distance from the center. Its balanced nondiver-
gent wind and potential temperature fields are shown in Fig. 2b. Homogeneous Dirichlet lateral boundary conditions are prescribed on Ψ and Φ at x = ±2000 km and y = ±2000 km. Klein (1957) and Thorpe (1985, 1986) have presented solutions for similar localized PV anomalies. Here our objective is the inversion of portions of the anomaly field. For simplicity, some fraction of the anomaly shown in Fig. 2 will be inverted by the different methods outlined above, and solutions will be compared. In other words, the total anomaly will be artificially considered as composed of two collocated pieces of fractional size α and 1−α, to use the notation of section 2. By considering both qa to be collocated, the influence of one piece on the other in the vicinity of the anomaly is maximized, thus providing a local upper bound on the departures between piecewise inversion methods. The value of α that will be chosen is one-third, and the corresponding wind field, obtained from the FL method, is precisely one-third of that shown in Fig. 2b.

Figure 3 shows differences between the various methods with α = 1/3 for the geostrophic wind. Only the right half of each section is shown because of the symmetry about x = 0. The interesting point about the geostrophic wind solution is that the important differences are away from the level of the anomaly. Consistent with the results in section 2, the FL – ST and FL – AM fields (Figs. 3b and 3c, respectively) are roughly equal and opposite.

Figure 4 shows the same fields as Fig. 3, but for the nondivergent wind. As with the geostrophic solutions, the departures of the nonlinear methods from the full linear method are nearly equal and opposite. Here, however, the maximum differences occur at the level of the anomaly. As in Fig. 3, the truncated linear inversion departs most from the full linear inversion. For a mean state at rest, the TL method nearly reduces to QG, except that we are inverting Erte’s PV perturbations. Hence, the difference between the FL and TL inversions (Fig. 3d) at the level of the PV perturbation mainly involves curvature effects, retained only in the FL approach.

To investigate the behavior of the differences as the Rossby number increases, anomalies with amplitudes of 3 PVU and 18 PVU, respectively, are considered. The anomalies are constructed such that the volume integrated perturbation EPV is linearly proportional to the central extremum. We again take α = 1/3. The departures of the ST and FL methods for inverting one-third of the 3 PVU and 18 PVU anomalies appear in Figs. 5a and 5b, respectively. These are qualitatively similar, but at small amplitude; the difference in velocity between the methods is concentrated near the EPV anomaly; at large amplitude, differences above and below the anomaly are nearly as large as those at the anomaly level.

To examine the relationship of the four piecewise inversion methods discussed above more systematically as nonlinearity increases, it is useful to define a few simple measures of the magnitude of the discrepancies. Based on the patterns in Figs. 3, 4, and 5, three parameters are used to quantify the differences among inversion methods. Our concern lies only with the nondivergent wind in the following. First, the maximum wind difference at the level of the anomaly, normalized by the maximum wind of the FL solution at that level, is used. Second, the maximum difference in surface wind, normalized by the maximum surface wind in the FL solution, is used. This parameter measures relative differences one may expect in the advection of the surrounding EPV field. A related practical situation is the advection of surface potential temperature by an
upper-level EPV anomaly, an important component of midlatitude cyclogenesis (HMR; Davis 1990). The third parameter is the volume-average, rms nondivergent wind in the difference field, normalized by the rms velocity in the FL solution.

Table 1 presents the dependence of the normalized difference between the FL and the other three inversion methods on the EPV anomaly magnitude. The ST and FL methods agree most closely; the TL solution has the least agreement. The key result is that the normalized differences in anomaly strength and volume-integrated kinetic energy do not increase with increasing amplitude, at least between the FL and the nonlinear inversion methods. This is despite the increasing size of the nonlinear terms. The TL method’s disagreement with the other methods grows because its winds are geostrophic. At the surface, on the other hand, the relative disagreement between all methods grows steadily. These results imply that one may be more certain about the strength of an anomaly locally than about its ability to advect EPV at a distance. Furthermore, differences away from an anomaly may be exacerbated by the presence of other EPV perturbations, an effect not considered in the above analysis.

b. Vertical penetration

Many aspects of the differences among piecewise inversion methods may be understood in terms of the vertical penetration of EPV anomaly circulations. Vertical penetration refers to the strength of an anomaly circulation at a distance from the anomaly (at the surface in this case) relative to the strength near the anomaly. With QG inversion, there is a well-defined penetration depth that is an external parameter. Because (1.3) and (1.4) are nonlinear, the vertical pen-
etration in the balance equations depends on the solution itself. To quantify this behavior, we perform a sequence of inversions for anomalies with the same shape as in Fig. 2a, but with maxima (\( Q \)) of 1, 2, 3, 4, 6, 8, 9, 12, and 18 PVU. Note that the full balance equations (1.3) and (1.4) are solved in each case; these are not piecewise inversions. The vertical penetration is then defined as the maximum velocity at the surface normalized by the maximum velocity at the level of the anomaly.

Figure 6 shows the dependence of vertical penetration on the magnitude of the EPV anomaly. The two curves represent the set of nine anomalies at different levels, 414 mb (dashed line) and 570 mb (solid line). The general trend is for the penetration to increase with amplitude, an effect inferred by HMR and Thorpe (1986). For the higher anomaly, penetration increases almost linearly within the range shown; however, one expects the curve to approach unity for very large \( Q \). In fact, the curve representing the lower set of anomalies shows this trend. For more intense EPV anomalies, the balanced flow becomes more barotropic, and the structure is therefore less sensitive to the location of the anomaly. As the EPV anomaly becomes very weak, one expects the normalized penetration to agree with the constant value predicted by QG, about 0.27 and 0.12 for the lower and higher anomalies, respectively.

It has been suggested by Raymond (1992) that for intense EPV anomalies the maximum velocity should scale as the magnitude of the EPV anomaly to some fractional power, probably one-third. It turns out that the maximum winds and vorticity increase more slowly than linearly in our parameter range, although not as slowly as a cube root. Some of the increase of vertical penetration is due to the fact that the Rossby number
in the vicinity of the anomaly is larger than near the surface. As a result, the wind maximum increases slower with \( Q \) than the surface wind maximum. Perhaps a more general definition of vertical penetration would be the maximum surface wind normalized by \( Q \) itself. If this quantity is plotted against \( Q \) (not shown), the curves have roughly the same shape as in Fig. 6, although there is no apparent reason for bounded asymptotic behavior at large \( Q \).

The increase of vertical penetration with \( Q \) is generally consistent with the differences between inversion methods above and below the EPV anomaly in Figs. 3-5. With an EPV anomaly of magnitude \( Q \) on a resting basic state, the AM method inverts an anomaly \( aQ \) and, because the mean state is at rest, all of the balanced motion comprises the AM solution. The ST method inverts anomalies of amplitude \( Q \) and \((1 - \alpha)Q \) and subtracts the solutions. Because vertical penetration increases with the strength of the anomaly, there is more penetration in the difference between the inversions of \( Q \) and \((1 - \alpha)Q \) anomalies than in the AM solution. Thus, if the AM solution is subtracted from the ST solution, the signature of enhanced penetration above and below the anomaly appears as cyclonic circulation for \( Q > 0 \). Although Figs. 3-5 do not show the differences between ST and AM winds, they may be inferred and are consistent with the preceding discussion.

Wind differences between the ST and AM methods at the level of the anomaly can be largely attributed to the slower-than-linear increase of maximum wind with \( Q \). The difference in maximum wind between inversions of EPV anomalies of magnitude \( Q \) and \((1 - \alpha)Q \) is smaller than the maximum wind associated with an anomaly of magnitude \( aQ \). Thus, if \( Q > 0 \), the ST solution should exhibit weaker cyclonic flow at the level of the anomaly than the AM solution.

It has been noted that the FL method is nearly the average of the nonlinear methods. In the algebraic sys-

---

**Table 1. Departures of nondivergent winds.** Columns are as follows: \( aQ \) = \( q_a \) in PVU; \( \alpha = 1/3 \); \( \delta_{\text{max}} \); maximum vorticity associated with the anomaly of magnitude \( Q \) (units \( 10^{-4} \) s\(^{-1}) \); \( V_{\text{rms}} \); normalized maximum nondivergent wind difference at the anomaly level; \( V_{\text{rms}} \); normalized rms nondivergent wind (volume averaged); and \( V_{\text{rms}} \); same as \( V_{\text{rms}} \) but evaluated at the lower boundary.

<table>
<thead>
<tr>
<th>( aQ )</th>
<th>( \delta_{\text{max}} )</th>
<th>( V_{\text{max}} )</th>
<th>( V_{\text{rms}} )</th>
<th>( V_{\text{rms}} ); same as ( V_{\text{rms}} ) but evaluated at the lower boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.3( ^a )</td>
<td>-0.99</td>
<td>-0.17</td>
<td>0.13</td>
<td>0.19</td>
</tr>
<tr>
<td>1.0</td>
<td>1.87</td>
<td>0.20</td>
<td>0.08</td>
<td>-0.10</td>
</tr>
<tr>
<td>2.0</td>
<td>2.88</td>
<td>0.23</td>
<td>0.12</td>
<td>-0.20</td>
</tr>
<tr>
<td>3.0</td>
<td>3.70</td>
<td>0.24</td>
<td>0.13</td>
<td>-0.33</td>
</tr>
<tr>
<td>4.0</td>
<td>4.66</td>
<td>0.22</td>
<td>0.13</td>
<td>-0.42</td>
</tr>
<tr>
<td>6.0</td>
<td>6.02</td>
<td>0.19</td>
<td>0.12</td>
<td>-0.50</td>
</tr>
</tbody>
</table>

**Normalized FL - ST**

<table>
<thead>
<tr>
<th>( aQ )</th>
<th>( V_{\text{rms}} ); same as ( V_{\text{rms}} ) but evaluated at the lower boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.3</td>
<td>-0.12</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.27</td>
</tr>
<tr>
<td>2.0</td>
<td>-0.33</td>
</tr>
<tr>
<td>3.0</td>
<td>-0.35</td>
</tr>
<tr>
<td>4.0</td>
<td>-0.35</td>
</tr>
<tr>
<td>6.0</td>
<td>-0.32</td>
</tr>
</tbody>
</table>

**Normalized FL - AM**

<table>
<thead>
<tr>
<th>( aQ )</th>
<th>( V_{\text{rms}} ); same as ( V_{\text{rms}} ) but evaluated at the lower boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.3</td>
<td>-0.18</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.66</td>
</tr>
<tr>
<td>2.0</td>
<td>-0.90</td>
</tr>
<tr>
<td>3.0</td>
<td>-1.09</td>
</tr>
<tr>
<td>4.0</td>
<td>-1.23</td>
</tr>
<tr>
<td>6.0</td>
<td>-1.42</td>
</tr>
</tbody>
</table>

\( ^a \) The negative perturbation EPV magnitude can only be compared to the positive anomalies in terms of volume integrated PV. Thus, \( -aQ \) implies there is \( -1.3 \) times as much integrated perturbation EPV as in the weakest positive anomaly listed. The negative anomaly has a smaller extreme value than the +1-PVU anomaly and occupies a volume about three times larger.
the piecewise inversion methods reveals essentially the same difference fields for the nondivergent wind as shown in Fig. 4. The patterns were qualitatively the same even though the flow direction itself was reversed. This is because the nonlinear terms in the balance equation have the same sign regardless of the sign of the EPV anomaly. The largest differences were collocated with the maximum anticyclonic flow of the anomaly. With $\alpha = 1/3$ as before, the normalized differences were about the same as for a cyclonic anomaly with comparable circulation strength (see Table 1). An exception was the truncated linear inversion, in which differences appeared to be relatively small.

The nonlinear terms alter the vertical penetration of the anomaly in a way opposite to that for positive anomalies. The normalized penetration decreases with increasing amplitude. This effect was inferred by Thorpe (1986). The regime where the penetration decreased rapidly with increasing amplitude was out of our reach, suggesting that the vertical penetration of anticyclonic anomalies is governed primarily by the reference state.

d. *QG inversion*

As mentioned in the Introduction, QG inversion has the advantage of being linear and the disadvantage of being inaccurate for large Rossby number. A crucial difference between QG inversion and the TL method described above is that each uses a different type of potential vorticity as its base. PPV may be calculated from the mass field alone, which for the above inversions is the balanced mass field. Thus, inversion of the total PPV distribution is constrained to yield the same balanced mass field. However, PPV and EPV may have different spatial distributions; hence, results of piecewise inversions generally will not be the same.

Figure 7 displays the perturbation PPV field (Coriolis term removed) calculated from the balanced geopotential solutions for increasingly intense EPV anomalies. The EPV anomaly fields all have the same shape as that in Fig. 2 but have amplitudes of 1, 3, 9, and 18 PVU. As one might expect, when the amplitude is small, the PPV and EPV perturbations are almost identical. The spatial patterns depart markedly at large amplitude, however (note that PPV fields are normalized by the amplitude of the corresponding EPV anomaly). Some differences are expected, because it is the variation of EPV along an isentrope that approximates the variation at constant pressure of PPV (if $R_0$ is small) (Charney and Stern 1962). However, the isentropes become horizontal near the boundary, and on these surfaces EPV is constant, while PPV is clearly not constant horizontally. This merely restates that the Rossby number is large. The central point is that the location of the PV anomaly depends on what type of PV one uses. For example, if a nonconservative process creates a localized EPV anomaly with a balanced flow at large Rossby number, the corresponding PPV...
anomaly will not be localized at all. Thus, there will be an apparent source of PPV at other levels unaccounted for by the explicit source terms. In fairness to QG, the maximum relative vorticity in the balanced solutions must be two to three times the Coriolis parameter for these distortions to be significant. However, mesoscale PV anomalies in frontal zones can easily attain this intensity (Kuo et al. 1991).

4. PV Inversions from observations

The idealized inversions presented above provide some information about the differences obtained from various inversion methods. By considering co-located anomalies, we provided an upper bound on differences near the anomaly itself. Since these were 15%–25%, uncertainties concerning the strength of EPV anomalies seem generally small. We also noted that the relative uncertainty in circulations away from an anomaly was greater and increased with amplitude. In observed atmospheric flows, other EPV anomalies may be present away from the anomaly of interest and may further influence the differences among inversion methods.

One application of piecewise inversion to observations is assessing the strength of PV anomalies created by nonconservative processes, particularly latent heat release. By inverting such anomalies separately, one can assess both the instantaneous contribution of diabatically produced PV to the balanced circulation and how the winds associated with this PV may affect the rest of the PV distribution.

Several studies (Manabe 1956; Kleinschmidt 1957; Kuo and Reed 1988; Whitaker et al. 1988; Hoskins and Berrisford 1988; Kuo et al. 1991) have given examples of positive PV anomalies produced at low levels by condensation. To conserve the mass-weighted volume integral of EPV, there is also a negative anomaly produced at upper levels. The low-level positive anomaly has much more influence on the strength of the instantaneous near-surface circulation. Davis and Emanuel calculate the contribution of such an EPV
feature to the low-level circulation of an observed cyclone and find it to be about 40% (using the FL inversion method). However, one may ask how much this answer depends on the inversion method used or on whether one uses EPV or PPV.

In considering a near-surface EPV anomaly, the lower boundary condition of uniform potential temperature (translating to homogeneous Neumann conditions on $\psi$ and $\Phi$) becomes very important. For instance, if the 9-PVU anomaly of Fig. 2 is centered at the second lowest interior level (763 mb), the associated nondivergent wind maximum is increased by almost 50% (not shown). The lower boundary condition does not allow the isentropes to bow upward underneath the anomaly. As a result, the static stability perturbation at the center is smaller and the absolute vorticity larger than for an anomaly well removed from the boundary. If the bottom surface is artificially lowered 4 km and the anomaly is kept at 763 mb, the balanced wind field reverts back to the one depicted in Fig. 2b. Comparison of piecewise inversion methods with the anomaly near the lower boundary gives nearly the same results as in section 3. The difference fields in this case look like the fields in section 3 at and above the anomaly level (i.e., only the top “half” of the structures presented in section 3 is seen).

Two examples of observed low-level EPV perturbations, each taken from a different case of cyclone development, will now be discussed. The synoptic times considered are 1200 UTC 15 December 1987 (case 1) and 1200 UTC 5 February 1988 (case 2), near the time of maximum intensity for both cyclones. Figure 8 shows the perturbation EPV and relative humidity in west–east cross sections for both cases. The analyses were constructed from the mandatory-level, final, global analyses from the National Meteorological Center on a 2.5° × 2.5° latitude–longitude grid. The synoptic aspects of case 1 appear in Schneider (1990), and DE discuss case 2. For each case, the total and mean PV are inverted by solving (1.3) and (1.4) as discussed in DE. Means are defined as an average in time over a period nearly coincident with the period of the cyclone waves. In case 2, the large-scale trend over the wave period is also added to the mean such that mean fields actually evolve slowly in time. Thus, the mean states considered here are considerably more complicated than the mean state in the idealized inversions.

In the absence of reliable trajectories, the PV created by condensation is considered to be a positive perturbation EPV lying in regions of relative humidity greater than 70%. The threshold of 70% was chosen as a compromise to include EPV that advects out of the ascent region and into weakly subsiding (and hence, subsaturated) air and excludes EPV that is clearly of stratospheric origin. Only EPV perturbations below 500 mb are included.

Figure 9a shows the nondivergent wind associated with the lower EPV anomaly in case 1 obtained using the FL method, and Figs. 9b, 9c, and 9d display the difference between this solution and the ST, AM, and TL inversions, respectively. The patterns all resemble those for the idealized inversions at and above the anomaly level, and the largest relative differences between the FL and nonlinear methods are again about 15%–20%. However, there is a noticeable eastward shift of the difference fields at low levels. In Figs. 9b and 9c, the maximum differences coincide with the zero velocity of the FL solution. This indicates an excellent agreement of the maximum wind speeds but a small shift of the zero line in each field. This shift may be due to the presence of other anomalies, perhaps the intense perturbations of surface potential temperature present in this case. The smallness of differences aloft is probably due to the presence of a lowered tropopause, above which solutions decay rapidly with height.
In Fig. 10, the same fields as in Fig. 9 are shown, but for the low-level PV anomaly in case 2. The strength of this anomaly, as measured by the winds, is almost twice the strength of the perturbation in case 1. The normalized differences between the full linear and the ST and TL inversions are very similar to those in case 1, being about 15% and 50%, respectively. However, the discrepancy between the FL and AM inversion methods has the same sign as that between FL and ST methods. In all previous examples, the full linear inversion was nearly the average of the nonlinear methods, but that is not the case here. It is still true, though, that the magnitudes in Fig. 10c are about 15% of the perturbation wind maximum in Fig. 10a. Thus, in the inversion of these observed EPV anomalies, the associated flow that one obtains does not depend significantly on the method of inversion, except in the case of TL inversion.

The answer does depend on whether one uses EPV or PPV as the base for the inversion. Figure 11 shows the FL, ST, TL, and QG geostrophic wind solutions for the low-level PV anomaly in case 1. Geostrophic winds are compared to remove effects of flow curvature that are excluded in QG. While the geostrophic winds never differed by more than 25% or so among the EPV inversion methods, QG differs by roughly a factor of two. The QG inversion departs similarly for the inversion of the case 2 PV anomaly.

As suggested by DE, QG inversion seems to differ most from EPV inversion for PV in frontal zones and typically overemphasizes the strength of positive anomalies located there, especially in cyclonic flow. To see the reason for this, consider a hydrostatic, Boussinesq, f-plane fluid with a constant reference-state stratification $\alpha \partial \theta / \partial z$. One may formulate the difference between EPV and PPV as

$$\rho \alpha \frac{\partial \theta}{\partial z} q_p = \frac{\partial \theta}{\partial z} \left( \zeta' - \zeta'_p \right) + \zeta' \frac{\partial \theta'}{\partial z} \frac{\partial v'}{\partial z} + \frac{\partial u'}{\partial y} \frac{\partial \theta'}{\partial y}. \quad (4.1)$$
Here primes denote departures from a reference state. The first term on the right-hand side is negative where there is cyclonic curvature vorticity, and the last two terms are negative definite for a flow in thermal-wind balance. The term $\zeta \partial^2 \theta / \partial z^2$ is negative in a uniform-PV front, but it may have either sign when the PV is nonuniform, depending on the sign of $\partial^2 \theta / \partial z^2$. For a localized EPV anomaly, $\partial^2 \theta / \partial z^2$ changes sign above and below the anomaly from its value at the center. Therefore, the stratification perturbations of spatially separate, like-signed PV anomalies tend to cancel near the anomalies if they are aligned vertically. In the observational examples shown above, $\partial^2 \theta / \partial z^2$ associated with a positive surface $\theta$ anomaly opposed $\partial^2 \theta / \partial z^2$ of the interior PV. The cancellation was enough to render the right-hand side of (4.1) negative and, therefore, the interior PPV perturbation was stronger than the corresponding EPV perturbation. However, one would expect that if $\partial^2 \theta / \partial z^2$ associated with the interior PV were much larger than $\partial^2 \theta / \partial z^2$ for surrounding EPV anomalies, the right-hand side of (4.1) could become positive. This is depicted in Fig. 7, where there is only one PV anomaly and the EPV is locally larger.

5. Conclusions

Piecewise inversion of potential vorticity is uniquely defined only in the case of linear inversion operators, such as QG pseudo–potential vorticity. For large-scale motions, QG can provide an accurate description of potential vorticity dynamics. However, there is increasing interest in applying “PV thinking” to large Rossby number flows and the study of phenomena on scales of a few hundred kilometers or less. The dynamical framework of Ertel’s PV and associated nonlinear balance may provide considerable insight into the behavior of systems traditionally classified as mesoscale. However, PV thinking relies on an ability to identify the salient PV features in the flow and treat them independently. The nonlinear nature of (1.3) and (1.4) makes this approach inherently ambiguous.
Because there is no unique method for inverting EPV perturbations, this paper has examined the relationship of several possible techniques. As long as equally plausible methods yield nearly the same flow associated with a given anomaly, there is no serious dilemma. We have found that it is crucial to account for the nonlinear terms in the balance equations. Three methods have been examined that retain such terms, and the relative differences between methods near an anomaly do not grow with increasing Rossby number. The disadvantage of the nonlinear piecewise inversion methods (ST and AM) is that the sum of solutions for each PV anomaly does not equal the total perturbation flow. If the residual flow is not dynamically inert, interpretation problems can arise. The full linear (FL) method, which hides the nonlinear terms in a nonconstant coefficient linear differential operator, is defined such that the sum of the piecewise solutions equals the total. This property allows one to think of the inversion process as a Green's function technique, even though one may not be able to formally write the Green's function for the operators defined in (3.1) and (3.2).

As the strength of an EPV anomaly increases, the spatial pattern of the differences between methods changes. At large Rossby number, the largest differences (between the FL and nonlinear methods) appear well away from the anomaly itself, a result of the dependence of vertical penetration on amplitude. The fact that differences away from the anomaly grow as nonlinearity increases implies that advection of other EPV perturbations, and hence implied dynamical evolution, may depend more seriously on the choice of piecewise inversion method.

The dependence of penetration on EPV anomaly sign and amplitude was investigated further. For positive, midlevel EPV perturbations, the normalized surface wind and pressure perturbation both increase with $R_0$ (and the opposite for negative anomalies). Because the Rossby number for negative anomalies is approximately bounded by unity, these nonlinear effects will be manifested mostly in positive, large amplitude PV anomalies. This behavior may have a practical implication for midlatitude cyclogenesis. The coupling between cyclonic anomalies associated with positive PV
perturbations at the tropopause and locally warm surface potential temperatures (positive PV equivalent) will increase with amplitude faster than linearly. This may be an important component of deep, cyclonic wrapups that often accompany atmospheric cyclogenesis.

We should point out that the dependence of vertical penetration on Rossby number depends on how the Rossby number is varied. As shown by Raymond (1992), solutions to the balance equations without rotation have no vertical penetration. If one increases the Rossby number by reducing the ambient rotation, the penetration depth decreases.

A final comment concerns the practical use of piecewise inversion techniques to diagnose the behavior of the real atmosphere or numerical simulations. The use of PPV is attractive because the invertibility relation is linear and computationally inexpensive to solve. However, when the Rossby number is large, the distribution of PPV can be qualitatively different from the distribution of EPV.

Although more accurate, the piecewise EPV inversion methods are more complicated than piecewise QG inversion because one must solve the full nonlinear system (1.3) and (1.4) twice, regardless of the method chosen. In addition, the FL and TL methods require the solution of a linear system for the perturbation. The computational costs of each piecewise EPV inversion method are comparable, the cheapest being dependent on the particular problem. As one may expect, the solution of the nonlinear system is generally more costly than solution of the linear system. Thus, if the number of PV perturbations (\( n \)) is large, the FL or TL methods may be cheapest. In this case, one is solving two nonlinear systems and \( n \) linear systems versus \( n + 1 \) nonlinear inversions with the ST and AM methods. At the other extreme, when one is inverting a single EPV perturbation, the nonlinear ST or AM methods are probably cheaper, especially if the perturbation is weak.

Perhaps the most inexpensive computational approach would be to assume that the observed (or modeled) mass and nondivergent wind fields are in balance and to proceed directly to the FL piecewise inversion using (3.1) and (3.2). There are errors in this assumption that will result in the sum of piecewise solutions not equaling the total perturbation flow. Discrepancies will be on the order of the departure of the nondivergent wind and mass fields from balance, which appears to be generally small. With this approach, the FL inversion method has nearly the same computational expense as QG inversion, but retains its greater accuracy.

Acknowledgments. The author would like to thank Drs. Chris Snyder and Jeff Whitaker of the National Center for Atmospheric Research (NCAR) and Professor David Raymond of New Mexico School of Mining and Technology for their helpful comments during the course of this work. This paper was written while the author was in the Advanced Studies Program at NCAR.

REFERENCES