Chapter 7

Ensemble Synoptic Analysis

GREGORY J. HAKIM AND RYAN D. TORN

University of Washington, Seattle, Washington

(Manuscript received 9 September 2004, in final form 17 May 2005)

Hakim **Torn**

ABSTRACT

Synoptic and mesoscale meteorology underwent a revolution in the 1940s and 1950s with the widespread deployment of novel weather observations, such as the radiosonde network and the advent of weather radar. These observations provoked a rapid increase in our understanding of the structure and dynamics of the atmosphere by pioneering analysts such as Fred Sanders. The authors argue that we may be approaching an analogous revolution in our ability to study the structure and dynamics of atmospheric phenomena with the advent of probabilistic objective analyses. These probabilistic analyses provide not only best estimates of the state of the atmosphere (e.g., the expected value) and the uncertainty about this state (e.g., the variance), but also the relationships between all locations and all variables at that instant in time. Up until now, these relationships have been determined by sampling in time by, for example, case studies, composites, and time-series analysis. Here the authors propose a new approach, ensemble synoptic analysis, which exploits the information contained in probabilistic samples of analyses at one or more instants in time.

One source of probabilistic analyses is ensemble-based state-estimation methods, such as ensemble-based Kalman filters. Analyses from such a filter may be used to study atmospheric phenomena and the relationships between fields and locations at one or more instants in time. After a brief overview of a research-based ensemble Kalman filter, illustrative examples of ensemble synoptic analysis are given for an extratropical cyclone, including relationships between the cyclone minimum sea level pressure and other synoptic features, statistically determined operators for potential-vorticity inversion, and ensemble-based sensitivity analysis.

1. Introduction

Progress in synoptic and dynamic meteorology is often marked by the advent of new or improved observing systems and novel methods for analyzing and understanding these data. For example, deployment of the routine radiosonde network and the emergence of quasigeostrophic (QG) theory provided an opportunity for analysts like Fred Sanders to both better document and understand synoptic-scale weather systems during the latter half of the twentieth century. We propose that emerging techniques in state estimation (‘‘data assimilation’’) may offer new opportunities to analyze and understand atmospheric phenomena. We group these opportunities under the title ‘‘ensemble synoptic analysis’’ (ESA), which derives from the fact that the methods pertain to ensembles of analyses valid at an instant (or

Corresponding author address: Gregory J. Hakim, Department of Atmospheric Sciences, Box 351640, University of Washington, Seattle, WA 98195-1640.
E-mail: hakim@atmos.washington.edu

multiple instants) in time. Our goal is to outline the analysis techniques available to ESA and to demonstrate these techniques through illustrative examples.

In deriving basic understanding of atmospheric phenomena, the analysis often revolves around discovering and exploiting relationships between fields and between locations; for example, for extratropical cyclones, geostrophic balance relates wind and pressure, and tropopause disturbances make important contributions to surface development. Up until now, these relationships have been determined by sampling methods involving long periods of time, such as multiple case studies, composites, and time series analysis. A central attribute of ESA that distinguishes it from these other methods is that it uses a probabilistic estimate for the analysis, rather than a single deterministic analysis. Specifically, an ensemble approach is used to generate a sample of analyses valid at an instant (or multiple instants) in time. Although not yet available from operational centers, these probabilistic analyses may reach operational deployment in the future. (A pseudo-operational system has been available online at www.atmos.washington.edu/~enkf/ since December 2004.)

Current three-dimensional variational data assimilation (e.g., 3DVAR) systems rely upon knowledge of established dynamical relationships, such as hydrostatic and geostrophic balance, to specify covariance relationships and to determine the state (analysis). Here we reverse the process so that *data assimilation is used as a tool for revealing dynamical relationships.* The technique used to generate probabilistic analyses is an ensemble Kalman filter (EnKF), which we will describe in more detail in section 2. This technique applies an ensemble of nonlinear forecasts to approximate the extended Kalman filter, which for Gaussian errors and linear operators provides the maximum likelihood state that also has minimum error variance. The EnKF provides the best estimate of the state of the atmosphere (ensemble mean), state error (ensemble variance), and the relationships between all locations and all variables (ensemble covariance).

Although it is tempting to perform ESA on operational ensemble forecasts, the forecast lead times must be chosen such that the memory of the initial perturbations are lost, since the analysis ensembles are currently specified by ad hoc methods that are not designed to sample the probability distribution of the analysis (e.g., total-energy singular vectors and bred-grown modes). As such, the initial ensemble covariance may be inappropriate for use in ESA; therefore, we apply an EnKF to determine ensemble analyses and forecasts.

After a brief overview of the EnKF in section 2, ESA is defined in section 3. Ensemble analyses and select covariance relationships are discussed for an extratropical cyclone in section 4. Section 5 is devoted to using ESA for potential vorticity (PV) inversion, and section 6 is devoted to using ESA for sensitivity analysis. A summary is provided in section 7.

2. Ensemble Kalman filters

Modern state estimation involves the synthesis of observations and a model's estimate of these observations by appropriately weighting these two pieces of information. This weighting depends on the error associated with the observations relative to the model estimate of the observations. Current operational state estimation systems are deterministic, and as such they do not provide the probabilistic, flow-dependent data needed for ESA. Such probabilistic analyses are now being explored by the research community, including those generated by EnKFs (e.g., Evensen 1994; Houtekamer and Mitchell 1998; Hamill and Snyder 2000). A brief overview of state estimation and the EnKF is provided here; the interested reader may find more background information in Daley (1993), Kalnay (2002), and Hamill (2006).

In mathematical terms, the maximum likelihood analysis, assuming Gaussian statistics, is determined by

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y}^0 - \mathcal{H}(\mathbf{x}^b)], \tag{1}$$

where $\mathbf{x}^a$ is the analysis state vector,[1] $\mathbf{x}^b$ is the background state vector, $\mathbf{y}^0$ is the observation vector, and $\mathcal{H}(\mathbf{x}^b)$ is a vector-valued function that returns a column vector of observations, as estimated by the background state. Normally the background is given by a short-term model forecast, and therefore $\mathcal{H}(\mathbf{x}^b)$ is the model's estimate of the observations. This operation can be as simple as linear interpolation from model grid points to observation locations, but it may also be a complicated nonlinear function of the state, such as a radar equation.

The Kalman gain matrix is $\mathbf{K}$, which is given by

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{2}$$

where $\mathbf{P}^b$ is the background-error covariance matrix, $\mathbf{R}$ is the observation-error covariance matrix, and superscript T denotes the matrix transpose. The diagonal elements of $\mathbf{P}^b$ and $\mathbf{R}$ contain the error variance in the background and observations, respectively, and the off-diagonal elements indicate the covariance relationships between model state variables and observations, respectively. Matrix $\mathbf{H}$ is the linearization of $\mathcal{H}$ about the background state. The error covariance matrix for the background estimate of the observations is $\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}}$ (i.e., the same as $\mathbf{R}$ except that it applies to the model's estimate of the observations). Essentially, $\mathbf{K}$ determines the weight given to the new observational information [i.e., the innovation $\mathbf{y}^0 - \mathcal{H}(\mathbf{x}^b)$] relative to the background estimate. Considering just a single observation, so that $\mathbf{R}$ and $\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}}$ are scalars, observations with large errors relative to the background ($\mathbf{R} \gg \mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}}$) have

[1] The vectors $\mathbf{x}^a$ and $\mathbf{x}^b$ are constructed from the multivariate three-dimensional arrays of model gridpoint values by packing all gridpoint values into a one-dimensional array (a column vector) following a chosen procedure. The details of the procedure may be chosen arbitrarily, provided that it is invertible; the three-dimensional arrays must be recoverable from the vector values.

small $\mathbf{K}$; therefore, the analysis is weighted toward the background estimate.

The crucial difference between the EnKF and 3DVAR involves the assumptions for $\mathbf{P}^b$. For 3DVAR, $\mathbf{P}^b$ is typically fixed to assumed relationships (e.g., background errors are approximately time independent and spatially homogeneous and isotropic), whereas for the EnKF, $\mathbf{P}^b$ depends on time and space and is estimated with an ensemble of nonlinear forecasts by

$$\mathbf{P}^b = \frac{1}{M-1}\mathbf{X}^{b\prime}\mathbf{X}^{b\prime\mathrm{T}}. \tag{3}$$

Here $\mathbf{X}^{b\prime}$ is a matrix containing an ensemble background state estimate in M column vectors, with one ensemble member in each column vector; the superscript prime notation indicates that the ensemble-mean state has been removed from each ensemble member. Because $\mathbf{P}^b$ is flow dependent, the influence of an observation is expected to reduce analysis error more than the 3DVAR flow-independent $\mathbf{P}^b$. The EnKF comes at the added expense of having to integrate the full nonlinear forecast model M times, rather than once, as for 3DVAR. These M forecasts are then all updated with new observations using an ensemble square root filter as described in Whitaker and Hamill (2002). This analysis ensemble is then immediately available for initializing a fresh ensemble forecast and another assimilation step. Note that there is no need to generate synthetic perturbations around a single deterministic analysis as is the current procedure for populating ensembles at operational centers.[2]

Our implementation of an EnKF at the University of Washington utilizes the Weather Research and Forecasting model, version 2.0.2 (WRF; Michalakes et al. 2001), in a perfect model scenario. A ''truth'' integration is performed first, which is then sampled to generate observations that are assimilated with the EnKF to produce analysis ensembles. Each of the 100 ensemble members utilizes the same model configuration: ~100-km horizontal grid spacing on a 90 × 90 grid, with 28 vertical levels. Model physical parameterizations include warm-rain microphysics, the Medium-Range Forecast (MRF) planetary boundary layer scheme (Hong and Pan 1996), and the convective parameterization scheme of Janjic (1994). Observations consist of 250 randomly spaced surface pressure observations sampled from the truth run, which employs Global Forecast System (GFS) analysis lateral boundary conditions. The uniform observing system used here is, of course, unrealistic but has the advantage of being straightforward to implement and allows for unambiguously defined analysis errors. Surface pressure observations have been shown to effectively constrain tropospheric analysis errors on the synoptic scale (Whitaker et al. 2004) and are useful for quickly generating the ensemble analyses needed for ESA. Although the results depend on this choice of observation network, our main purpose here is to illustrate ESA techniques on a suitable dataset. The filter is initialized at 0000 UTC 24 March 2003, about 4 days prior to the time of interest; the analysis error statistics come into equilibrium with observation errors within about 36 h (i.e., by about 1200 UTC 25 March 2003).

[2] As a technical matter, an initial ensemble is required to start an EnKF, although the specific choice becomes unimportant after the filter has cycled through several assimilation steps. Here we populate the initial ensemble with forecasts of different lead time verifying at the same time (0000 UTC 24 March 2003).

3. Ensemble synoptic analysis

ESA is explored here using *linear* relationships and Gaussian statistics. ESA is not necessarily limited by these assumptions, but it seems logical to explore these prior to more complicated nonlinear relationships and non-Gaussian statistics. Given two multivariate ensemble samples of data that have had the ensemble mean removed (indicated by superscripted primes), the individual ensemble members are stored in column vectors of the matrices $\mathbf{X}'$ and $\mathbf{Y}'$, which are size $n \times M$ and $N \times M$, respectively. Each of the M columns represents an ensemble member, which has $n - 1$ and $N - 1$ degrees of freedom for $\mathbf{X}'$ and $\mathbf{Y}'$, respectively. A specific example to be considered more thoroughly in section 5 applies to potential vorticity inversion, where $\mathbf{X}'$ represents the potential vorticity at n grid points, and $\mathbf{Y}'$ represents the state variables that are recovered from the PV inversion (e.g., u, v, and T at all the grids points gives $N = 3n$).

A linear relationship between $\mathbf{X}'$ and $\mathbf{Y}'$ may be expressed as

$$\mathbf{Y}' = \mathbf{L}\mathbf{X}', \tag{4}$$

where $\mathbf{L}$ is an $N \times n$ linear operator that maps $\mathbf{X}'$ into $\mathbf{Y}'$; that is, the relationship between $\mathbf{X}'$ and $\mathbf{Y}'$. The goal here is to recover $\mathbf{L}$ given the samples for $\mathbf{X}'$ and $\mathbf{Y}'$, which is a standard problem in statistics. The solution

$$\mathbf{L} = \mathbf{Y}'\mathbf{X}'^{-1} \tag{5}$$

assumes that $\mathbf{X}'$ is invertible. When $\mathbf{X}'$ is singular or not square, a pseudoinverse may still exist (e.g., Golub and Van Loan 1996, p. 257), which will be further explained below; hereafter, references to inverse matrices are understood to include the pseudoinverse. Prior to exploring the properties of this solution, we consider an alternative expression for (5). Right multiplying (4) by $\mathbf{X}^{\mathrm{T}}$ gives

$$\mathbf{Y}'\mathbf{X}'^{\mathrm{T}} = \mathbf{L}\mathbf{X}'\mathbf{X}'^{\mathrm{T}} \tag{6}$$

or, equivalently,

$$\mathrm{cov}(\mathbf{Y}', \mathbf{X}') = \mathbf{L}\,\mathrm{cov}(\mathbf{X}', \mathbf{X}'). \tag{7}$$

This reveals that

$$\mathbf{L} = \mathrm{cov}(\mathbf{Y}', \mathbf{X}')\mathrm{cov}(\mathbf{X}', \mathbf{X}')^{-1}, \tag{8}$$

where cov is shorthand notation for a covariance matrix

of the indicated arguments. This expression shows that the operator $\mathbf{L}$ may be understood as a linear regression of $\mathbf{Y}'$ on $\mathbf{X}'$. Although (8) provides a useful interpretation for $\mathbf{L}$, it is not computationally efficient.

Returning to (5), a formidable calculation might be anticipated given the size of the state vector for the problems we consider ($N \sim 10^6$). A major simplification to this potentially large calculation is available because of the small size of the ensembles that we consider ($M \sim 100$). Using the singular value decomposition (SVD) for $\mathbf{X}'$ (e.g., Golub and Van Loan 1996),

$$\mathbf{X}' = \mathbf{USV}^{\mathrm{T}}, \tag{9}$$

which applies to any matrix; the inverse of $\mathbf{X}'$ is given by

$$\mathbf{X}'^{-1} = \mathbf{VS}^{-1}\mathbf{U}^{\mathrm{T}}. \tag{10}$$

Here we have used the fact that $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices (transpose is the inverse), and $\mathbf{S}$ is diagonal. An even better expression for the inverse derives from using (9) to eliminate $\mathbf{U}$ in favor of $\mathbf{X}'$,

$$\mathbf{X}'^{-1} = \mathbf{VS}^{-2}\mathbf{V}^{\mathrm{T}}\mathbf{X}'^{\mathrm{T}}, \tag{11}$$

so that from (5) we find

$$\mathbf{L} = \mathbf{Y}'\mathbf{VS}^{-2}\mathbf{V}^{\mathrm{T}}\mathbf{X}'^{\mathrm{T}}. \tag{12}$$

Thus in order to determine $\mathbf{L}$ from the known data $\mathbf{X}'$ and $\mathbf{Y}'$, only matrices $\mathbf{V}$ and $\mathbf{S}$ need to be calculated. From (9) we find that

$$\mathbf{X}'^{\mathrm{T}}\mathbf{X}' = \mathbf{VS}^{2}\mathbf{V}^{\mathrm{T}}, \tag{13}$$

so that $\mathbf{V}$ and $\mathbf{S}$ are the eigenvector and eigenvalue matrices, respectively, of $\mathbf{X}'^{\mathrm{T}}\mathbf{X}'$. Of greatest importance is that $\mathbf{V}$, $\mathbf{S}$, and $\mathbf{X}'^{\mathrm{T}}\mathbf{X}'$ are all $M \times M$, or in the present application, 100×100, so they are trivial to manipulate numerically. However, this simplification also reflects the fact that there are only 100 independent degrees of freedom available for the analysis, and this is one factor that will impose a limit on the insight that may be gained from ESA.

These ideas are illustrated for an extratropical cyclone using two examples: potential vorticity inversion and sensitivity analysis. Before discussing the results, it will prove helpful to first consider a brief overview of the case and of the ensemble-determined covariance relationships.

4. ESA applied to an extratropical cyclone

An extratropical cyclone is chosen for analysis because of its familiarity and the fact that Fred Sanders made important contributions to our understanding of these features (e.g., Sanders and Gyakum 1980; Sanders 1986, 1987). The (randomly) chosen case occurred on 28–29 March 2003. During this time, a cyclone moved from the central United States to Michigan as it deepened. An overview of the vertical structure of the system on 0600 UTC 29 March is provided in Fig. 1. The developing low is marked by well-defined warm and cold fronts that are characterized by abrupt horizontal potential temperature gradients and troughs in the surface pressure field (Fig. 1a). A broad and meridionally extensive 500-hPa trough is located to the west of the surface low, as may be expected for a developing surface cyclone (Fig. 1b). What is less clear is the role played by smaller-scale shortwave disturbances, which are more obvious in the 500-hPa PV distribution (Fig. 1b, red lines). As will be shown below, the southern disturbance is apparently important to the surface development of the low over Michigan despite its distant location. On the dynamic tropopause,[3] an elongated region of potentially cold air is located in the region of the 500-hPa trough, as well as a stronger gradient in potential temperature near Texas (Fig. 1c).

Figure 1 also introduces the first example regarding the importance of considering analyses probabilistically by showing all ensemble members (gray lines) in addition to the ensemble mean (solid black lines). All ensemble members must be regarded as equally likely realizations of the state, so that considering just one is potentially misleading (e.g., Leith 1974).

Having considered the ensemble distribution and mean value for selected parameters, we now explore aspects of the ensemble analysis covariance matrix, which is at the heart of ESA for Gaussian statistics. To simplify the presentation and to remain focused on the surface cyclone, covariance relationships are considered for a single grid point: the point of lowest surface pressure, which will be referred to as the cyclone central pressure. The sample of cyclone central pressure values is normalized by the ensemble standard deviation (2.7 hPa), so that the covariances carry the value of the covarying field. Defining this ensemble sample of M normalized surface cyclone central pressure values by the *row vector* $\mathbf{y}'$ and the full $N \times M$ ensemble state matrix by $\mathbf{X}'$, the expression we evaluate is

$$\mathbf{c} = -\mathrm{cov}(\mathbf{X}', \mathbf{y}') = -\mathbf{X}'\mathbf{y}'^{\mathrm{T}}. \tag{14}$$

The prime notation indicates that the ensemble mean has been removed, and the negative sign in (14) is applied to the covariances so that the fields reflect a deeper cyclone. Note that $\mathbf{c}$ is an $N \times 1$ column vector, with each value reflecting the linear relationship between the metric of interest (cyclone central pressure) and all other state variables. Figures 2–4 represent selected parts of vector **c.**

Figure 2 shows the covariance of the cyclone central pressure with the surface pressure and wind fields. When the surface cyclone central pressure is lower than the ensemble mean, as it is for some ensemble members, the pressure at nearby points is also lower; that is, these points covary strongly, as may be expected. The pressure covariance field also shows interesting patterns

[3] Defined here by the 1.5×10^{-6} m^2 K kg^{-1} s^{-1} (PVU) surface.

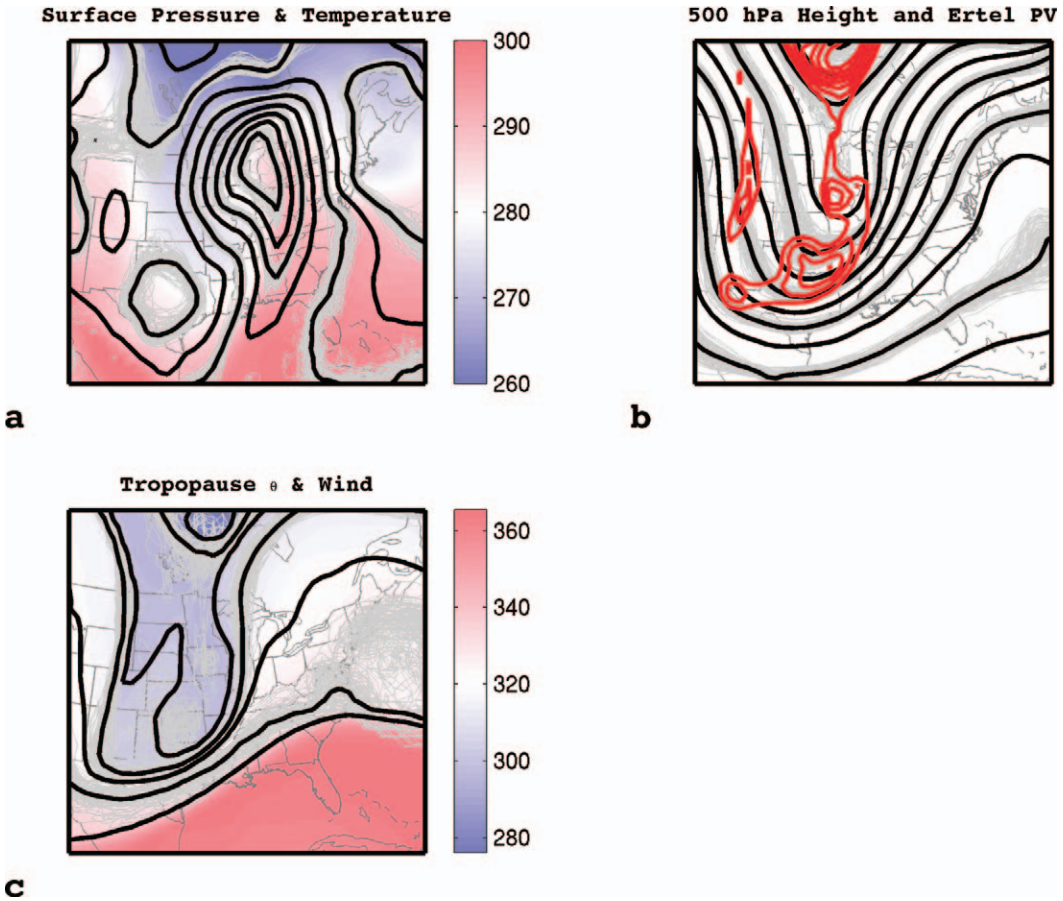

FIG. 1. Ensemble-mean fields at 0600 UTC 29 Mar 2003 of (a) surface pressure (black lines every 4 hPa) and potential temperature (colors); (b) 500-hPa geopotential height (black lines, every 60 m) and Ertel potential vorticity (red lines every 0.25 PVU, starting with 1 PVU); and (c) dynamic tropopause potential temperature (black lines every 10 K and colors). Gray lines show all 100 ensemble members at twice the contour interval of the ensemble-mean field.

away from the low center. Specifically, near the cold front there is a dipole with positive values west of the cold front and negative values to the east. This pattern suggests a phase shift in the location of the pressure trough associated with the cold front such that the surface cold front is displaced eastward when the surface low is deeper; this is intuitively appealing because a deeper low should be associated with stronger cyclonic circulation. The wind field covariance reflects this inference, with a local maximum in wind speed along the front. Note also that the wind vectors have a component directed toward lower pressure, qualitatively consistent with the effect of surface friction on geostrophic flow (e.g., Holton 2004).

At 500 hPa, the geopotential height field covariance with the cyclone central pressure shows that when the surface low is deeper by 2.7 hPa, there is a dipole response at 500 hPa, with lower heights upstream (−10 m) and higher heights downstream (10 m; Fig. 3). This pattern is appealing based on quasigeostrophic reasoning in that the upper-level wave is amplified above the surface low and the dominant wavelength shortened. Interestingly, the 500-hPa trough located over Texas also covaries strongly with the surface low, suggesting that it also contributes to deepening the cyclone despite being located much farther away than the trough over Illinois. The wind field covariance shows clear qualitative evidence of balance, with the vectors following the height field covariance and the magnitude generally proportional to the geopotential height covariance gradient. Balance issues will be examined more closely in the next section on PV inversion.

Moving up to the tropopause (Fig. 4), the potential temperature field shows that where the 500-hPa heights are lower (higher) the tropopause is colder (warmer). A particularly interesting result is the band of warmer air along the subtropical jet stream that extends from Mexico to North Carolina. It is unclear why this region is dynamically related to the surface low over Michigan, although one possibility is that, on the large scale, the surface low is near the left exit region of a planetary-scale jet. The strong horizontal potential temperature gradients both south and north of the cyclone are en-

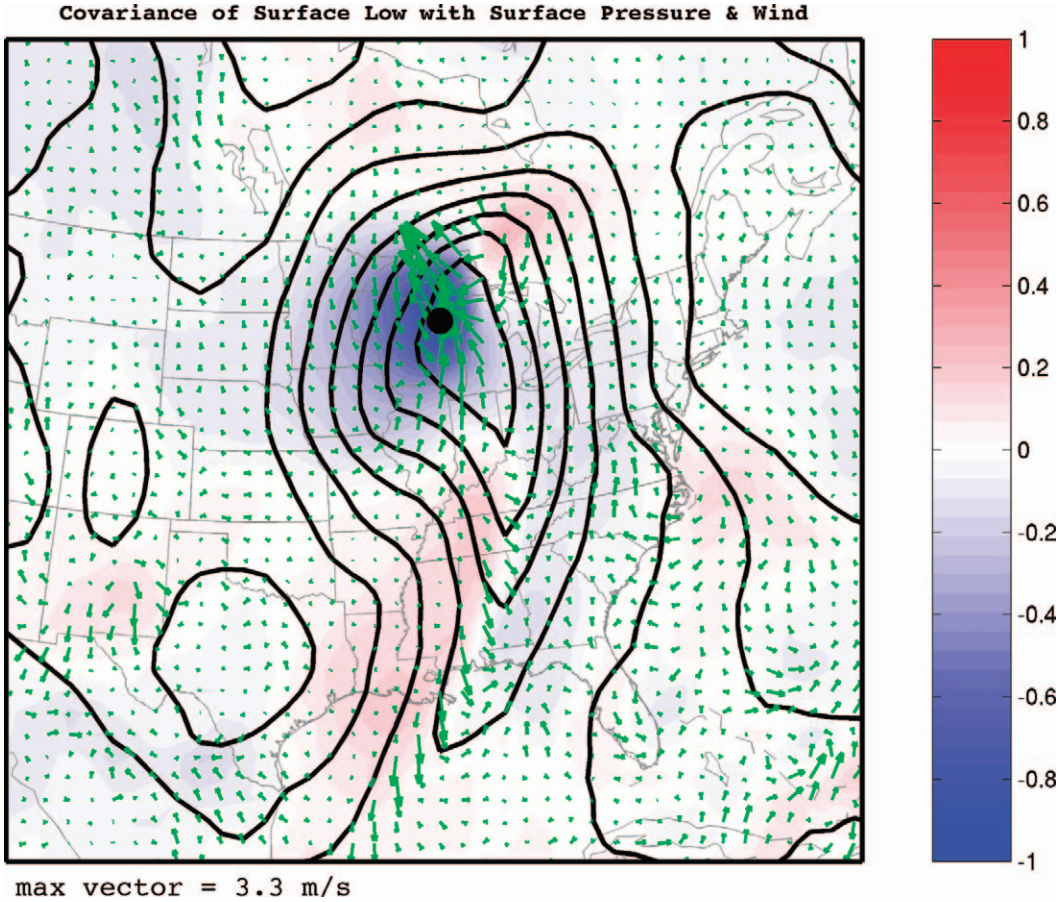

FIG. 2. Ensemble-based covariance fields between the normalized cyclone central pressure and the surface pressure (colors, hPa) and wind (vectors) at 0600 UTC 29 Mar 2003. The ensemble-mean surface pressure is given in solid lines every 4 hPa, and the sample of cyclone central pressure values are normalized by the ensemble standard deviation. The location of lowest pressure in the surface cyclone position is denoted by the black dot.

hanced for a deeper cyclone, which suggests a greater tendency to form upper-level fronts.

To summarize this section, the covariance relationships are generally qualitatively consistent with expectations for an extratropical cyclone, although they also suggest some interesting relationships that may not be obvious using other analysis techniques, for example, the relative importance of the two upper-level disturbances to the surface low and the relationship between the surface low and the subtropical jet. We proceed to use these statistical relationships to perform a more complicated diagnostic: piecewise PV inversion.

5. Statistical potential vorticity inversion

Typically, PV inversion involves the specification of balance constraints between wind, mass, and temperature; boundary conditions; and numerical approximations (e.g., Hoskins et al. 1985; Davis and Emanuel 1991). Here we determine the inversion operator *statistically* using ESA methods outlined in section 3. In the earlier notation the linear operator $\mathbf{L}$ is estimated by

$$\mathbf{L} = \mathbf{X}'\mathbf{P}'^{-1}, \tag{15}$$

where $\mathbf{X}'$ is an ensemble state matrix, $\mathbf{P}'$ is an ensemble PV matrix, and the matrix inverse is calculated as described in section 3. Note also that the symbol $\mathbf{X}'$ represented independent variables in section 3, but here the state vector is dependent on the potential vorticity. A complete specification of $\mathbf{X}'$ and $\mathbf{P}'$ will be given below when individual experiments are discussed. Piecewise PV inversion may then proceed given the operator $\mathbf{L}$ and some subsample of the PV,

$$\mathbf{x}' = \mathbf{L}\mathbf{p}', \tag{16}$$

where $\mathbf{x}'$ is a column vector for the state that is attributable to the column vector of PV, $\mathbf{p}'$. For example, $\mathbf{p}'$ may consist of one column of $\mathbf{P}'$, with all entries but one set to zero, in which case $\mathbf{x}'$ is a (discrete) Green's function. Repeating this example for all rows of $\mathbf{P}'$ but one set to zero yields an ensemble sample for the Green's function, which may be used to estimate a mean solution and variance (i.e., error).

Because the Ertel PV has nonlinear terms, (15) may

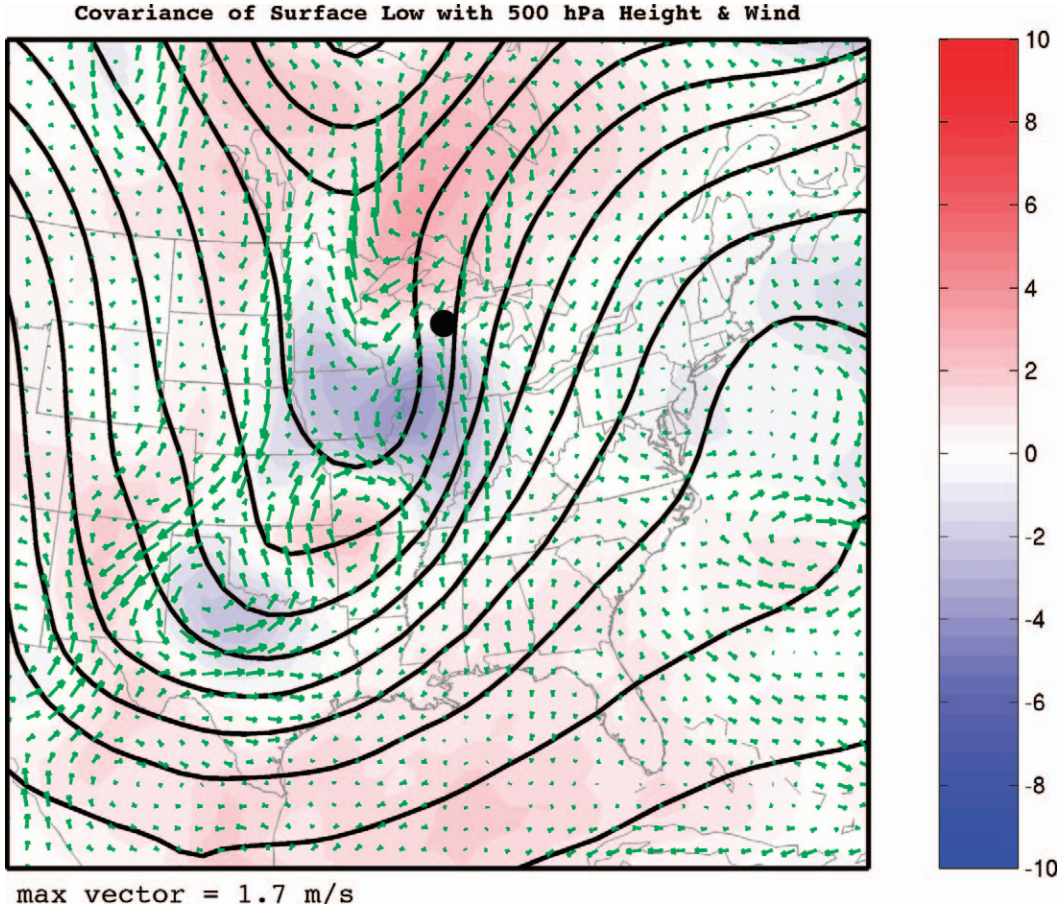

FIG. 3. Ensemble-based covariance fields at 0600 UTC 29 Mar between the normalized cyclone central pressure and the 500-hPa geopotential height (colors, m) and wind (vectors). The ensemble-mean 500-hPa geopotential height field is given by solid lines every 60 m. The sample of cyclone central pressure values are normalized by the ensemble standard deviation, and the surface cyclone position is denoted by the black dot.

seem like an inappropriate estimator for this field. Recall that we have removed the ensemble mean from matrices $\mathbf{X}'$ and $\mathbf{P}'$, so that $\mathbf{L}$ represents a linearization about the ensemble mean. As with all piecewise Ertel PV inversions, some linearization is necessary, and the ensemble linearization is particularly useful because it requires no space or time averaging.

Recall that M, the number of ensemble members (here 100), places a restriction on exactly how much information can be extracted by ESA, and this plays an important role here. One way to think about the problem is that there are only 100 degrees of freedom in the ensemble, so we cannot expect to constrain all $O(10^6)$ degrees of freedom in the numerical model given a small ensemble. Consider a simple example in three-dimensional Cartesian coordinates (x, y, z) for a two-member ensemble with vectors lying in the (x, y) plane. Because there is no information normal to the plane, the z direction is ''invisible'' to the ensemble statistics; in mathematical terms, it does not lie in the span of the ensemble—no linear combination of ensemble vectors in the (x, y) plane project off the plane. For PV inversion, our results are limited to the space spanned by the 100 singular vectors of the potential vorticity field ($\mathbf{V}$ in the notation of section 3), which means that there will be many directions that are invisible. A second way to think about this problem has to do with noise in the covariance calculation at distances far from a particular point. Close to this particular point, the covariance relationships should be better than those far from this point where, by random chance, a small ensemble may suggest a strong relationship. These spurious relationships far from the point introduce noise into the calculation that may adversely affect the inversion. Mathematically this means that covariances between fields physically separated by large distances are potentially spurious; one possible solution often used in the EnKF literature is to simply zero out these entries (e.g., Houtekamer and Mitchell 1998; Hamill et al. 2001). This procedure increases the rank of the covariance matrix (effectively, the size of the ensemble), but tests of this idea are beyond the scope of the present analysis. Our goal here is simply to illustrate proof of concept; a more thorough analysis will be published elsewhere.

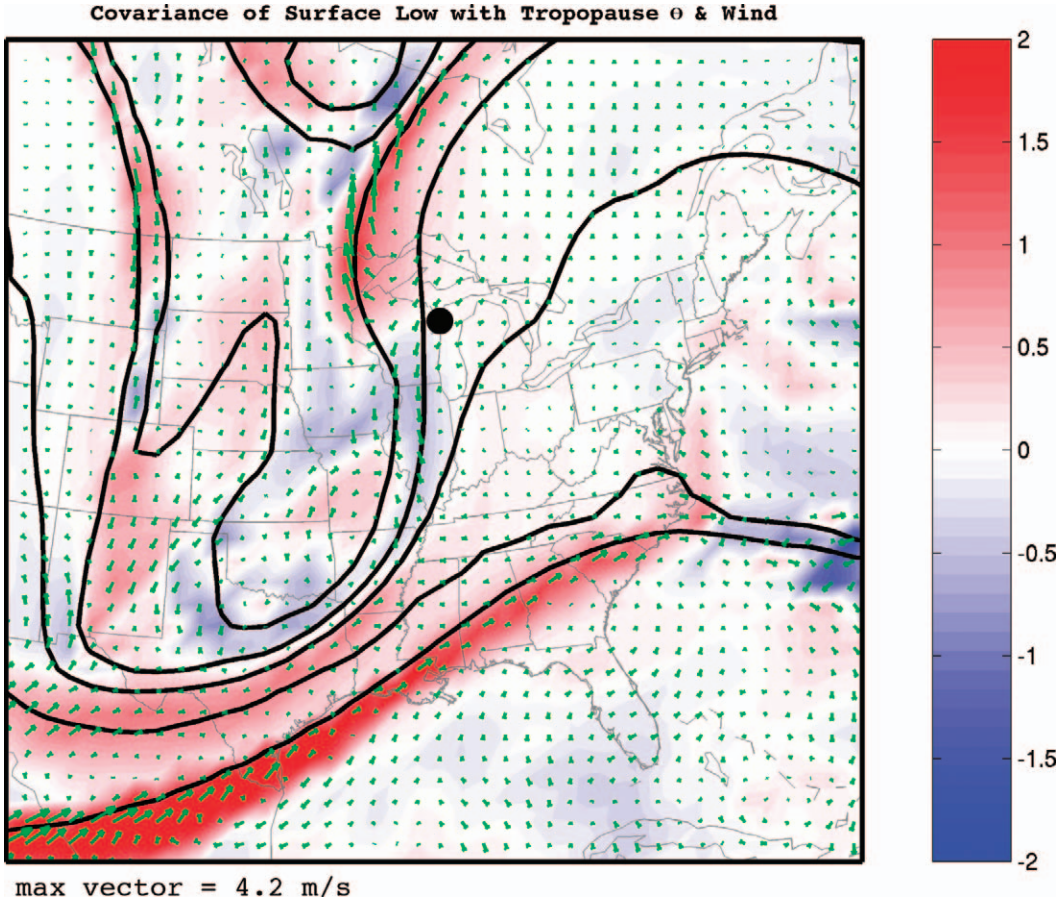

FIG. 4. Ensemble-based covariance fields at 0600 UTC 29 Mar between the normalized cyclone central pressure and the dynamical tropopause potential temperature (colors, °C) and wind (vectors). The ensemble-mean tropopause potential temperature is given by solid lines every 10°C. The sample of cyclone central pressure values are normalized by the ensemble standard deviation, and the surface cyclone position is denoted by the black dot.

Mindful of the problems described above, piecewise PV inversion may be applied to large-scale features over the entire grid or to smaller features in a window around a local region. We choose the latter application by limiting the data to lie close to a point at 500 hPa for two examples: a potential vorticity blob at 500 hPa and an east–west cross section through the same location.

The first example is illustrated in Figs. 5 and 6 for piecewise PV inversion of a blob of PV located mainly over Oklahoma. Inversion (15) is calculated over 21 × 21 grid of points independently for the geopotential height, u, and v fields; that is, three operators are calculated. A piecewise inversion is then accomplished by applying these operators on a subset of the 21 × 21 PV field by choosing a vector that is zero everywhere but for the grid points indicated by the red lines in Fig. 5. The results show that the height field reaches a local minimum near the center of the PV anomaly, and the wind field has cyclonic circulation around the PV blob, with maximum speeds near the edges of the blob, and decays toward zero at larger distances. These patterns are qualitatively in accord with QG theory, which has essentially a Laplacian relationship between the geopotential height field and the potential vorticity field, so that the geopotential height minimum is near the PV maximum. Because no assumptions are made regarding balance, ageostrophic wind vectors may be computed by the difference between the geostrophic wind determined from the inverted height field and the inverted full wind field. The results show anticyclonic circulation about the low geopotential height region; as expected (e.g., Holton 2004, p. 67), the flow is subgeostrophic around the low (Fig. 6).

The present example is more exploratory for the technique rather than definitive of its quantitative properties; nevertheless, a brief analysis of the sensitivity of the results is provided. Increasing the grid size to 31 × 31 points yields a geopotential height minimum of −74 m as compared with −75 m for the 21 × 21 grid and a maximum wind speed of 17 m s^{-1} as compared with 21 m s^{-1} for the 21 × 21 grid. Further increasing the grid to 41 × 41 gives a height minimum of −75 m and a maximum wind speed of 16 m s^{-1}; moreover the plotted fields are very similar to those shown in Fig. 5. The

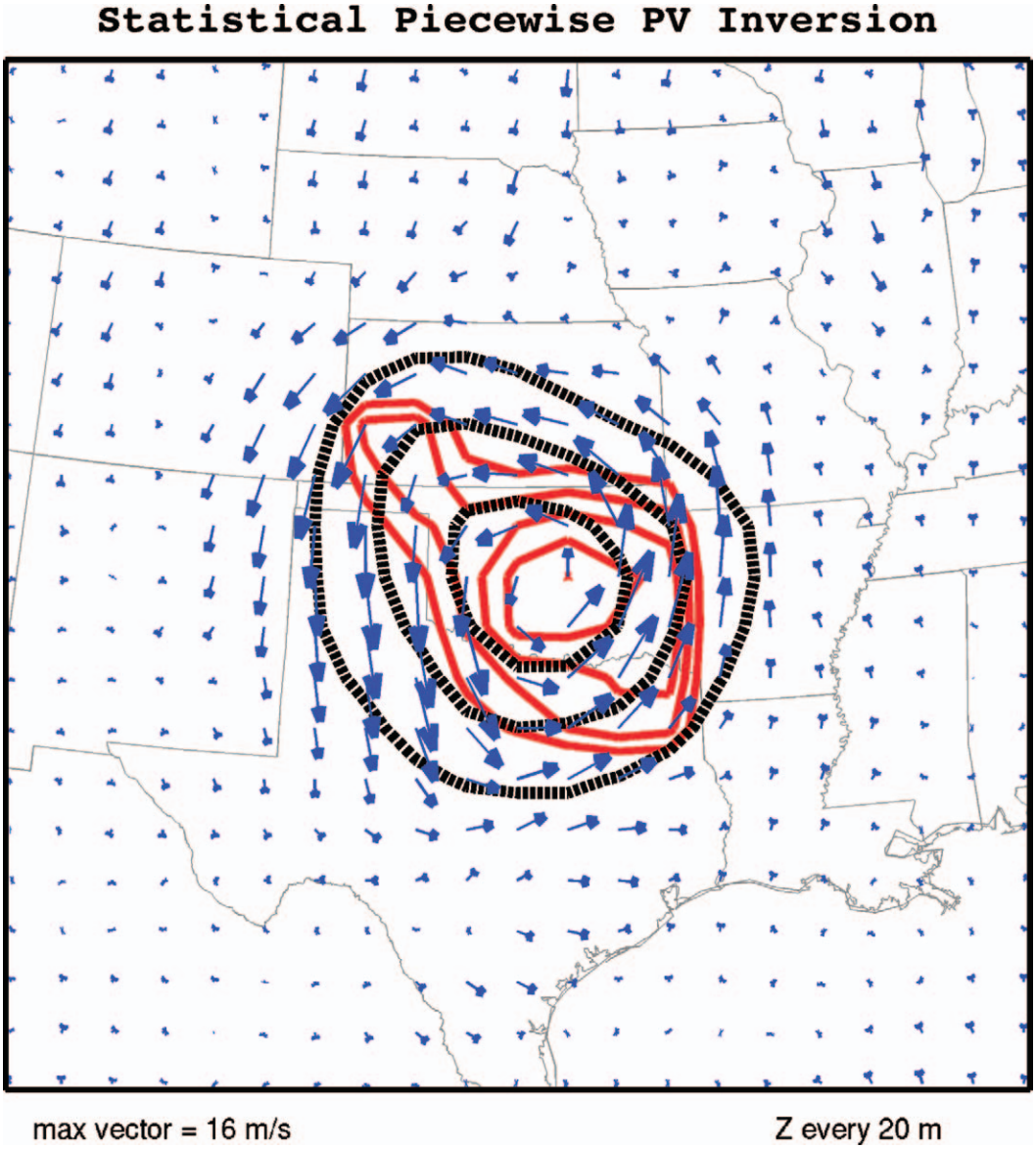

FIG. 5. Piecewise Ertel PV inversion for a region of 500-hPa PV given by the red contours (every 0.5 PVU, starting with 1 PVU) at 0600 UTC 29 Mar. Dashed black lines give the inverted 500-hPa geopotential height field every 20 m, and the vectors give the inverted wind field.

results start to degrade for grid sizes larger than 61 × 61, at which point spurious noise in the covariances probably starts to pollute the local signal in the ensemble. A test of the results with respect to the number of ensemble members indicates that 75 members give results very similar to Fig. 5, and the results do not start to degrade considerably until fewer than 50 members are used (not shown). A convergence test in the space of PV singular vectors shows that 25 singular vectors achieve about the same results as approximately 75 ensemble members, which reflects the optimal dimensional reduction of SVD (not shown).

The above inversion procedure is repeated for the second example, which applies to a zonal cross section through the same PV blob as in Fig. 5. Piecewise PV inversion on PV in the 450–550-hPa layer of the cross section yields a familiar pattern (e.g., Hoskins et al. 1985, their Fig. 15). A local minimum in geopotential height is found near the PV maximum, with warm air above and cold air below the anomaly (Fig. 7). This distribution is qualitatively as expected from hydrostatic balance, although hydrostatic balance is not assumed in the calculation. The wind field normal to the cross section shows the strongest winds at the edges of the PV anomaly, which decrease with increasing distance. Note that because boundary conditions are not specified, there

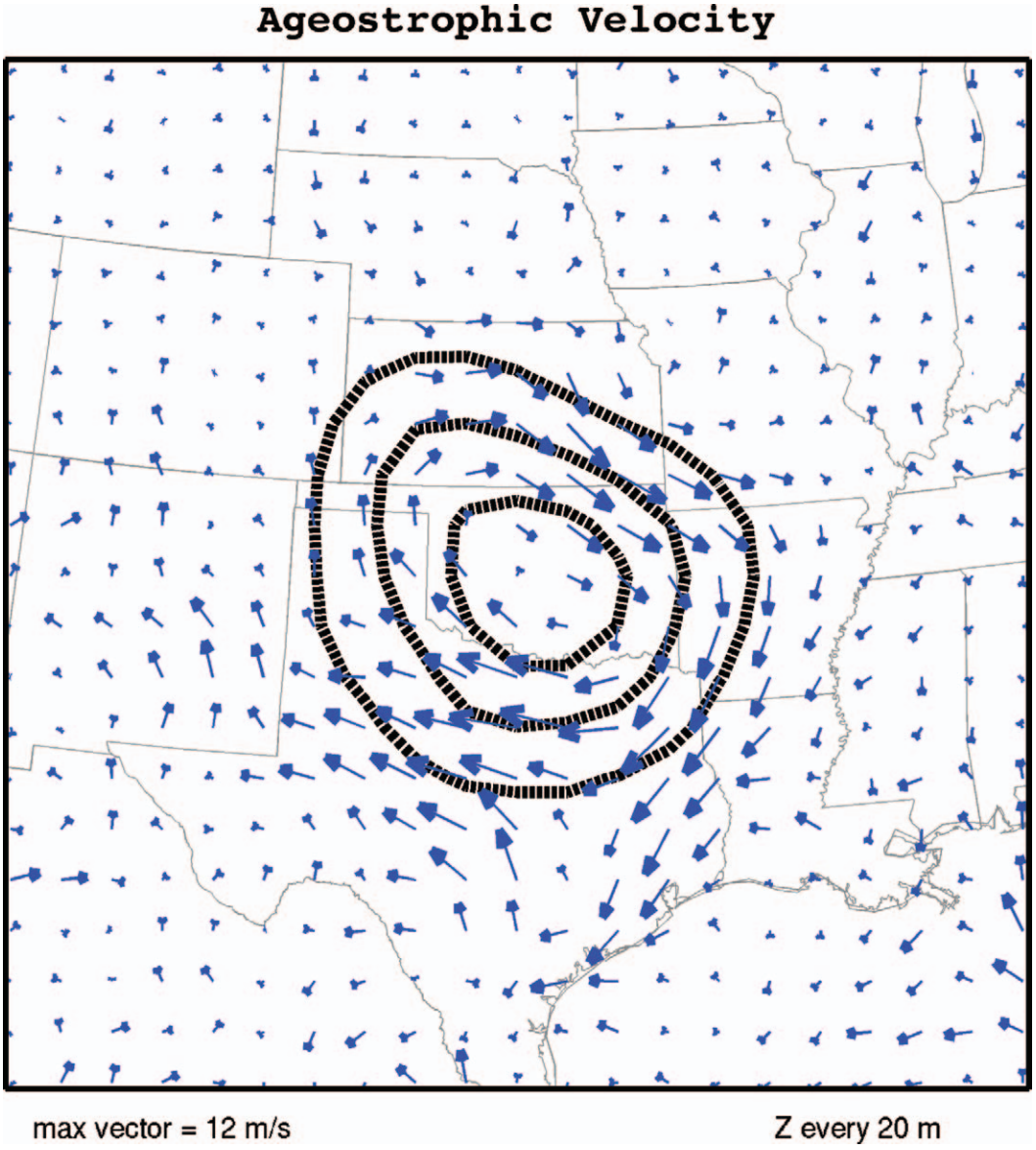

FIG. 6. Ageostrophic velocity field (vectors) for the piecewise PV inversion shown in Fig. 5. Geopotential height is given by dashed lines every 20 m. Note the area of low geopotential is associated with anticyclonic ageostrophic circulation; the total wind around the low is subgeostrophic.

is no ambiguity in the solution near the surface. This is an attractive property given the strong boundary condition influence on piecewise PV inversion (Hakim et al. 1996).

These two examples suggest that ESA PV inversion is a potentially useful tool. Advantages of this technique include inverting the unapproximated Ertel PV and freedom from balance assumptions, boundary conditions, map factors, etc. As such, irregular grids and boundaries are no more difficult than regular domains. Furthermore, the inversion is linear and thus superposition is rigorous and straightforward, and the implementation is very easy. The foremost disadvantage of the technique is the nonuniqueness of the inversion operator due to ensemble size, which may need to be large when considering piecewise PV inversions for anomalies extending over a large number of grid points. This problem may be overcome by larger ensembles or, more realistically, by techniques designed to increase the effective ensemble size (e.g., covariance localization) (Houtekamer and Mitchell 1998; Hamill et al. 2001).

We note that although the operator $\mathbf{L}$ has a null space because of ensemble subsampling there is a more fundamental space that may prove useful for studying, or even defining, balance dynamics and balance models. This space, the orthogonal complement of $\mathbf{L}$, consists

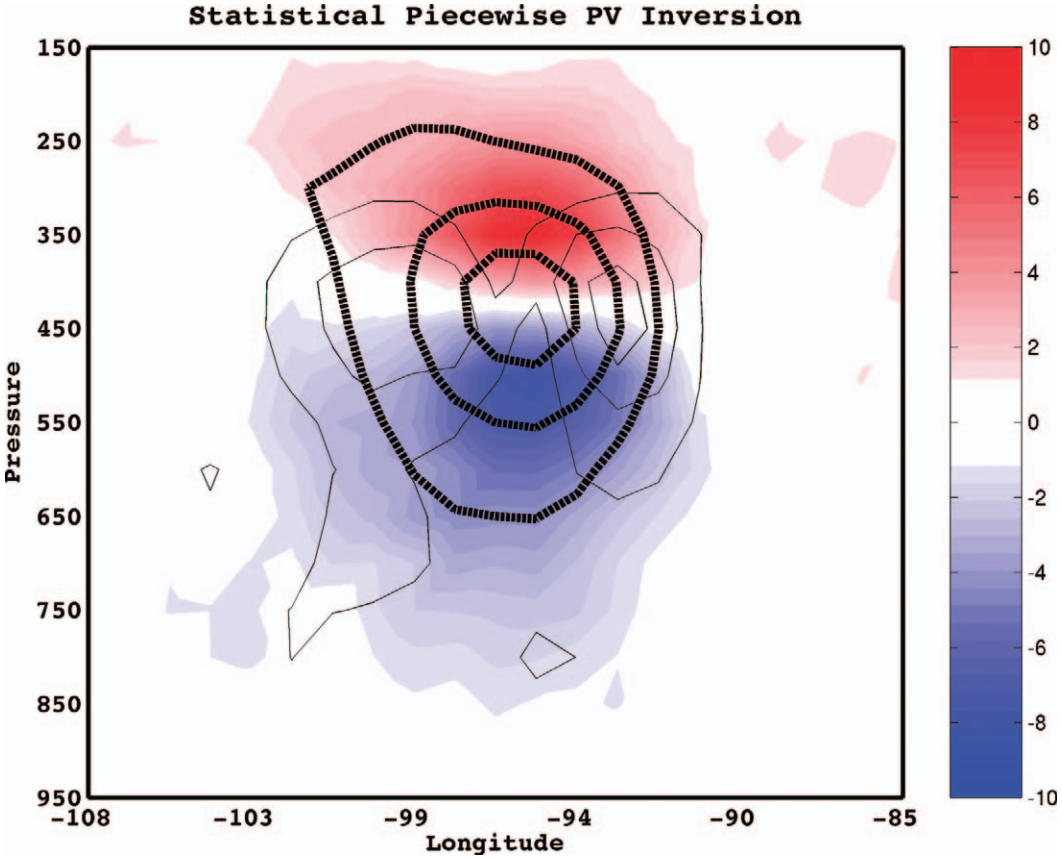

FIG. 7. East–west cross section of piecewise Ertel PV inversion for a blob of PV bounded in the vertical by the 450–550-hPa layer and in the horizontal by values greater than about 1 PVU. Thick solid black lines give the geopotential every 20 m, and thin solid black lines give the wind speed normal to the cross section every 4 m s^{-1}. Colors give the potential temperature field (°C).

of the subspace of the model state **X**, which cannot be recovered by PV inversion, even for full-rank ensembles. Consider a hydrostatic incompressible model having n grid points for model velocities u and v and thermodynamic variable T for a total of $N = 3n$ degrees of freedom. A full-rank PV matrix has at most n degrees of freedom, leaving $2n$ degrees of freedom orthogonal to the space spanned by the PV inversion operator. What is the dynamical basis for this large model subspace? In the case of linearized dynamics around a state of rest, these directions correspond to inertia–gravity waves and the n modes resolved by the PV correspond to Rossby waves. What is interesting is that the ensemble approach permits a natural extension of this separation to finite amplitude states, with the orthogonal complement corresponding to unbalanced modes linearized about the ensemble-mean state. This interpretation suggests a statistical definition for balance: Balance dynamics are defined by the state subspace that covaries with potential vorticity.

We close this section by noting that there is nothing in principle that limits ESA PV inversion to the wind and mass field as is typically the case; one could also determine, for example, the precipitation field ''attributable'' to a particular PV anomaly.

6. Statistical dynamical sensitivity analysis

The final ESA example concerns sensitivity analysis, which addresses problems such as determining how changes to an initial condition affect a subsequent forecast. This analysis is equivalent to determining the factors most important to the dynamics of a weather system. For the cyclone examined here, an application of these ideas concerns the factors that control the cyclone central pressure at 0600 UTC 29 March, which is the metric chosen for the analysis. A 24-h time interval is selected for examination, and sensitivity is defined by the covariance between the chosen metric and the 0600 UTC 28 March analysis. We shall refer to the ensemble solution as ''the control.'' Specifically, the control ensemble-mean analysis at 0600 UTC 28 March is perturbed by

$$\mathbf{x}_0' = \alpha \operatorname{cov}(\mathbf{X}_0', \mathbf{y}') = \alpha \mathbf{X}_0' \mathbf{y}'^{\mathrm{T}}, \tag{17}$$

where $\mathbf{x}_0'$ is the perturbation state vector at 0600 UTC

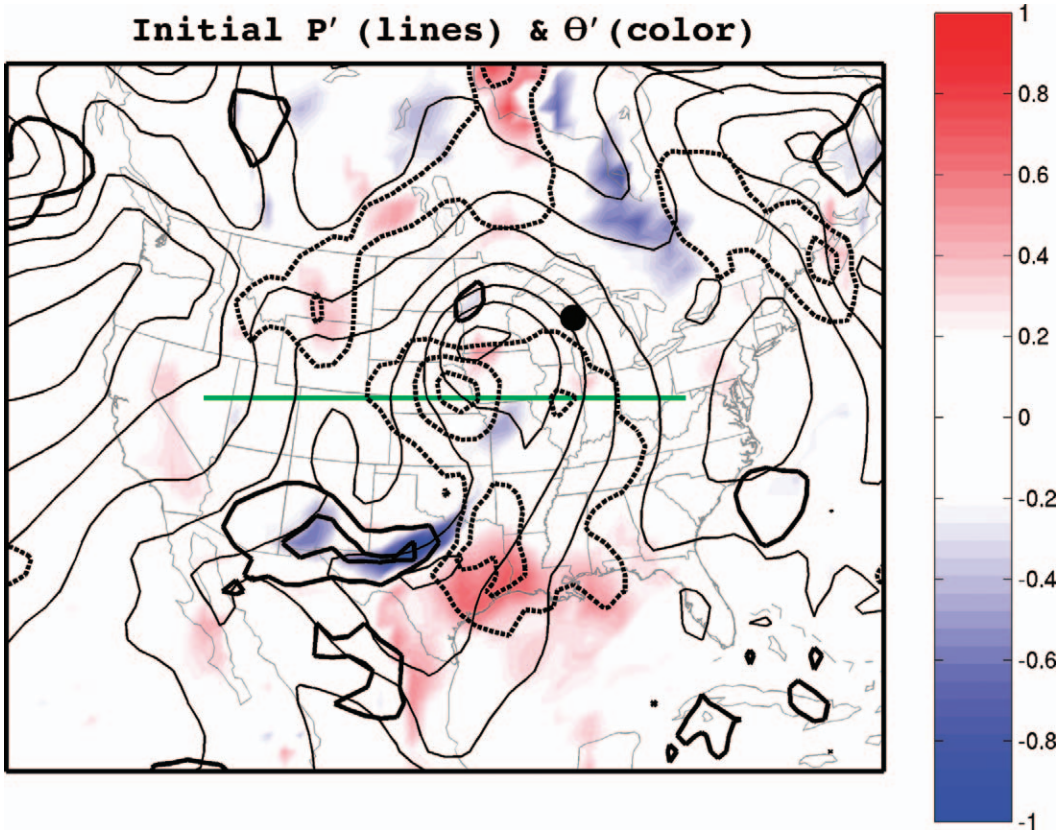

FIG. 8. Sensitivity fields at 0600 UTC 28 Mar pertaining to the cyclone central pressure at 0600 UTC 29 Mar as determined by ESA for $\alpha = -2.5$; the predicted change in the cyclone central pressure at 0600 UTC 29 Mar is -6 hPa. Thick black lines show the perturbation surface pressure field every 0.25 hPa, and the colors show the surface potential temperature perturbations. Thin solid lines show the control ensemble-mean surface pressure at 0600 UTC 28 Mar every 4 hPa.

28 March ($N \times 1$), $\mathbf{X}_0'$ is the control ensemble state matrix at 0600 UTC 28 March ($N \times M$; ensemble mean removed), and $\mathbf{y}'$ is the normalized control ensemble sample of cyclone central pressure at 0600 UTC 29 March ($1 \times M$; mean removed). As in prior calculations [see (14)], normalization of the central pressure is by the ensemble standard deviation. Parameter α controls the amplitude of the initial perturbations.

One possible factor that may be expected to affect the 0600 UTC 29 March cyclone central pressure is the cyclone central pressure at 0600 UTC 28 March; that is, a deeper initial cyclone may be expected to produce a deeper ending cyclone. In fact, the results show that the positioning of the cyclone relative to other flow features is more important than the initial amplitude (Fig. 8), with the pressure covariance field indicating a displacement of the low center farther south. An east–west cross section through this feature exhibits a nearly barotropic structure, with essentially no vertical tilt and largest amplitude at the surface and the tropopause (Fig. 9). This structure stands in contrast to adjoint sensitivity analyses, which often give structures that are highly tilted in the vertical (e.g., Langland et al. 1995). A second disturbance is found farther to the south, near a weak frontal wave in Texas. The surface covariances indicate that a displacement of this feature northeastward along the front is an important factor for the intensity of the parent cyclone. Moreover, the surface temperature covariance field indicates that a stronger cross-front temperature gradient near the frontal wave is also linked to the intensity of the main surface low 24 h later. Overall, this particular case leaves the impression that changes in the phase relationships of existing features are the most important control on the central pressure of the surface cyclone.

With $\alpha = -2.2$, a perturbed ensemble-mean initial condition defined by (17) is evolved in the full WRF model. The WRF solution for this initial condition gives a cyclone central pressure that is 4.6 hPa lower than in the control case at 0600 UTC 29 March, as compared with 6 hPa lower as predicted by the ensemble covariance (Fig. 10). A more stringent test of the method derives from the following covariance considerations at 0600 UTC 29 March. The linearized ensemble dynamics obey

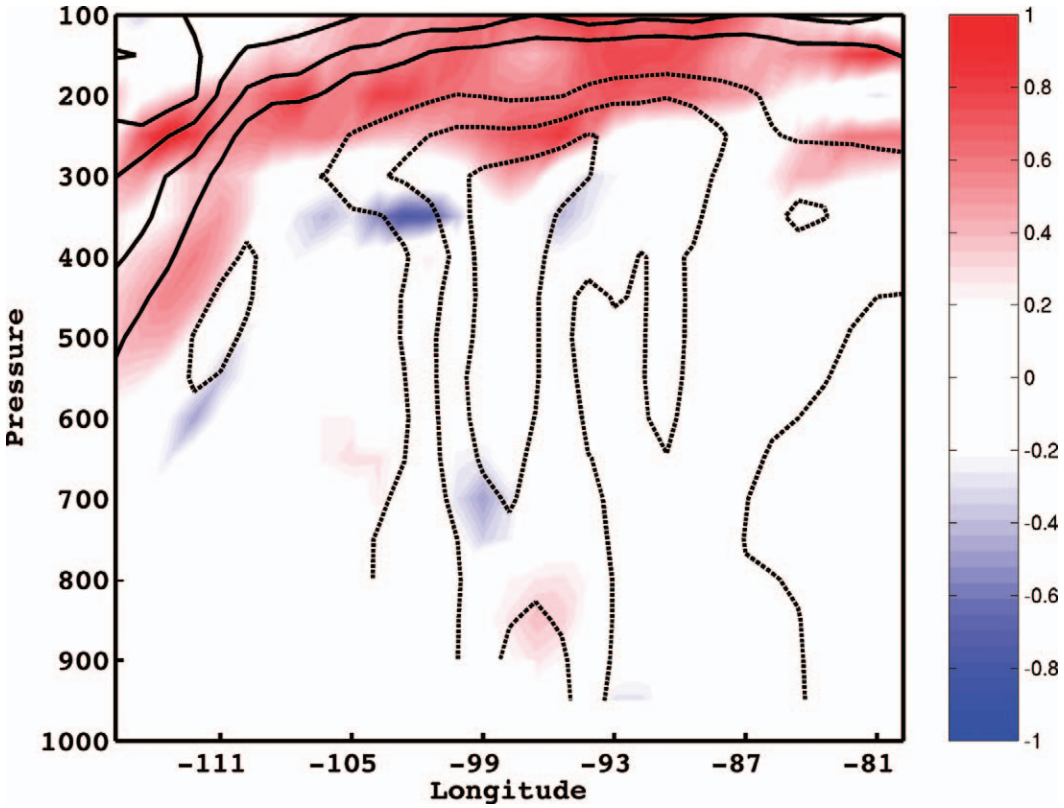

FIG. 9. Cross section near 40°N through the sensitivity geopotential height field (black lines every 2 m) and potential temperature field (colors).

$$\mathbf{X}'_t = \mathbf{W}\mathbf{X}'_0, \tag{18}$$

where $\mathbf{X}'_t$ is the ensemble control solution at 0600 UTC 29 March, $\mathbf{X}'_0$ is the ensemble control initial condition, and $\mathbf{W}$ represents the tangent-linear WRF model.[4] Right multiplying (18) by $\mathbf{y}^{\mathrm{T}}$ gives

$$\mathbf{X}'_t\mathbf{y}'^{\mathrm{T}} = \mathbf{W}\mathbf{X}'_0\mathbf{y}'^{\mathrm{T}}. \tag{19}$$

Using (17), we find that

$$\mathbf{x}'_t = \mathbf{W}\mathbf{x}'_0 = \alpha\ \mathrm{cov}(\mathbf{X}'_t, \mathbf{y}'). \tag{20}$$

Equation (20) says that the perturbed solution at 0600 UTC 29 March given perturbed initial condition (17) is equal to the control ensemble covariance, scaled by α. Thus, unlike adjoint sensitivity analysis, ESA sensitivity analysis predicts the changes in both the chosen metric and in the full solution fields. Comparing Figs. 10 and 2 we find very similar fields, as predicted by (20).

Finally, the above perturbation procedure is repeated for other values of α to determine the range of linear dynamics for this event in the chosen metric. There is a good linear relationship between predictions based on the covariance calculation and WRF solutions over the range -8 to 8 hPa in perturbation cyclone central pressure (Fig. 11, light blue dashed line). Note, however, that the linear relationship falls below the main diagonal, which indicates that the WRF response in the chosen metric is systematically smaller than the covariance prediction. One possible explanation for this result is that the ensemble has too much variance in the estimate of the forecast cyclone central pressure. Renormalizing the ensemble sample of cyclone central pressure values [i.e., $\mathbf{y}'$ in (17)] with this correction yields a very good linear response (Fig. 11, red dashed line). This plot also clearly indicates that linear perturbation dynamics are violated when the initial perturbations are scaled to produce deviations larger than about 8 hPa in cyclone central pressure at 0600 UTC 29 March. Note also the asymmetry between positive and negative perturbations, which suggests that the cyclone can be deepened by at most ~10 hPa, whereas it may be weakened by nearly 15 hPa.

7. Summary

ESA is proposed as a new tool for investigating atmospheric phenomena that is based upon probabilistic analyses, which are estimated from an ensemble. Unlike deterministic analyses, the probabilistic samples dis-

[4] Note that $\mathbf{W}$ is not known in matrix form but rather only as a computer program. Nevertheless, the action of the model may be represented symbolically as in (18). Furthermore, we do not have the tangent linear model for WRF; however, for small-amplitude perturbations, the difference between two nonlinear runs provides a good approximation to the true tangent linear model.

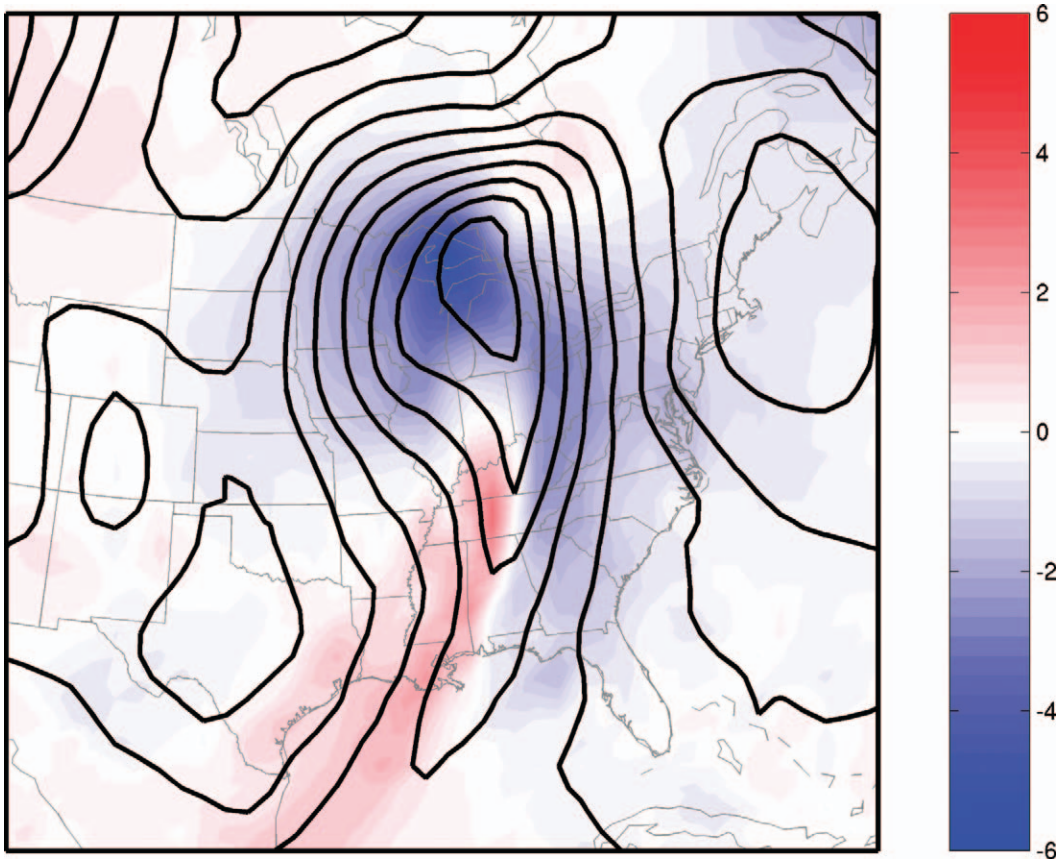

FIG. 10. Difference surface pressure field at 0600 UTC 29 Mar relative to the control for the perturbed initial condition for $\alpha = -2.5$. Pressure perturbations (colors, hPa) are defined as the difference between the perturbed solution at 0600 UTC 29 Mar, as determined by integration of the perturbed initial condition in the WRF model, and the control solution at 0600 UTC 29 Mar. The surface pressure field for the perturbed solution is given by thin solid lines every 4 hPa.

cussed here admit, and quantify, the inherent errors in atmospheric analyses. These ensemble analyses also seamlessly integrate with ensemble prediction, since it is not necessary to create analysis perturbations. Our goal here has been to propose analysis methods that utilize the wealth of information in these probabilistic analyses and, in doing so, to illustrate that state estimation may be used as a tool for atmospheric dynamics. The analysis ensemble is determined by a research-based ensemble Kalman filter consisting of 100 members and is applied to a case of cyclogenesis.

The ensemble covariance relationships qualitatively confirm known linkages between surface cyclones and upper-level disturbances. The results also suggest perhaps less obvious relationships between the cyclone and the surface cold front, a second upper-level shortwave, and the subtropical jet stream. These novel insights underscore the potential for this technique to provide useful new perspectives into phenomena that are less well understood than extratropical cyclones.

An application of ESA to potential vorticity inversion suggests the viability of this approach, which is appealing given the short algorithm, and freedom from traditional assumptions, such as balance relationships. Furthermore, the analysis motivates a new definition for balance dynamics in terms of the state subspace that covaries with potential vorticity.

A second ESA example addressed dynamical sensitivity analysis and factors that control the intensity of the surface cyclone. Unlike adjoint sensitivity analysis, this technique does not require an adjoint model and is very easy to perform numerically; both this calculation and the PV inversion calculation are easily performed with short software programs. Moreover, the method implicitly accounts for errors in the initial and end states, and for a given initial perturbation, it predicts the solution response in both the chosen metric and the full solution field. Continuing this comparison, the ESA results for this case show that the initial condition sensitivity is highly organized around existing flow features, and the perturbation fields exhibit little vertical tilt. Adjoint-based sensitivity analyses often indicate widespread regions of sensitivity and complicated vertical and horizontal structures (e.g., Langland et al.

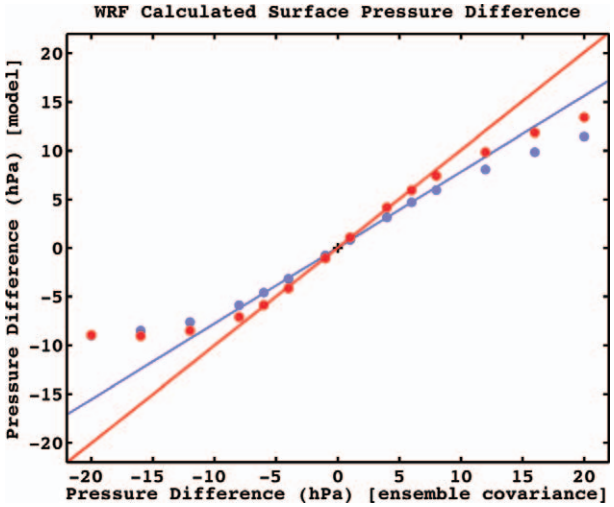

FIG. 11. Cyclone central pressure differences as determined by perturbed integrations of the WRF model (ordinate) against those predicted by the control ensemble covariance statistics (abscissa). The dashed blue line applies to the ensemble covariance perturbations and the dashed red line to the renormalized ensemble covariance perturbations (see text for details). Values along the main diagonal indicate agreement between the covariance predictions and solutions from WRF integrations.

1995; Zou et al. 1998). We caution that this comparison is based on one case, and further research will be needed to compare these two techniques.

Results for ESA sensitivity analysis show that the primary factors affecting the strength of the surface cyclone involve a phase shift of existing flow features, including a frontal wave well removed from the primary cyclone. These features were determined by the covariance between the chosen metric (surface cyclone central pressure) and the full model state 24 h earlier. To check these covariance predictions, the full WRF model was used to integrate covariance-perturbed initial conditions. The solutions show a remarkably good correspondence with the covariance predictions of the cyclone central pressure.

Although the example illustrations given here all apply to a case of cyclogenesis, that choice was mainly based on familiarity and recognition of Fred Sanders' contributions to our understand of this phenomenon. In fact, ESA may be applied to any problem where ensemble samples of data are available.

Acknowledgments. We thank Tom Hamill and an anonymous reviewer for their helpful comments on an earlier version of the manuscript. This research was supported by National Science Foundation Grants ITR-0205648, ATM-0228804, and CMG-0327658 awarded to the University of Washington.

REFERENCES

Daley, R., 1993: *Atmospheric Data Analysis.* Cambridge University Press, 471 pp.

Davis, C. A., and K. A. Emanuel, 1991: Potential vorticity diagnosis of cyclogenesis. *Mon. Wea. Rev.,* **119,** 1929–1953.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

Golub, G. H., and C. F. Van Loan, 1996: *Matrix Computations.* Johns Hopkins University Press, 698 pp.

Hakim, G. J., D. Keyser, and L. F. Bosart, 1996: The Ohio Valley wave-merger cyclogenesis event of 25–26 January 1978. Part II: Diagnosis using quasigeostrophic potential vorticity inversion. *Mon. Wea. Rev.,* **124,** 2176–2205.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation: A tutorial. *Predictability of Weather and Climate,* T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 702 pp.

——, and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.,* **128,** 2905–2919.

——, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790.

Holton, J. R., 2004: *An Introduction to Dynamic Meteorology.* Elsevier Science and Technology Books, 560 pp.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322–2339.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer and turbulence closure schemes. *Mon. Wea. Rev.,* **122,** 927–945.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 364 pp.

Langland, R. H., R. L. Elsberry, and R. M. Errico, 1995: Evaluation of physical processes in an idealized extratropical cyclone using adjoint sensitivity. *Quart. J. Roy. Meteor. Soc.,* **121,** 1349–1386.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.,* **102,** 409–418.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2001: Development of a next generation regional weather research and forecast model. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology,* W. Zwieflhofer and N. Kreitz, Eds., World Scientific, 269–276.

Sanders, F., 1986: Explosive cyclogenesis in the west-central North Atlantic Ocean, 1981–84. Part I: Composite structure and mean behavior. *Mon. Wea. Rev.,* **114,** 1781–1794.

——, 1987: Study of 500-mb vorticity maxima crossing the east coast of North America and associated surface cyclogenesis. *Wea. Forecasting,* **2,** 70–83.

——, and J. R. Gyakum, 1980: Synoptic-dynamic climatology of the "bomb." *Mon. Wea. Rev.,* **108,** 1589–1606.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913–1924.

——, G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.,* **132,** 1190–1200.

Zou, X., Y.-H. Kuo, and S. Low-Nam, 1998: Medium-range prediction of an extratropical oceanic cyclone: Impact of initial state. *Mon. Wea. Rev.,* **126,** 2737–2763.