JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 101, NO. D5, PAGES 9471–9478, APRIL 27, 1996

An objective determination of the polar vortex using Ertel's potential vorticity

Eric R. Nash,[1] Paul A. Newman,[2] Joan E. Rosenfield,[3] and Mark R. Schoeberl[2]

Abstract. We have developed objective criteria for choosing the location of the northern hemisphere polar vortex boundary region and the onset and breakup dates of the vortex. By determining the distribution of Ertel's potential vorticity (Epv) on equivalent latitudes, we define the vortex edge as the location of maximum gradient of Epv constrained by the location of the maximum wind jet calculated along Epv isolines. We define the vortex boundary region to be at the local maximum convex and concave curvature in the Epv distribution surrounding the edge. We have determined that the onset and breakup dates of the vortex on the 450 K isentropic surface occur when the maximum wind speed calculated along Epv isolines rises above and falls below approximately 15.2 m s^{-1}. We use 1992–1993 as a test case to study the onset and breakup periods, and we find that the increase of polar vortex Epv values is associated with the dominance of the term in the potential vorticity equation involving the movement of air through the surface due to the diabatic circulation. We also find that the decrease is associated with the dominance of the term involving radiatively induced changes in the stability of the atmosphere.

Introduction

The polar vortex is the dominant dynamical feature of the stratospheric winter circulation in both northern and southern hemispheres. A typical vortex is marked by strong circumpolar winds and cold temperatures in the interior of this vortex region. As the circumpolar winds begin to increase in the autumn, the polar vortex consequently spins up, and the values of Ertel's potential vorticity (Epv) rise at the pole. These large polar Epv values form a relatively coherent air mass, and this air mass is commonly referred to as the polar vortex. The polar vortex is conventionally described using global maps of Epv on isentropic surfaces. The mapping of Epv has become one of the primary diagnostic tools for the analysis of stratospheric transport and dynamic processes.

Sampling of the polar vortex air in both hemispheres by high-altitude aircraft and satellite instruments has conclusively demonstrated that the chemical composition is substantially different from midlatitude air outside of the vortex [*Leovy et al.*, 1985; *Proffitt et al.*, 1989]. For example, Microwave Limb Sounder (MLS) data from the Upper Atmosphere Research Satellite (UARS) have shown enhanced values of ClO contained within the interior of the polar vortex [*Waters et al.*, 1993]. Aircraft and satellite measurements of long-lived stratospheric trace gases show substantially lower values inside the polar vortex [*Loewenstein et al.*, 1989; *Russell et al.*, 1993]. Also, gradients of long-lived constituent tracers are highly correlated with the high gradients of Epv at the vortex edge [*Schoeberl et al.*, 1992; *Loewenstein et al.*, 1990].

Irreversible transport of material out of the polar vortex can be easily observed using Epv as a diagnostic. *McIntyre and Palmer* [1983] stimulated this use of Epv by demonstrating its utility in their discussions of planetary wave breaking. They showed that the winter Epv distribution had high values in the polar region (the main vortex), with a broad region of weak Epv gradient surrounding the main vortex (the surf zone). Since Epv is generally conserved on an isentropic surface over weekly time periods, this separation into a main vortex and surf zone was noted as a material separation, and the breaking of planetary waves was observed to cause irreversible separation of material from the vortex [*McIntyre and Palmer*, 1984; *Clough et al.*, 1985; *Waugh et al.*, 1994]. It is now recognized that irreversible transport may also occasionally move material into the vortex under certain conditions of persistent tropospheric blocking [*Plumb et al.*, 1994]. This irreversible transport can also lead to a reduction in Epv as tongues of material are pulled off the vortex and stretched until the area of the tongue can no longer be resolved by the observational networks [*McIntyre and Palmer*, 1983].

Radiative effects also have a large impact on the development of the stratospheric polar vortex. *O'Neill and Pope* [1993a, b] showed that radiation can be a prime contributor to the development of the polar vortex using the analysis of circulation integrals over constant Epv values. They were able to show, using a numerical simulation of the 1980–1981 winter season, that the vortex in the middle stratosphere developed through the radiative term in the circulation budget and that the decay of the vortex was also largely driven by this radiative term.

Because of the importance of the polar vortex, precise knowledge of the vortex boundary region and the vortex evolution is extremely important. The vortex boundary region was originally defined as the area between the main vortex and the surf zone where the Epv gradient attained large values. *Butchart and Remsberg* [1986] used Epv analyses to show that the short-timescale changes in area of the vortex on the 850 K isentropic surface were related to planetary wave breaking. These changes were superimposed upon the slower changes of Epv related to the diabatic heating. Various modeling studies

[1]Applied Research Corporation, Landover, Maryland.
[2]NASA Goddard Space Flight Center, Greenbelt, Maryland.
[3]General Sciences Corporation, Laurel, Maryland.

Paper number 96JD00066.
0148-0227/96/96JD-00066$05.00

have shown that the polar vortex is relatively isolated from the midlatitudes [*Rood et al.*, 1992; *Waugh et al.*, 1994]. Further, *Bowman* [1993], *Pierce and Fairlie* [1993], and *Manney et al.* [1994b] have shown that the vortex boundary region acts as a barrier to transport.

McIntyre and Palmer [1984] and *Butchart and Remsberg* [1986] first presented a way of looking at the vortex using the area under Epv isolines. More recently, *Manney et al.* [1994c] have used this viewpoint to study the polar vortex of 1992–1993. In this paper we use this same viewpoint combined with several analyses techniques to provide a comprehensive description of the gross features of the northern hemisphere polar vortex. We define objective techniques for determining the extent of the vortex region using the gradients of Epv. We also define a method to determine the dates of vortex formation and dissipation. Using these definitions, we will focus our attention on the lower stratospheric northern hemisphere winter of 1992–1993 as a case study of the vortex spinup and breakdown. We also analyze the Epv budget of the polar vortex and show that the physics of the lower stratospheric vortex are in good agreement with the middle stratospheric results of *O'Neill and Pope* [1993a, b].

Data

The National Meteorological Center's (NMC) Climate Analysis Center provided temperature and geopotential height data on pressure surfaces as 65 × 65 polar stereographic grids. These data are a combination of radiosondes and satellite retrievals from the TIROS operational vertical sounder (TOVS) aboard the National Oceanic and Atmospheric Administration's polar orbiter series [*Randel*, 1987; *Nagatani et al.*, 1988, 1990]. At 100 hPa and below the analyses are from the Global Data Acquisition System (GDAS) [*McPherson et al.*, 1979]. Above 100 hPa the analyses are calculated using a scheme based on the successive corrections method [*Cressman*, 1959]. The data quality has been analyzed by *Trenberth and Olson* [1988] for the GDAS data and by *Gelman et al.* [1986] for the stratospheric data. The analyses were of good quality for the period that we considered, and we have interpolated over any obviously poor data. We interpolated all of the data used for this study to the 450 K isentropic surface. This level corresponds to pressures between 100 hPa and 50 hPa.

Ertel's potential vorticity Q is defined as

$$Q = \rho^{-1}(\nabla \times \mathbf{u} + 2\boldsymbol{\Omega}) \cdot \nabla\theta,$$

where ρ is the density, $\mathbf{u}$ is the three-dimensional velocity field, $\boldsymbol{\Omega}$ is the Earth's angular velocity, and θ is potential temperature. We calculated Epv in the form

$$Q = -g(\zeta_p + f)\frac{\partial\theta}{\partial p}, \tag{1}$$

where g is the acceleration due to gravity, p is pressure, ζ_p is the vertical component of the relative vorticity given by

$$\zeta_p = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \tag{2}$$

f is the Coriolis parameter given by

$$f = 2|\boldsymbol{\Omega}| \sin\phi,$$

ϕ is latitude, and u and v are the zonal and meridional wind components along the x and y Cartesian coordinates.

The potential vorticity equation on an isentropic surface is given by

$$\frac{\partial Q}{\partial t} + \mathbf{v} \cdot \nabla_\theta Q = \frac{Q}{\sigma}\frac{\partial}{\partial\theta}(\sigma H) - \sigma^{-1}\mathbf{k} \cdot \nabla_\theta \times \left(H\frac{\partial \mathbf{v}}{\partial\theta}\right) + F, \tag{3}$$

where t is time, $\mathbf{v}$ is the horizontal velocity field, H is the diabatic heating rate ($= D\theta/Dt$), F is the frictional forcing, and σ is defined as

$$\sigma \equiv -g^{-1}(\partial p/\partial\theta).$$

Our use of Epv (Q) is somewhat analogous to the mixing ratio of a chemical substance. The "PV-substance," which is analogous to the corresponding chemical substance, has no source or sink in a layer bounded by two isentropic surfaces and cannot be transported across these surfaces. However, Q does change within a layer because mass can be transported into or out of the layer. For a general discussion of the conservation of Q and the quantity "σQ," see *Haynes and McIntyre* [1987, 1990], *McIntyre* [1992], and references cited therein.

From the geopotential heights we calculated balanced winds and relative vorticities on the 65 × 65 grid [*Randel*, 1987; *Newman et al.*, 1988]. We then interpolated the temperatures, winds, and relative vorticities to a 5° longitude by 2° latitude grid. This method produced more realistic polar winds and vorticity fields than if the heights were first interpolated to the latitude-longitude grid. We calculated $\partial\theta/\partial p$ using log pressure for the vertical differencing.

We also used data from an assimilation model produced by the Data Assimilation Office at NASA Goddard Space Flight Center [*Rood et al.*, 1990]. This model (GEOS1) used an optimal interpolation scheme that advects information from data poor to data rich regions. The derived fields were constrained to be consistent with each other. Excessive inertial gravity waves were avoided by using the incremental analysis update procedure which gradually inserted data over a 6-hour period as forcing terms. The data were provided to us on a 5° longitude by 4° latitude grid.

Methodology

Determining the Vortex Boundary Region and Vortex Edge

The boundary between polar vortex and midlatitude air is indicated by the position of the strong horizontal gradient of Epv on an isentropic surface [*McIntyre*, 1982; *McIntyre and Palmer*, 1983, 1984]. This area defines the polar vortex boundary region with high Epv values found poleward and a region of smaller Epv and gradients (the surf zone) in midlatitudes. *Plumb et al.* [1994] have shown that weak mixing into the vortex can occur under conditions of persistent tropospheric blocking. However, in general, most recent studies [e.g., *Juckes and McIntyre*, 1987; *Bowman*, 1993; *Pierce and Fairlie*, 1993; *Manney et al.*, 1994b] have shown that there is a kinematic barrier to transport of material into the polar vortex, effectively isolating the vortex from midlatitude material.

Determining the vortex boundary region can be difficult. In plotting the Epv gradient against the Epv values for each grid point on the latitude-longitude grid, there is a scattering of points with a peak value that can be determined by fitting Gaussian, polynomials, and similar functions to the data and finding the maximum. This technique is highly susceptible to

noise, analysis errors, and the choice of functions to fit the data.

We view the Epv data in an alternative way, as first presented by *McIntyre and Palmer* [1984] and *Butchart and Remsberg* [1986]. We began by contouring the Epv values on an equivalent area map projection. Then we calculated the area enclosed by each isoline. Expressed as a fraction of the total hemispheric surface area, we rearranged these areas on the hemisphere so that they ranged from 1 at the equator to 0 at the pole. The transformed Epv isolines can be thought of as being symmetrically arranged around the pole and monotonically increasing from the equator to the pole. The latitudes on this transformed map projection are referred to as "equivalent latitudes."

Plotting Epv against equivalent latitude results in a distribution that has an "S" shape in the vortex region. Our method for determining the extent of the boundary region is based on both this shape and the averaged wind speed along Epv isolines. Since the wind is not single-valued around the vortex edge, we averaged the wind along Epv isolines to determine the location of the jet maximum relative to the area under the Epv isolines. We define the vortex edge to be the location of the highest Epv gradient, constrained by the proximity of a reasonably strong westerly jet. We note that from (1) and (2), Epv is proportional to the negative meridional gradient of zonal wind, and thus the gradient of Epv with equivalent latitude is proportional to the negative second derivative of zonal wind with respect to equivalent latitude. Therefore the broad-scale jet is a useful constraint on the location of the vortex edge, since the wind maximum is strongly correlated with the Epv gradient maximum. However, while the wind maximum is a surrogate for the vortex edge, it cannot serve as a precise definition because the quasi-conservative Epv also contains the $\partial\theta/\partial p$ term and the wind maximum is typically a few degrees equatorward of the vortex edge.

There are often multiple peaks in the distributions of the Epv gradient and the wind versus equivalent latitude, in addition to the strong peak at the vortex edge. Although some of these other peaks may also be large, they usually appear only in one distribution but not in both. However, at the true vortex edge, both distributions have strong peaks. Therefore we multiplied together the values of the peaks in the Epv distribution with associated nearby peaks in the wind distribution and chose the largest resultant value in the Epv distribution as the true vortex peak. We also applied a sloping filter that ranges between 1 at 80° and 0 at 90° equivalent latitude. There are few points that add to these small areas, and slight variations in the number of points can produce these large gradients.

We define the vortex boundary region to be located between the local maximum convex and concave curvature in the Epv curve that surrounds the vortex edge. As an example, Figure 1a shows a plot of the NMC Epv versus equivalent latitude for the northern hemisphere on January 1, 1993. Figure 1b shows that the peak gradient is located at 68°, and this is where we determined the vortex edge to be located. The value of Epv at this edge is 2.9×10^{-5} K m^2 kg^{-1} s^{-1} (indicated by the dashed line). There was a corresponding peak in the wind distribution that confirmed the peak in the Epv gradient (Figure 1c). The second derivative of Epv (not shown) had a local minimum and maximum at 66° and 70° with associated values of 2.6 and 3.2×10^{-5} K m^2 kg^{-1} s^{-1}, respectively (indicated by the dotted lines.)

This method considerably reduces the noise and analysis

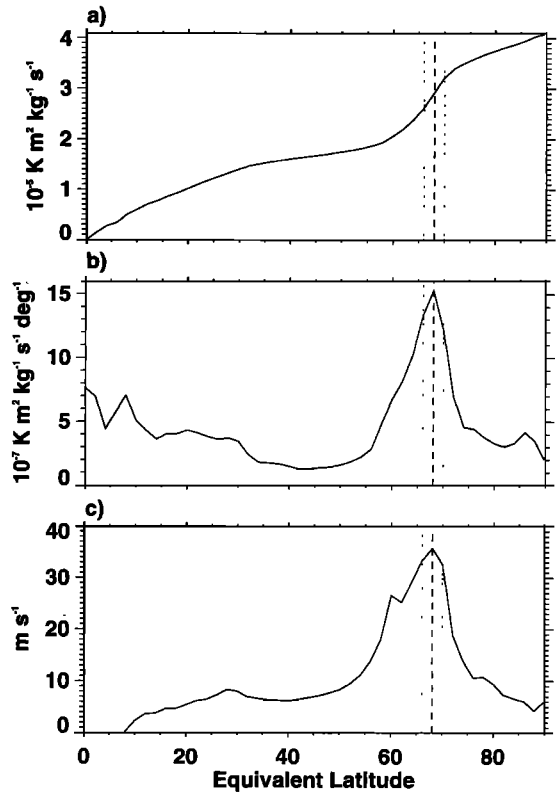

Figure 1. (a) Ertel's potential vorticity (Epv), (b) the gradient of Epv, and (c) the average wind speed along Epv isolines versus equivalent latitude for the 450 K isentropic surface in the northern hemisphere on January 1, 1993. The dashed line indicates the calculated vortex edge, and the dotted lines indicate the calculated boundary region.

errors and completely eliminates the need for choosing functional fits. However, there is a choice to be made in the resolution of the equivalent latitudes. Choosing the steps for equivalent latitude is the same as choosing steps of Epv isolines. We have found that choosing equivalent latitudes with the same latitude spacing as the original data provides the best estimates for the boundary calculations. Choosing a larger interval (less latitude bands or Epv isolines) broadens the derivatives so much that an accurate placement of the vortex cannot be made. Choosing smaller intervals leads to "ringing" in the derivatives, where local differences in the distribution of Epv versus equivalent latitude give large local differences in the derivatives.

Other studies have used a particular value of Epv throughout the season to define the vortex edge (see, for example, *Manney et al.* [1994c]). We note that the maximum gradient of Epv changes as the vortex develops and dissipates and that our technique is more sensitive to these changes. Also, our definition of a vortex boundary region better allows for studies of constituent transport through this finite area.

We have utilized the vortex boundary calculations to analyze the development and breakdown of the polar vortex on the 450 K isentropic surface. In the case of the summer circulation, the winds in the stratosphere are generally easterly. Hence averaging of the wind component parallel to the Epv isolines yields easterlies that ill define an edge. As the wind reverses direction during the autumn, the wind average on Epv isolines also

reverses direction, and a weak Epv gradient maximum becomes evident.

Determining a date for the formation and breakup of the vortex is a relatively subjective process. *Manney et al.* [1994c] determined the dates to be when the equivalent latitude under a chosen Epv contour was greater then 80°. Our formation and breakup dates were determined for most of the years of data by using winds and Epv. For this date determination, we considered changes in (1) the Epv gradient, (2) the distribution of winds around Epv isolines, (3) the maximum winds near the vortex edge, and (4) the area of the vortex. We found that the maximum wind speed integrated around Epv isolines was in best agreement with the subjectively determined dates. We calculated the average of these maximum wind speeds to be 15.2 m s^{-1} with a standard deviation of 3.2 m s^{-1} for both the formation and breakdown dates. We used the 15.2 m s^{-1} value as the threshold for determining the existence of the vortex on the 450 K surface.

The vortex builds up slowly in late autumn. Using the standard deviation range for the wind speeds provides a useful range of dates during which the vortex is formed. As the vortex decreases in strength between winter and spring, the winds along the vortex edge also decrease, and the vortex usually dissipates quickly. The range of dates associated with wind speed range is much smaller than that in the buildup period. Clearly this relationship of the threshold wind to the strength of the Epv gradient is a consequence of the buildup of westerly momentum and kinetic energy of the flow field during the autumn and of the decrease to easterlies during the spring.

Determining the Change in Area

The evolution of the polar vortex can be analyzed by understanding the physical processes that control the growth, maintenance, and decay of the area of the vortex. Plate 1 shows a typical example of the evolution of the polar vortex over the course of the winter. In midautumn (October 20, 1992) Epv values are low, the Epv gradient between mid-latitudes and the polar region is small, and the areal coverage (indicated by the yellow to red colors) is also small. By late January the Epv values and areal coverage are quite large, and the gradient is very strong. The areal coverage is a useful indicator of the polar vortex temporal behavior.

The area as a function of time, $A(t)$, enclosed by a given Epv isoline Q_*, will remain constant in the absence of diabatic heating, divergent winds, or frictional forces. Changes in the area provide a diagnostic tool to study the vortex behavior. Using (3), the time rate of change of the area $A(t)$ is given by

$$\frac{d}{dt}A(t)_{\hat{Q}\geq Q_*} \approx \oint_{\hat{Q}=Q_*} Q\frac{\partial H}{\partial \theta}\frac{ds}{|\nabla_\theta \hat{Q}|} - \oint_{\hat{Q}=Q_*} H\frac{\partial \hat{Q}}{\partial \theta}\frac{ds}{|\nabla_\theta \hat{Q}|} + \int_{A(t)\hat{Q}\geq Q_*} \nabla_\theta \cdot \hat{\mathbf{v}}\, dA - \oint_{\hat{Q}=Q_*} \mathbf{v}\cdot\nabla_\theta Q'\frac{ds}{|\nabla_\theta \hat{Q}|} + \oint_{\hat{Q}=Q_*} F\frac{ds}{|\nabla_\theta \hat{Q}|} \quad (4)$$

where ds is a length element along an isoline, the circumflex denotes the observed values, and the prime denotes the deviation of the observed values from the true values (generally spatial scales below the resolution of the analyses) [*Butchart and Remsberg*, 1986]. Line integrals are taken at a particular Epv isoline ($\hat{Q} = Q_*$), and the area integral is taken over all values of Epv greater than or equal to the particular isoline ($\hat{Q} \geq Q_*$). We calculated the heating rates from the radiation model of *Rosenfield et al.* [1987], updated following *Rosenfield* [1991].

The first term on the right-hand side represents changes in Epv associated with diabatically induced changes in the stability. It is usually negative at this level around the northern hemisphere vortex region, since Epv is positive and the heating rate decreases (cooling is larger) with increasing potential temperature, thus reducing the stability. This term acts to dissipate the polar vortex and can be conceptually understood as a vorticity dilution term, where diabatic processes act to compress the material and hence spin down the Epv.

The second term on the right-hand side represents changes in Epv from the movement of air through the surface due to the diabatic circulation. In the northern hemisphere polar vortex, Epv increases with height, so that radiative cooling then leads to increases in Epv and the area enclosed by Epv isolines.

The third term on the right-hand side denotes the change in area due to the divergence of the wind field. We could not accurately determine its direct contribution since the calculation of this term was very unstable. We instead used a high-resolution trajectory model to determine how the area would change in the absence of the diabatic terms.

The fourth term denotes unresolvable contributions from subgrid scale processes. These processes can act in either direction to expand or shrink the vortex. Epv isolines become stretched into long thin filaments that wrap around the vortex during the breaking of Rossby waves [*Polvani and Plumb*, 1992; *Waugh*, 1993a, b]. This appears to usually be a one-way process in which air is stripped from the vortex into the surf zone. The large-scale satellite observations cannot resolve these fine-scale features and may represent a smearing of the Epv over a region of alternating high and low Epv values or may miss these features completely. Smearing would result in a widening of the horizontal Epv gradient in the areas of the fine-scale features, while missing a feature would, at most, be reflected in a slight change in the local Epv values.

The final term involves the contribution of frictional forces. We note that gravity waves can have a large impact on both the heating rate and momentum balance. The effect of this frictional drag would be to reduce the Epv gradients and shrink the vortex area.

There are other terms which involve horizontal derivatives of the heating rate and vertical winds. We estimated from the assimilation and NMC data that these terms were at least 2 orders of magnitude smaller than the terms given in (4), and so we have ignored them for this study.

Examples and Discussion

We used the winter vortex of 1992–1993 as a case study to show the applicability of calculating the vortex boundary region and edge, and to diagnose the processes responsible for the development and decay of the polar vortex. Plate 1 shows six different Epv maps on key dates in the season. The map for October 20 shows the state of the atmosphere before formation of the polar vortex. The map for November 18 indicates the conditions during early vortex formation. The map on January 21 shows a well-formed vortex in midseason. The maps on February 16 and March 20 point out the deformation of the

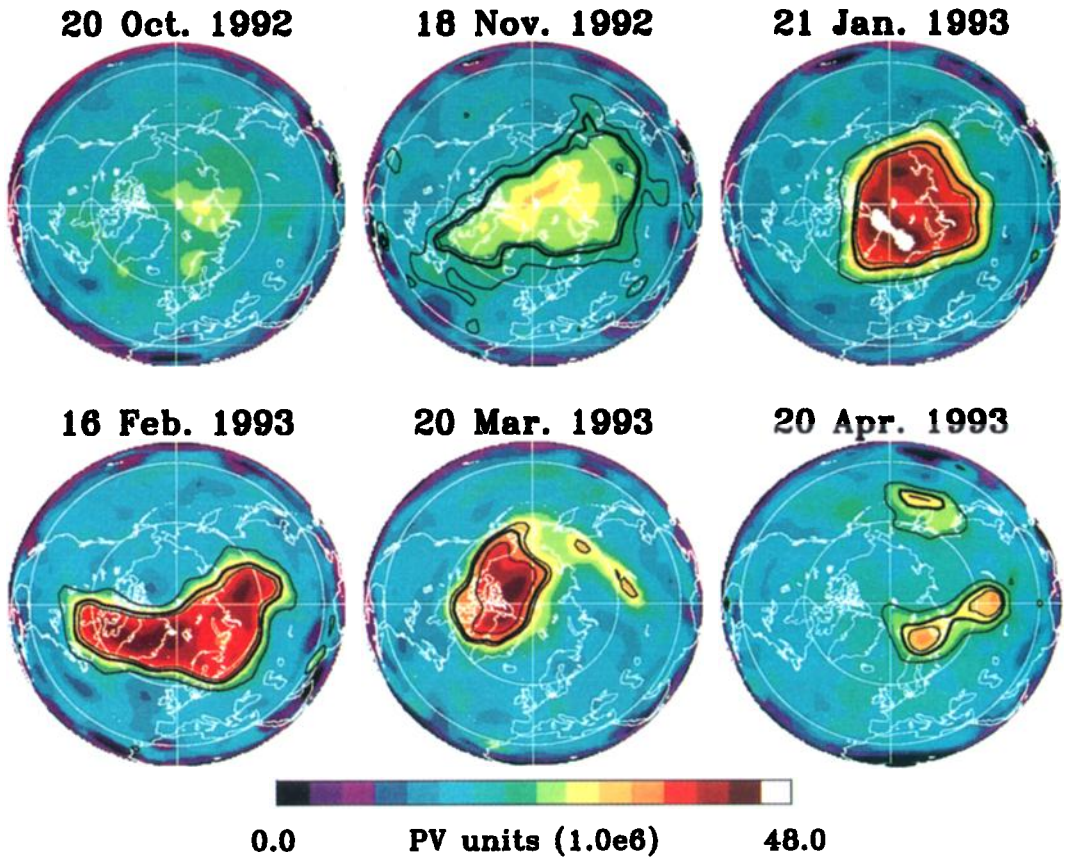

Plate 1. Northern hemisphere Ertel's potential vorticity on the 450 K isentropic surface. The Greenwich meridian points downward. The heavy black line is the computed vortex edge, and the thin black lines are the computed vortex boundary region. Units are 10^{-6} K m^2 kg^{-1} s^{-1}.

vortex during two warming events. Finally, the April 20 map illustrates the conditions at the vortex breakup. *Manney et al.* [1994a, c] describe the 1992–1993 winter polar vortex using data from the Upper Atmosphere Research Satellite (UARS) and Epv derived from NMC data. Our findings are in general agreement with their descriptions.

Formation of the Winter 1992–1993 Vortex

From a visual inspection of Epv charts, we can clearly see a well-formed vortex in late November, whereas the vortex is relatively indistinct in early November. Figure 2 shows the distribution of maximum winds along Epv isolines over the winter. There is a daily scatter about a steadily increasing jet maximum and then a falloff after mid-January. Using 5-day smoothed values, we determined that the vortex formed on November 10, when the wind maximum reached more than 15.2 m s^{-1}. Applying our standard deviation limits, we note that the wind reached 12.0 m s^{-1} on October 27 and 18.4 m s^{-1} on November 18, a range of 23 days.

We found slightly different results with the assimilation data. The wind maximum reached 15.2 m s^{-1} on November 18, and the 12 m s^{-1} and 18.4 m s^{-1} limits were reached on November 7 and December 2, respectively, a range of 26 days. The assimilation winds were consistently lower then the NMC winds during this period and resulted in the later dates of formation. The late part of the NMC range and early part of the assimilation range overlapped. Given that the vortex was slow to form, these ranges are in reasonable agreement.

In the early part of November the calculated vortex edge was located around 76° (Figure 3) equivalent latitude. In mid-November, calculated edge values dropped to around 64°. The more northerly values in early November were a result of local peaks in the Epv gradient and wind maximum fields. As the vortex began to spin up in mid-November, the Epv gradients

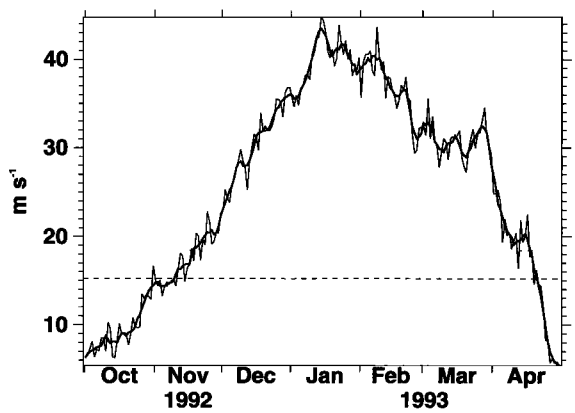

Figure 2. Maximum of average winds along Epv isolines in the northern hemisphere on the 450 K isentropic surface (thin line). The heavy line is a 5-day running mean of the wind maximums. The dashed line is at 15.2 m s^{-1}, which is the determining value for the onset and breakdown of the polar vortex. Dotted lines are at the 1 standard deviation limits of 12.0 and 18.4 m s^{-1}.

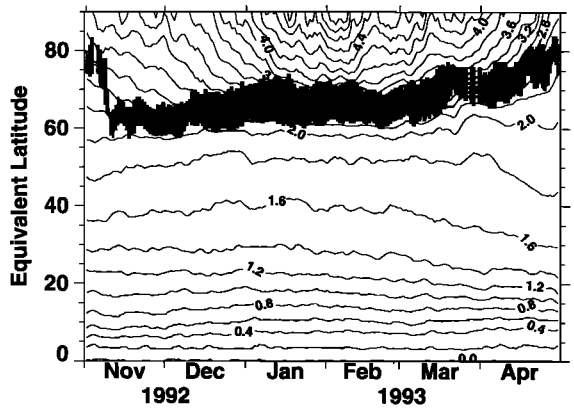

Figure 3. Isolines of Ertel's potential vorticity over equivalent latitude and time on the 450 K isentropic surface in the northern hemisphere. The heavy line is the calculated vortex edge, and the shaded area is the calculated vortex boundary region. Isolines are plotted every 0.2×10^{-5} K m^2 kg^{-1} s^{-1}. The data shown were smoothed in time with a 5-day running mean.

and wind maximum increased, and we began to see the peaks in these fields at the vortex location of about 64°.

The area enclosed by the calculated vortex increased slightly between mid-December and mid-February (see also *Manney et al.* [1994c]). However, there was a steady rise in the Epv value at the vortex edge from 2.2 to 3.2×10^{-5} K m^2 kg^{-1} s^{-1}. The gradient across the vortex edge tripled by mid-January from the value at formation in mid-November.

The development of the polar vortex is mainly a result of the movement of air through the surface due to the diabatic circulation. To show this, we calculated the change in area expected from the first two terms on the right-hand side of (4) enclosed by the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} isoline (Figure 4). These two terms together are the contribution from diabatic processes. This isoline first appeared in late October at 90° and then showed a rapid increase in area until early December. It was near the vortex edge from early January throughout the rest of the season. During the formation period, both diabatic terms increased in magnitude (negative and positive in sign, respectively), but the second term dominated and caused the rapid increase in the area enclosed by the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} isoline. There is good agreement between the observed change in the area under the isoline and the area calculated from the diabatic heating terms during this period. In early December the area tendency of the isoline started to decrease, and the rate of increase of Epv inside the vortex slowed down. These results are similar to those found by *O'Neill and Pope* [1993a, b], whose data were at a higher potential temperature level for the 1980–1981 season.

The Midwinter 1992–1993 Vortex

During the 1992–1993 winter season there were three warming events, although, since the zonal wind never reversed below 10 hPa these could not be classified as major warmings. The first event was dominated by wave 1 and occurred in mid-February. On the 450 K isentropic surface, the temperature difference between the pole and midlatitudes decreased in magnitude but did not reverse. This remained true up through the 30 hPa level. At 10 hPa there was a reversal to a warm pole and this remained so through the rest of the season. On the 450 K surface the maximum polar warming occurred on February 16 (bottom left pane of Plate 1). The warming occurred over the western Pacific and displaced the vortex to the opposite hemisphere. During this time the area of the vortex stayed nearly constant, however, inner vortex Epv values decreased rapidly from 5.2 to 4.6×10^{-5} K m^2 kg^{-1} s^{-1} (Figure 3).

The second warming event was a minor warming in which the Aleutian anticyclone intensified over Canada and displaced the vortex off the pole toward the northern Russian coast. The maximum warming occurred on March 6 (not shown), and there was a reversal to a warm pole on the 450 K surface and above. The total area of the vortex and the maximum values inside the vortex began to decrease before this event, but during the peak of the warming these values increased slightly, propagating from 90° to to the polar boundary region over a few days.

In late March there was a dramatic change in the vortex position. In a reversal from the second event, the vortex slid across to the Canadian side of the hemisphere. The peak warming occurred on March 20 (bottom center pane of Plate 1). During this event the winds increased, while the pole warmed. The area of the vortex decreased rapidly, but the highest Epv values inside the vortex increased slightly. This event was wave 1 dominated in geopotential height, but wave 2 was nearly as large in the temperatures.

In early February the area tendency from the diabatic circulation was negative and by late February the tendency from the NMC data was also negative (Figure 4). Both diabatic terms decreased in absolute magnitude during this time, but the dissipation term began to become stronger, which resulted in the shrinkage of the vortex area.

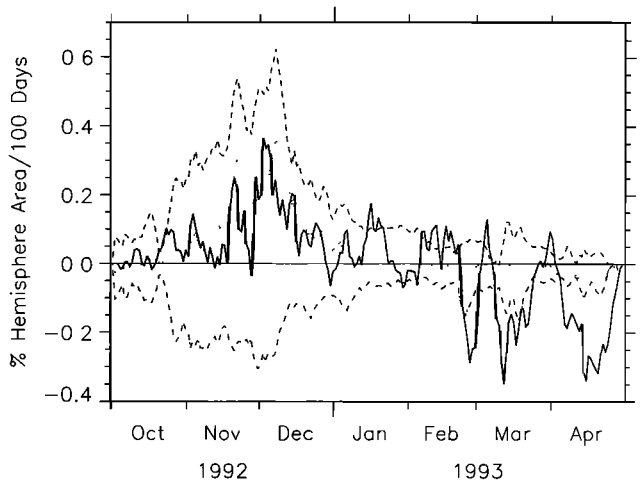

Figure 4. Area tendency on the 450 K isentropic surface in the northern hemisphere for the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} isoline of Ertel's potential vorticity. The solid line is the calculated tendency from the NMC data. These data were smoothed with a 7-day running mean. The dashed line is the component of the tendency associated with the diabatically induced movement of air through the isentropic layer, and the dashed-dotted line is the component associated with diabatically induced changes in the stability. The dotted line is the sum of these two terms and is the total time tendency calculated from the diabatic circulation. See the text for definitions of these terms.

Breakup of the Spring 1993 Vortex

The final warming on the 450 K isentropic surface started on April 5, when the 80°–50° latitude temperature difference became positive for remainder of the season. The zonal winds decreased from a mid-March maximum to a minimum on April 24, although the winds were still slightly westerly.

The vortex broke up on April 20, 1993 (bottom right pane of Plate 1). Normally, as the vortex breaks down, it fragments into two or more separate pieces of high Epv material. Visual inspection of Epv maps in early April showed an intact coherent vortex, whereas maps in late April showed a weak fragmented Epv field. Figure 2 shows the steadily declining values of the wind maximum around Epv isolines. The wind maximum went below 15.2 m s^{-1} on April 20, and this corresponded well with our visual inspection of Epv plots. Again, as a reference, we note that the wind maximum fell below 18.4 m s^{-1} on April 17 and below 12 m s^{-1} on April 22, a range of 6 days. This is a much tighter range than for the formation period.

The assimilation data showed the vortex breakup on April 17. The 18.4 m s^{-1} and 12 m s^{-1} limits were reached on April 5 and April 21, respectively, a range of 17 days. Just as in the build up period, the assimilation wind speeds were less than the NMC wind speeds, so the break up date was earlier. There was an overlap of the entire NMC range with the later part of the assimilation range. Again, the data are in reasonable agreement.

During April the area contained by Epv isolines rapidly decreased. The calculated diabatic terms were small and the resulting expected area tendency of the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} isoline was much less than the observed tendency by a factor of 3 to 6 (Figure 4). These results are similar to those of *O'Neill and Pope* [1993a, b], who found that in the middle stratosphere the radiation terms accounted for about half of the decrease in area during major vortex shrinkage events. They proposed that the vertical smoothing of satellite data may account for much of the difference. We made estimates of the limits of possible changes due to the diabatic heating terms by running the radiation model with various sonde, aircraft, and satellite data profiles. We found that the differences in the diabatic heating provided, at most, a change of 0.05 fraction of hemispheric area per 100 days. This is still too small by a factor of 2 to 3 to explain the decrease in Epv area.

We used a high-resolution trajectory model [*Schoeberl et al.*, 1993a, b] to calculate the evolution of the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} Epv isoline through April. The results showed that the Epv field was not sheared into unresolvable scales. The vortex area contained by this adiabatic advection of the 2.8×10^{-5} K m^2 kg^{-1} s^{-1} isoline remained relatively stable over the entire period. This implies that the divergence of the wind field was small and the Epv was not sheared out into scales of motion that were unresolvable by the synoptic network. This differs from the results found by *Juckes and O'Neill* [1988], who estimated that half of the decrease they observed was due to this effect. However, their data were for the winter of 1982–1983 on the 850 K surface.

Gravity waves may have a large effect on Epv amounts. There were indications that significant activity took place on the southern boundary of the vortex during April (J. T. Bacmeister, personal communication, 1994). The effect of this activity would be to reduce the area of the vortex, and this may have accounted for a significant portion of the area decrease which we found.

Summary

We have presented an objective method for determining the location of the vortex boundary region and vortex edge using the gradient of Ertel's potential vorticity and the wind maximum around Epv isolines. The wind maximum is nearly colocated with the maximum Epv gradient on equivalent latitudes. Using the wind maximum to constrain the location of the Epv gradient peaks allows us to unambiguously determine the polar vortex edge.

We have shown that the formation and breakup dates of the vortex can be determined from the coarse grid of the NMC data. We have determined that on the 450 K isentropic surface, the polar vortex forms when the wind maximum rises above 15.2 m s^{-1} in the autumn and that breakup occurs when the wind maximum drops below 15.2 m s^{-1} in the spring. During the winter of 1992–1993, we estimate that the lower stratospheric vortex formed on November 10, 1992, and broke down on April 20, 1993. The NMC data set should be useful enough to produce a climatology of formation and ending dates for the period of 1979 to the present.

We have shown the importance of diabatic heating in the formation, expansion, and shrinkage of the polar vortex. From the potential vorticity equation (3) we find that the autumn build up of the lower stratospheric vortex is primarily associated with the movement of air through the surface due to the diabatic circulation in the polar region. As the polar region becomes coldest in December and January, the polar lower stratosphere radiative relaxation timescales become very long, and diabatic heating is reduced from the late autumn maximum. Hence the movement of air through the layer is also reduced, and the vortex stops growing. In late winter and spring the dissipation term (diabatically induced changes in the stability) becomes larger and is associated with the shrinking of the vortex. During the period close to breakup, all of the calculated terms are too small to account for the change in area of the vortex. Frictional dissipation of the vortex by gravity waves may account for much of this disparity.

Acknowledgments. We would like to thank J. Bacmeister for advising us of the possibility of strong gravity wave activity during the breakup period. We would also like to thank an anonymous reviewer for helpful comments on the manuscript.

References

Bowman, K. P., Large-scale isentropic mixing properties of the Antarctic polar vortex from analyzed winds, *J. Geophys. Res.*, *98*, 23,013–23,027, 1993.

Butchart, N., and E. E. Remsberg, The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface, *J. Atmos. Sci.*, *43*, 1319–1339, 1986.

Clough, S. A., N. S. Grahame, and A. O'Neill, Potential vorticity in the stratosphere derived using data from satellites, *Q. J. R. Meteorol. Soc.*, *111*, 335–358, 1985.

Cressman, G. P., An operational objective analysis system, *Mon. Weather Rev.*, *87*, 367–374, 1959.

Gelman, M. E., A. J. Miller, K. W. Johnson, and R. M. Nagatani, Detection of long-term trends in global stratospheric temperature from NMC analyses derived from NOAA satellite data, *Adv. Space Res.*, *6*(10), 17–26, 1986.

Haynes, P. H., and M. E. McIntyre, On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces, *J. Atmos. Sci.*, *44*, 828–841, 1987.

Haynes, P. H., and M. E. McIntyre, On the conservation and impermeability theorems for potential vorticity, *J. Atmos. Sci.*, *47*, 2021–2031, 1990.

Juckes, M., and M. McIntyre, A high resolution, one-layer model of breaking planetary waves in the stratosphere, *Nature*, *328*, 590–596, 1987.

Juckes, M. N., and A. O'Neill, Early winter in the northern stratosphere, *Q. J. R. Meteorol. Soc.*, *114*, 1111–1125, 1988.

Leovy, C. B., C.-R. Sun, M. H. Hitchman, E. E. Remsberg, J. M. Russell, L. L. Gordley, J. C. Gille, and L. V. Lyjak, Transport of ozone in the middle stratosphere: Evidence for planetary wave breaking, *J. Atmos. Sci.*, *42*, 230–244, 1985.

Loewenstein, M., J. R. Podolske, K. R. Chan, and S. E. Strahan, Nitrous oxide as a dynamical tracer in the 1987 Airborne Antarctic Ozone Experiment, *J. Geophys. Res.*, *94*, 11,589–11,598, 1989.

Loewenstein, M., J. R. Podolske, K. R. Chan, and S. E. Strahan, N_2O as a dynamical tracer in the Arctic vortex, *Geophys. Res. Lett.*, *17*, 477–480, 1990.

Manney, G. L., R. W. Zurek, A. O'Neill, R. Swinbank, J. B. Kumer, J. L. Mergenthaler, and A. E. Roche, Stratospheric warmings during February and March 1993, *Geophys. Res. Lett.*, *21*, 813–816, 1994a.

Manney, G. L., R. W. Zurek, A. O'Neill, and R. Swinbank, On the motion of air through the stratospheric polar vortex, *J. Atmos. Sci.*, *51*, 2973–2994, 1994b.

Manney, G. L., R. W. Zurek, M. E. Gelman, A. J. Miller, and R. Nagatani, The anomalous Arctic lower stratospheric polar vortex of 1992–1993, *Geophys. Res. Lett.*, *21*, 2405–2408, 1994c.

McIntyre, M. E., How well do we understand the dynamics of stratospheric warmings?, *J. Meteorol. Soc. Jpn.*, *60*, 37–65, 1982.

McIntyre, M. E., Atmospheric dynamics: Some fundamentals, with observational implications, in *The Use of EOS for Studies of Atmospheric Physics*, edited by J. Gille and G. Visconti, pp. 313–386, North-Holland, New York, 1992.

McIntyre, M. E., and T. N. Palmer, Breaking planetary waves in the stratosphere, *Nature*, *305*, 593–600, 1983.

McIntyre, M. E., and T. N. Palmer, The 'surf zone' in the stratosphere, *J. Atmos. Terr. Phys.*, *46*, 825–849, 1984.

McPherson, R. D., K. H. Bergman, R. E. Kistler, G. E. Rasch, and D. S. Gordon, The NMC operational global data assimilation system, *Mon. Weather Rev.*, *107*, 1445–1461, 1979.

Nagatani, R. M., A. J. Miller, M. E. Gelman, and P. A. Newman, A comparison of Arctic lower stratospheric winter temperatures for 1988–1989 with temperatures since 1964, *NOAA Tech. Rep. NWS 40*, 125 pp., Natl. Weather Serv., Washington, D. C., 1988.

Nagatani, R. M., A. J. Miller, K. W. Johnson, and M. E. Gelman, An eight-year climatology of meteorological and SBUV ozone data, *NOAA Tech. Rep. 17*, 333 pp., Natl. Oceanic and Atmos. Admin., Silver Spring, Md., 1990.

Newman, P. A., L. R. Lait, and M. R. Schoeberl, The morphology and meteorology of southern hemisphere spring total ozone mini-holes, *Geophys. Res. Lett.*, *15*, 923–926, 1988.

O'Neill, A., and V. D. Pope, The coupling between radiation and dynamics in the stratosphere, *Adv. Space Res.*, *13*(1), 351–358, 1993a.

O'Neill, A., and V. D. Pope, The effects of resolution on diagnostic calculations of coupling between radiation and dynamics, *Adv. Space Res.*, *13*(1), 359–362, 1993b.

Pierce, R. B., and T. D. A. Fairlie, Chaotic advection in the stratosphere: Implications for the dispersal of chemically perturbed air from the polar vortex, *J. Geophys. Res.*, *98*, 18,589–18,595, 1993.

Plumb, R. A., D. W. Waugh, R. J. Atkinson, P. A. Newman, L. R. Lait, M. R. Schoeberl, E. V. Browell, A. J. Simmons, and M. Loewenstein, Intrusions into the lower stratospheric Arctic vortex during the winter of 1991–1992, *J. Geophys. Res.*, *99*, 1089–1105, 1994.

Polvani, L. M., and R. A. Plumb, Rossby wave breaking, microbreaking, filamentation, and secondary vortex formation: The dynamics of a perturbed vortex, *J. Atmos. Sci.*, *49*, 462–476, 1992.

Proffitt, M. H., et al., A chemical definition of the boundary of the Antarctic ozone hole, *J. Geophys. Res.*, *94*, 11,437–11,448, 1989.

Randel, W. J., The evaluation of winds from geopotential height data in the stratosphere, *J. Atmos. Sci.*, *44*, 3097–3120, 1987.

Rood, R. B., P. A. Newman, L. R. Lait, D. J. Lamich, and K. R. Chan, Stratospheric temperatures during AASE: Results from STRATAN, *Geophys. Res. Lett.*, *17*, 337–340, 1990.

Rood, R. B., R. S. Stolarski, A. R. Douglass, J. A. Kaye, and D. J. Allen, Episodic total ozone minima and associated effects on heterogeneous chemistry and lower stratospheric transport, *J. Geophys. Res.*, *97*, 7979–7996, 1992.

Rosenfield, J. E., A simple parameterization of ozone infrared absorption for atmospheric heating rate calculations, *J. Geophys. Res.*, *96*, 9065–9074, 1991.

Rosenfield, J. E., M. R. Schoeberl, and M. A. Geller, A computation of the stratospheric diabatic circulation using an accurate radiative transfer model, *J. Atmos. Sci.*, *44*, 859–876, 1987.

Russell, J. M., III, A. F. Tuck, L. L. Gordley, J. H. Park, S. R. Drayson, J. E. Harries, R. J. Cicerone, and P. J. Crutzen, HALOE Antarctic observations in the spring of 1991, *Geophys. Res. Lett.*, *20*, 719–722, 1993.

Schoeberl, M. R., L. R. Lait, P. A. Newman, and J. E. Rosenfield, The structure of the polar vortex, *J. Geophys. Res.*, *97*, 7859–7882, 1992.

Schoeberl, M. R., S. D. Doiron, L. R. Lait, P. A. Newman, and A. J. Krueger, A simulation of the Cerro Hudson SO_2 cloud, *J. Geophys. Res.*, *98*, 2949–2955, 1993a.

Schoeberl, M. R., et al., The evolution of ClO and NO along air parcel trajectories, *Geophys. Res. Lett.*, *20*, 2511–2514, 1993b.

Trenberth, K. E., and J. G. Olson, Evaluation of NMC global analyses: 1979–1987, *Tech. Note NCAR/TN-299+STR*, 82 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., 1988.

Waters, J. W., L. Froidevaux, W. G. Read, G. L. Manney, L. S. Elson, D. A. Flower, R. F. Jarnot, and R. S. Harwood, Stratospheric ClO and ozone from the Microwave Limb Sounder on the Upper Atmosphere Research Satellite, *Nature*, *362*, 597–602, 1993.

Waugh, D. W., Contour surgery simulations of a forced polar vortex, *J. Atmos. Sci.*, *50*, 714–730, 1993a.

Waugh, D. W., Subtropical stratospheric mixing linked to disturbances in the polar vortices, *Nature*, *365*, 535–537, 1993b.

Waugh, D. W., R. A. Plumb, R. J. Atkinson, M. R. Schoeberl, L. R. Lait, P. A. Newman, M. Loewenstein, D. W. Toohey, and C. R. Webster, Transport out of the lower stratospheric Arctic vortex by Rossby wave breaking, *J. Geophys. Res.*, *99*, 1071–1088, 1994.

E. R. Nash, P. A. Newman, J. E. Rosenfield, and M. R. Schoeberl, Mail Code 916, NASA Goddard Space Flight Center, Greenbelt, MD 20771. (e-mail: nash@triffid.gsfc.nasa.gov)

(Received April 28, 1994; revised December 15, 1995; accepted December 15, 1995.)