A New Look at Stratospheric Sudden Warmings. Part III: Polar Vortex Evolution and Vertical Structure

N. J. MATTHEWMAN AND J. G. ESLER

Department of Mathematics, University College London, London, United Kingdom

A. J. CHARLTON-PEREZ

Department of Meteorology, University of Reading, Reading, United Kingdom

L. M. POLVANI

Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York

ABSTRACT

The evolution of the Arctic polar vortex during observed major midwinter stratospheric sudden warmings (SSWs) is investigated for the period 1957–2002, using 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) Ertel's potential vorticity (PV) and temperature fields. Time-lag composites of vertically weighted PV, calculated relative to the SSW onset time, are derived for both vortex-displacement SSWs and vortex-splitting SSWs, by averaging over the 15 recorded displacement and 13 splitting events. The evolving vertical structure of the polar vortex during a typical SSW of each type is clearly illustrated by plotting an isosurface of the composite PV field, and is shown to be very close to that observed during representative individual events. Results are verified by comparison with an elliptical diagnostic vortex moment technique.

For both types of SSW, little variation is found between individual events in the orientation of the developing vortex relative to the underlying topography; that is, the location of the vortex during SSWs of each type is largely fixed in relation to the earth's surface. During each type of SSW, the vortex is found to have a distinctive vertical structure. Vortex-splitting events are typically barotropic, with the vortex split occurring near simultaneously over a large altitude range (20–40 km). In the majority of cases, of the two daughter vortices formed, it is the "Siberian" vortex that dominates over its "Canadian" counterpart. In contrast, displacement events are characterized by a very clear baroclinic structure; the vortex tilts significantly westward with height, so that the top and bottom of the vortex are separated by nearly 180° longitude before the upper vortex is sheared away and destroyed.

1. Introduction

The current study is the third in a series investigating the observed dynamical behavior of major midwinter stratospheric sudden warmings (SSWs). Observational benchmarks derived from the studies in Charlton and Polvani (2007, hereafter Part I) and Charlton et al. (2007, hereafter Part II) and the current work will serve to characterize typical dynamical behavior in the polar stratosphere during both vortex-displacement and vortex-splitting SSW events. Part I of the series was primarily concerned with the identification and classification of all SSWs in the period 1957–2002, and on the development of diagnostic benchmarks. In particular, in Part I, the evolution of temperature, geopotential height, zonal wind, and eddy flux fields during SSW events of each type were documented using a composite method. In Part II, the diagnostic benchmarks developed in Part I were used to evaluate the performance of several

Corresponding author address: N. J. Matthewman, Department of Mathematics, University College London, 25 Gower Street, London WC1E 6BT, United Kingdom.
E-mail: jmatt@math.ucl.ac.uk

stratosphere-resolving models in simulating SSWs. Here, the aim is to extend the study further by analyzing SSWs from the perspective of three-dimensional vortex dynamics.

It is well established (e.g., McIntyre and Palmer 1983; Nash et al. 1996) that the Arctic polar vortex can be identified with an air mass of elevated Ertel's potential vorticity (PV), and that the vortex itself is strongly disrupted during an SSW (e.g., O'Neill 2003, and references therein). The aim of the current work is therefore to investigate the structure of the Arctic polar vortex during SSW events of each type, by a detailed analysis of the PV distribution in the Arctic polar region in the altitude range of 15–45 km. Following Limpasuvan et al. (2004) and Part I, time-lag composite diagnostics are used to characterize the development of the vortex leading up to and following each type of SSW. The aim is to answer the following questions:

- What is the vertical structure of the Arctic polar vortex during the vortex-displacement and vortex-splitting SSWs identified in Part I? Does each type of SSW event have a characteristic vertical structure, or does the vertical structure vary between individual events?
- What is the location and orientation of the Arctic polar vortex with respect to the earth's surface during the development of each type of SSW event? To what extent are SSW events "phase locked" to stationary tropospheric planetary waves generated by surface topography and land–sea contrast?
- How does the shape of the Arctic polar vortex develop prior to, during, and subsequent to the onset time of SSWs? Is there a characteristic life cycle associated with the development of the polar vortex during each type of SSW?

It is anticipated that our results, and the answers to the above questions, in particular, will be of assistance to modelers attempting to simulate SSWs in stratosphere-resolving general circulation models. Additionally, the results should be of use to researchers seeking to validate different dynamical theories of SSWs, as will be discussed in detail below.

A number of authors have analyzed the evolving vertical structure of the vortex during specific SSW events in both the Northern and Southern Hemisphere (e.g., Manney et al. 1994, 2005b,c). However, we are not aware of a study attempting a comprehensive analysis of the three-dimensional structure of the polar vortex during all recorded SSWs. One relevant, complementary investigation is that of Limpasuvan et al. (2004), who compiled composite diagnostics such as zonal mean winds and temperatures, and Eliassen–Palmer flux fields

during the life cycles of 48 SSW and minor warming events. Limpasuvan et al. (2004) did not, however, distinguish between vortex-displacement and vortex-splitting events, which from the perspective of vortex dynamics are fundamentally different phenomena. Nor did they consider the evolution of the vertical structure of the polar vortex itself during the SSW events. As will be seen below, the shape of the polar vortex undergoes robust and distinctive life cycles during each type of SSW, the details of which emerge when the above aspects are considered. Another relevant study is Liberato et al. (2007), who used NCEP–NCAR data to investigate the vertical structure of planetary wave energetics in relation to the variability of the Northern Hemisphere polar vortex during SSW events. In their work, composite analysis of displacement and splitting SSWs showed the dynamics of the two SSW types to be distinct. Displacement SSWs were forced by planetary wave energy anomalies with strong vertical dependence, whereas splitting SSWs were forced by planetary wave energy anomalies with predominantly barotropic structure in the vertical.

A detailed study of the climatological behavior of the Arctic and Antarctic polar vortices has been carried out by Waugh (1997), who introduced the methodology of "elliptical diagnostics" in order to produce time series describing the main aspects of the evolution of the Arctic and Antarctic polar vortex structure during their respective winter seasons. These time series were subjected to statistical analysis, resulting in reference climatologies for the position and the variability of the polar vortex centroid, aspect ratio, and orientation. Three key results of Waugh and Randel (1999, see in particular their Fig. 4) that are specifically relevant to the current study relate to the climatology of the Arctic polar vortex during the winter period (late November–March), when all of the SSWs identified in Part I occur. First, the centroid of the vortex is found to be displaced off the pole; at all levels throughout the winter it is located near 80°N (December) to 75°N (March). Second, the vortex is found to have a significant westward tilt with height; the longitude of the vortex centroid is located around 20°–30°E at low levels (450 K), near the 0° meridian in the midstratosphere (850 K), and around 30°–40°W in the upper stratosphere (1300 K). Finally, the aspect ratio of the vortex is generally in the 1.5–1.9 range, with the largest values occurring in February in the lower stratosphere. Results will be presented below with reference to this climatology.

Nevertheless, the focus in the Waugh (1997) and Waugh and Randel (1999) studies was not the three-dimensional evolution of the polar vortex during SSWs

themselves. Here, the elliptical diagnostic approach, which was adapted from techniques developed in the fluid dynamics literature to investigate two-dimensional vortex dynamics (Dritschel 1993; Legras and Dritschel 1993; Melander et al. 1986), is further adapted to study the structure of the Arctic polar vortex during SSW events of each type. One important novelty in the present paper is that we have removed the need to start by choosing the vortex edge, such as a specific isopleth of N_2O mixing ratio or of PV, before applying the elliptical diagnostics methodology. This modification allows independent results from the elliptical diagnostics method to be compared with those from a separate vortex edge-selecting algorithm. Results from the two techniques will therefore act to cross validate each other and help confirm the robustness of our results.

The remainder of the paper is structured as follows. Section 2 describes the datasets used to identify SSWs and the methodology employed to objectively define the boundary of the Arctic polar vortex and to generate the composite fields used to characterize SSWs of each type. The elliptical diagnostics are then introduced, the new aspects are described, and their utility is demonstrated. In section 3 the vertical structure, orientation, and elongation of the Arctic polar vortex are described for both vortex-splitting and vortex-displacement SSWs. Particular attention is given to the temporal evolution of these quantities relative to the SSW onset time. Both composites and representative individual events are compared and contrasted. In section 4, a list of proposed modeling benchmarks are presented and section 5 contains our conclusions.

2. Datasets and methodology

a. Dataset, SSW definition, and categorization

The methodology used to identify SSWs, and then classify them into vortex-displacement and vortex-splitting events, is described in detail in Part I. Part I uses data from both the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kistler et al. 2001) and the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) dataset (Uppala et al. 2005) to establish a climatology of SSW events between the winter seasons of 1957/58 and 2001/02. The results from the two reanalysis datasets are very similar, as expected, given that they are created using a largely common observational dataset. In the present study, the vertical structure of SSWs is investigated using Ertel's PV from the ERA-40

only, because PV is directly available at a higher vertical and temporal resolution in ERA-40 than in the NCEP–NCAR reanalysis.

Following Part I, our study is confined to major midwinter SSWs that occur during the extended boreal winter season (November–March). SSWs are defined to occur when the zonal mean zonal wind at 10 hPa and 60°N changes direction from westerly to easterly. An additional criterion—that the zonal mean zonal winds at this location recover to westerlies for at least 10 or more consecutive days following the SSW—is used to remove events that are final, as opposed to midwinter, warmings.[1] The dynamics of Northern Hemisphere final warmings have recently been discussed in detail by Black and McDaniel (2007). During the 1957–2002 period of investigation, 29 SSWs that satisfy these criteria are identified by Part I. These 29 events are then classified into vortex-displacement and vortex-splitting events, using an algorithm based on first identifying the vortex edge from the absolute vorticity field, and then comparing the size and amplitude of cyclonic vortices. Part I divides the 29 SSWs into 15 vortex-displacement events and 14 vortex-splitting events using this algorithm.

For the purposes of the current study, ERA-40's Ertel's potential vorticity and temperature fields were obtained from the British Atmospheric Data Centre (BADC; information online at http://badc.nerc.ac.uk/home/). Global fields were obtained at 1° × 1° horizontal resolution on 16 standard pressure levels between 400 and 1 hPa, with a vertical distance of approximately 2.6 km between each level. The temperature fields were used to interpolate the PV onto 36 isentropic levels in the range of 380–2850 K.

Note that Manney et al. (2005b,a) and Simmons et al. (2005) have reported an erroneous vertical oscillation in temperature in the polar regions in ERA-40. We have found that this oscillation, which is of a lower amplitude in the Arctic compared to the Antarctic and is only significant in the last few years of the period of interest, is not of sufficient amplitude to impact on our results. A direct comparison between ERA-40 and ERA operational analysis data for the February 2002 event (not shown) reveals no significant differences in any of our diagnostics.

---

[1] The Part I definition differs slightly from those in the existing literature in that it includes warming events that occur in March, and can be robustly ascertained not to be final warmings. In addition, a reviewer of this paper points out that some researchers may be uncomfortable with use of the term "midwinter warming" when discussing SSW events occurring in November, preferring to use "Canadian warming" for these events. For consistency with Part I, we have chosen to adhere to the dynamically based SSW definition given there, and have therefore included any November SSWs in Part I in our analysis.

b. Definition of the Arctic polar vortex boundary

To investigate the Arctic polar vortex during each SSW period, an objective definition of the vortex air mass is required. A natural way of defining the vortex air mass is to select a surface in three-dimensional space to define its outer boundary. The edges of the polar vortices throughout the stratosphere and lower mesosphere are identified with strong isentropic gradients in trace gas concentrations and in PV; hence, any surface defining the outer boundary of the vortex should coincide with these strong gradients where they exist. For the purposes of this study, the vortex is assumed to extend between the 400- and 1600-K isentropic surfaces (approximately 14 and 44 km) in the vertical. For each SSW identified in Part I, the vortex outer boundary within this range is selected to coincide with the region of strong PV gradient using the following method.

First, the Ertel PV q, on each isentropic surface θ, is expressed as a function of equivalent latitude ϕ_e (Butchart and Remsberg 1986) during the 20-day period surrounding each SSW. To obtain a quantitative snapshot of the vortex before the development of each SSW, $q(\phi_e, \theta)$ is considered 9 days before the SSW onset date identified in Table 1 of Part I. A time lag of 9 days was found to be sufficient for the vortex to be in a relatively undisturbed state prior to the SSW. To proceed, the equivalent latitude $\phi_m(\theta)$, corresponding to the maximum latitudinal gradient in PV, that is, max $|\partial q/\partial \phi_e|$, is identified on each isentropic level θ. At nearly all times and levels, the "vortex edge," defined by $\phi_m(\theta)$, lies in the equivalent latitude range 45°–80°N.

In theory, it might be desirable to proceed using the function $\phi_m(\theta)$ to identify the vortex edge during each SSW. In practice, however, there are numerical problems associated with discontinuities and noise in $\phi_m(\theta)$, resulting from, for example, (occasional) spurious maxima in the latitudinal gradient of PV. Additionally, such an approach requires a significant amount of information to be stored in conjunction with each SSW. Here, an alternative method, exploiting a transformation of the PV, is employed. Lait (1994) pointed out that the exponentially increasing component of PV q can be removed without affecting its conservation properties, by a transformation of the form

$$Q = q \left(\frac{\theta}{\theta_0} \right)^\alpha, \quad (1)$$

where θ_0 and α are constants. Müller and Günther (2003) point out that variation in the vertical temperature structure of the polar stratosphere is sufficiently large to warrant different values of α being used for different events. Hence, we aim to choose α so that a single isosurface of the vertically weighted PV, $Q = Q^*$, gives a good approximation to the vortex edge as defined by the equivalent latitude $\phi = \phi_m(\theta)$ on all isentropic surfaces. Without loss of generality, we can take $\theta_0 = 475$ K for all events. For each SSW, a particular choice of α and Q^* that leads to a good fit to the vortex edge is obtained by setting

$$Q^* = q[\phi_m(\theta_1), \theta_1] \left(\frac{\theta_1}{\theta_0} \right)^\alpha = q[\phi_m(\theta_2), \theta_2] \left(\frac{\theta_2}{\theta_0} \right)^\alpha, \quad (2)$$

that is, Q^* is set equal to the vertically weighted PV at the vortex edge (defined by ϕ_m) at two isentropic levels θ_1 and θ_2. The system (2) constitutes two equations in two unknowns, Q^* and α, and can be straightforwardly solved numerically to obtain a unique solution. We take $\theta_1 = 600$ K and $\theta_2 = 1400$ K to be representative of the vortex in the lower and upper stratosphere, respectively. For each SSW identified in Part I, except for the March 1971 event, the vortex edge is found to be well defined at both of these levels, and the surface $Q = Q^*$ lies close to the equivalent latitude ϕ_m throughout the potential temperature range of interest.[2]

The three-dimensional isosurface $Q = Q^*$, bounded below and above by the 400- and 1600-K isentropic surfaces, respectively, serves to define the outer boundary of the Arctic polar vortex throughout the period of each SSW. In addition to being a good fit to the location of the maximum latitudinal gradient of PV, the $Q = Q^*$ isosurface is straightforward to calculate and plot, requiring only the values of α and Q^* for each SSW. Following the above procedure for each event, the vortex edge was found to be within the range $Q^* = 38 \pm 7$ PVU (1 PVU $\equiv 10^{-6}$ K m^2 s^{-1} kg^{-1}), and the vertical scaling parameter $\alpha = -4.25 \pm 0.35$. The vertically weighted PV Q thus generated can also be used to generate composite SSWs, as will be described next.

It is worth noting that this method of deriving an isosurface of PV $Q = Q^*$ representing the vortex edge at any time is not restricted to SSW events. By generating a value of Q^* corresponding to the vortex edge, whether on a daily basis or by using a single value over a given time period, the three-dimensional evolution of the vortex can be easily represented at any time. However,

---

[2] For the March 1971 event the vortex was found to have an unusual structure in the upper stratosphere, and the above method could not be applied. This is most likely due to the March 1971 event occurring late in the winter, and to the polar vortex being relatively weak at upper levels. Although the algorithm of Part I identifies the event as a midwinter SSW, it in fact shares many dynamical features with final warmings. The March 1971 event is therefore excluded from our analysis below, leaving 15 displacement events and 13 splitting events.

the fact that PV is not conserved over long time periods in the stratosphere implies that at large times before or after the time at which Q^* is calculated, the isosurface $Q = Q^*$ may no longer correspond closely to the vortex edge.

c. Construction of composite fields

The evolution of the Arctic polar vortex during "typical" vortex-displacement and vortex-splitting events can be investigated using time-lag composites of the vertically weighted PV Q. The advantage of using Q to create the composites, rather than PV q, for example, is that the component of the PV field that increases exponentially with height has been removed from Q by the transformation (1) on a case-by-case basis. This ensures that the effects of the seasonal cycle and inter-annual variability on that part of the untransformed PV that grows exponentially with height are removed, and thus avoids biasing the resulting composites toward a few individual events.

To generate the time-lag composites, an onset time ($\tau = 0$) must be selected for each event. For the vortex-displacement events, $\tau = 0$ is taken to be the event onset date given in Part I. For vortex-splitting events, the dominant feature of the SSW is the time at which the vortex splits from a single vortex into two smaller daughter vortices. This time is not necessarily the onset time as given in Part I. More physically meaningful composites are generated if $\tau = 0$ is taken as the first recorded time at which the vortex, defined as the region $Q > Q^*$, splits into two pieces of comparable area on the 850-K isentropic surface, rather than using the onset date given in Part I. In most cases, this time differs by less than 2 days compared to the onset time identified by Part I, with six vortex-split SSWs having an observed onset before that given in Part I and seven vortex-split SSWs having an observed onset after that given in Part I.

Time-lag composites of Q can be used to examine the typical evolution of the vortex, relative to the SSW onset time, for events of each type. The composite PV field $\bar{Q}(\tau)$ is constructed by averaging the vertically weighted PV fields at time lags τ, relative to the onset time. Here, we focus on the time-lag range from $\tau = -10$ to $\tau = 10$ days.

A simple average of all relevant events is taken

$$\bar{Q} = \frac{\sum_{i=1}^{N} Q_i}{N}, \tag{3}$$

where N is the number of individual events of each type. For vortex-displacement events $N = 15$, corresponding to the events identified by Part I, and for

vortex-splitting events $N = 13$, corresponding to the 14 events identified by Part I minus the March 1971 event. For each event, the composite vortex edge is defined as $\bar{Q} = \bar{Q}^*$, where $\bar{Q}^*$ is the average of Q^* over all the events.

d. Vortex moment diagnostics and the equivalent ellipse

Next, an alternative methodology is described that can be used to generate objective measures of the vortex centroid, orientation, and ellipticity. The method consists of a novel adaptation of the elliptical diagnostics methodology of Waugh (1997). The basic idea is to exploit the approach of vortex moments (e.g., Melander et al. 1986), in order to define an "equivalent ellipse"; an elliptical vortex of uniform PV that has identical moment diagnostics to the polar vortex on a given isentropic level. The main modification to the technique as applied by Waugh (1997) is that we avoid the definition of a vortex edge prior to calculating the vortex moments altogether. A significant advantage of this modification to the technique is that the new diagnostics can serve to verify that our results are independent of the choice of vortex edge described above. Note, however, that Waugh (1997) in fact found that the results of his climatological study were not very sensitive to the choice of the vortex edge.

Specifically, the equivalent ellipse at a given time t, and on the isentropic surface with potential temperature θ, is calculated as follows. The Ertel's PV $q(\lambda,\phi,\theta,t)$ is considered initially as a function of longitude λ and latitude ϕ. Because we are interested in diagnosing the location, strength, and orientation of the polar vortex only, q is first modified to remove any anomalously low values of PV near the pole. Air with low PV is likely to have originated in the tropics and will have its own distinct dynamical significance, which could be investigated separately if required. A modified PV field $\hat{q}$ is defined by

$$\hat{q} = \begin{cases} q & \text{where} \quad q \geq q_b, \\ q_b & \text{where} \quad q < q_b. \end{cases} \tag{4}$$

Here $q_b(\theta,t)$ is a representative background value of PV calculated by averaging q poleward of 45°N. The moment diagnostics are then applied to the modified field $\hat{q}$, after first transforming to Cartesian coordinates (x,y), to facilitate the direct exploitation of the results of Melander et al. (1986). Although it is possible to generalize the moments approach to spherical polar coordinates (Dritschel 1993), Waugh (1997) found that the

transformation to Cartesian coordinates greatly simplified the moment calculations, and has minimal impact upon the results obtained. Hence, Lambert's azimuthal equal-area projection is employed, giving $x = R\cos\lambda$, $y = R\sin\lambda$, and $R = a\sqrt{2(1 - \sin\phi)}$, with a being the earth's radius. Note that this choice of R ensures that the mappings from spherical to Cartesian coordinates and back again are area preserving.

It is useful to introduce the absolute vortex moment $M_{kl}(\theta, t)$, given by

$$M_{kl} = \iint (\hat{q}(x, y) - q_b) x^k y^l \, dx \, dy, \qquad (5)$$

where the integral is taken out to the "equator" at $R = \sqrt{2}a$. Note that because $\hat{q} = q_b$ nearly everywhere where $R > a/\sqrt{2}$, the region between $R = a/\sqrt{2}$ and $R = \sqrt{2}a$ seldom contributes to the moment integrals. The equivalent ellipse is uniquely defined by its centroid $(\bar{x}, \bar{y})$, its aspect ratio r, its angle of orientation ψ and its area A. These are determined as follows. First, the centroid is obtained

$$(\bar{x}, \ \bar{y}) = \frac{1}{M_{00}}(M_{10}, \ M_{01}), \qquad (6)$$

allowing the relative vortex moments

$$J_{kl} = \iint [\hat{q}(x, y) - q_b] (x - \bar{x})^k (y - \bar{y})^l \, dx \, dy, \qquad (7)$$

to be defined. Then ψ, defined as the angle between the x axis and the major axis of the ellipse, and the aspect ratio r are obtained as

$$\psi = \frac{1}{2} \tan^{-1} \left(\frac{2 J_{11}}{J_{20} - J_{02}} \right), \qquad (8)$$

$$r = \left| \frac{(J_{20} + J_{02}) + \sqrt{4J_{11}^2 + (J_{20} - J_{02})^2}}{(J_{20} + J_{02}) - \sqrt{4J_{11}^2 + (J_{20} - J_{02})^2}} \right|^{1/2} . \qquad (9)$$

Importantly, the value of the aspect ratio r is invariant under both translations and rotations of the modified PV field $\hat{q}$.

Finally, the equivalent area of the ellipse is defined to be

$$A = \frac{M_{00}}{q_b}. \qquad (10)$$

The definition (10) implies that the equivalent area A is an objective measure of not only the actual area but also the intensity of the polar vortex. The area can there-

fore be regarded as an objective measure of the total strength, and hence dynamical significance, of the vortex on a given isentropic level. This is a substantial difference from Waugh (1997), where the area represents the area within a closed PV contour selected by a separate algorithm. Note that here the PV anomaly associated with the vortex is normalized by the background value of PV q_b, so that if the PV in the polar vortex is significantly less than $2q_b$, then the equivalent ellipse will appear smaller than the original vortex. Conversely, the equivalent ellipse will be larger if PV in the polar vortex significantly exceeds $2q_b$. Once calculated, the values of $\bar{x}$, $\bar{y}$, r, ψ, and A allow the coordinates of the boundary of the unique equivalent ellipse to be determined parametrically. The Cartesian coordinates of this boundary are then mapped back to longitude–latitude coordinates.

The result of the above procedure applied to the February 1979 vortex on the 850-K isentropic surface, 2 days before the warming onset time at 0600 UTC 21 February 1979, is shown in Fig. 1a, plotted together with the actual PV distribution at that time. It is clear that the equivalent ellipse corresponds closely to the actual polar vortex (note that the shaded region corresponds to the vortex as defined by the edge-finding method described above). The equivalent ellipse method therefore serves to reduce the essential elements of a two-dimensional field to a few key parameters.

The above method is found to work well whenever there is a single, relatively well-defined polar vortex. To investigate vortex-splitting SSWs, however, we also need to investigate situations where the vortex splits into two. The above procedure can be adapted when a vortex split has taken place, that is, when $\tau > 0$ during the vortex split SSWs, as follows. First, the vortex centroid $(\bar{x}, \bar{y})$ and the orientation of the modified PV field $\hat{q}$ are determined as above. Next, the Cartesian domain is divided into two regions, R_1 and R_2, by the straight line $(y - \bar{y}) = -(x - \bar{x}) \cot \psi$, that is, the straight line through the centroid that is perpendicular to the major axis of the $\hat{q}$ distribution. Two new modified PV distributions are then defined as

$$\hat{q}^{(1)} = \begin{cases} \hat{q} & \text{in } \mathcal{R}_1, \\ q_b & \text{in } \mathcal{R}_2, \end{cases} \quad \hat{q}^{(2)} = \begin{cases} \hat{q} & \text{in } \mathcal{R}_2, \\ q_b & \text{in } \mathcal{R}_1. \end{cases} \qquad (11)$$

The method described above is then applied to $\hat{q}^{(1)}$ and $\hat{q}^{(2)}$ in turn to derive two equivalent ellipses. The results of the application of this method to the polar vortex on 850 K on 21 February 1979, that is, at the SSW onset date ($\tau = 0$) 2 days later than Fig. 1a, is shown in Fig. 1b. The method clearly captures the location,
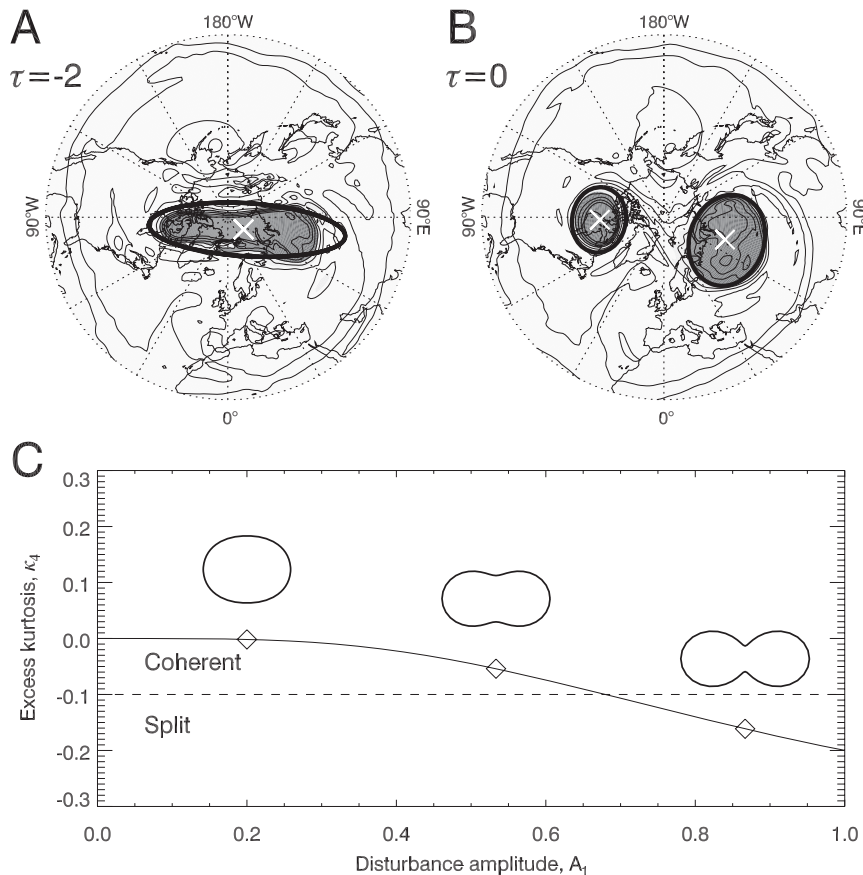
FIG. 1. (top) Ertel's potential vorticity q on the 850-K isentropic surface at (a) 0600 UTC 19 Feb 1979 , $\tau = -2$ days and (b) 0600 UTC 21 Feb 1979, $\tau = 0$ days. Shaded regions correspond to $q > q^*$, where $q^* = 370$ PVU (1 PVU $\equiv 10^{-6}$ K m^2 s^{-1} kg^{-1}) is the vortex edge value [for the vertically weighted PV, $Q^* = q^* (\theta/\theta_0)^\alpha = 27.5$ PVU, where $\alpha = -4.467$ for the February 1979 event]. The black ellipses show the "equivalent ellipses" calculated as described in section 2. The locations of the ellipse centroids are marked by white crosses. Note that in (b) a vortex split has occurred and therefore two ellipses are found following the algorithm described in the text. (c) The dependence of the excess kurtosis parameter κ_4 on the parameter A_1 [in Eq. (13)] that controls the extent to which an idealized vortex of uniform vorticity differs from an ellipse.

shape, size, and orientation of each of the two vortices following the split.

It is perhaps undesirable to be reliant on a vortex split definition based on the behavior of a particular vortex edge contour (e.g., the vortex edge contour on the 850-K surface as described above). Instead, can the vortex moments themselves be used to define when a vortex split takes place? To answer this question, higher-order moments of the $\hat{q}$ field must be considered. In statistics, the bipolarity of a given distribution with respect to a reference distribution is measured by the excess kurtosis $\kappa_4 = J_4/J_2^2 - F$, where J_4 and J_2 are the fourth- and second-order moments relative to the centroid, and F is determined by the reference distribution. Ideally, we would like an analogous quantity that measures the

bipolarity of the PV distribution, while satisfying the following properties:

- it is invariant to translations of $\hat{q}$;
- it is invariant to rotation of $\hat{q}$;
- it is independent of the size of the vortex (or, equivalently, the horizontal scale of the $\hat{q}$ distribution);
- it is equal to zero for a vortex of uniform PV that is exactly elliptical, while it takes negative values if the vortex becomes "pinched" in a manner that might lead to a split, and takes positive values (e.g.) for a diamond-shaped vortex or a vortex undergoing strong filamentation.

After some working, it can be shown that the following quantity satisfies the above five properties:

$$\kappa_4 = M_{00} \frac{J_{40} + 2J_{22} + J_{04}}{(J_{20} + J_{02})^2} - \frac{2}{3} \left[\frac{3r^4 + 2r^2 + 3}{(r^2 + 1)^2} \right], \quad (12)$$

where r is the aspect ratio (9). The second term in (12) serves to ensure that $\kappa_4 = 0$ for an ellipse of uniform vorticity. We refer to κ_4 as the "excess kurtosis" parameter, following the terminology for the fourth-order moments in statistics, and it serves as a measure of the bipolarity of the PV distribution $\hat{q}$. Figure 1c shows κ_4 as a function of A_1 for the idealized vortex patch of uniform vorticity with boundary in polar coordinates given by

$$r = 1 + A_1 \cos 2\theta. \quad (13)$$

As A_1 is increased (see Fig. 1c), the patch vortex becomes progressively more pinched, and the corresponding value of κ_4 falls uniformly. This allows us to define a practical criterion $\kappa_4 < -0.1$ that can serve as an indicator of whether or not a split has taken place. The evolution of κ_4 during actual vortex split SSWs will be discussed in the following section.

3. Results

a. Vortex structure prior to stratospheric sudden warmings

Figure 2 shows the vertically weighted PV field $\bar{Q}$, as a function of equivalent latitude and potential temperature, 9 days before the SSW onset. The aim is to illustrate the initial structure of the polar vortex prior to the SSW taking place, both in a mean sense, as illustrated by the splitting and displacement composites in Figs. 2a,b, and also for the January 1987 and February 1979 individual events in Figs. 2c,d. The advantage of choosing equivalent latitude and potential temperature is that when viewed in these coordinates the PV distribution is relatively insensitive to particular synoptic conditions, and the results are consequently largely independent of displacements or distortions of the polar vortex. The January 1987 and February 1979 events are chosen as case studies because they have been previously identified as archetypal SSWs of the displacement and splitting type, respectively. The February 1979 event, in particular, has been much studied (Andrews et al. 1985; Dunkerton and Delisi 1986; Jung et al. 2001; Labitzke 1981; Manney et al. 1994; Palmer 1981), although not primarily from a vortex dynamics perspective as here. Aspects of the January 1987 SSW are discussed by Manney et al. (2005b).

In each panel the polar vortex is clearly visible as a region of elevated $\bar{Q}$ at high latitudes. The thick black contour shows the vortex edge $\bar{Q} = \bar{Q}^*$ defined above, with the black crosses marked on Figs. 2c,d showing the equivalent latitudes $\phi_m(\theta)$ of the local maxima in PV gradient. In each case, the vortex edge isosurface $\bar{Q} = \bar{Q}^*$ is a good fit to the equivalent latitudes of the PV gradient maxima across a large altitude range. The region of elevated PV associated with the vortex is located largely between the 400- and 1600-K isentropic surfaces, justifying our choice of these levels as the upper and lower boundaries of our three-dimensional isosurface plots.

Does the vortex, viewed in this framework, have a characteristic structure that is distinct to the type of SSW that is about to take place? To answer this question comprehensively requires a detailed statistical study that is beyond the scope of the current work. However, Figs. 2a,b reveal that in a mean sense the vortex is somewhat stronger prior to splitting SSWs compared to displacement SSWs.[3] Nevertheless, there is substantial variability between individual events in the prewarming vortex structure, particularly in the slope of the vortex edge and in the vortex strength as a function of height.

The February 1979 and January 1987 vortices have some distinguishing features. Figure 2c shows that in February 1979, the polar vortex edge slopes significantly poleward with height, that is, the vortex has a distinctive conical shape. The conical shape may be a consequence of the dynamical activity "preconditioning" the vortex, as described by Labitzke (1981), and interestingly the Antarctic vortex had a similar conical shape prior to the September 2002 SSW (e.g., Esler et al. 2006). The February 1979 vortex also remains well defined at altitudes as low as 15 km (400 K). By contrast, the January 1987 vortex shown in Fig. 2d is poorly defined below 20 km, that is, there is relatively little contrast below 20 km between the vortex and surf-zone PV. The question of whether or not the prior vortex structure is important in determining the subsequent evolution of the polar vortex during the SSW will be the subject of further investigation.

b. Vortex-splitting stratospheric sudden warmings

The three-dimensional evolution of the Arctic polar vortex, defined here as the air mass enclosed by an isosurface of the vertically weighted PV between the 400- and 1600-K isentropes during vortex-splitting SSWs, is shown in Fig. 3. The left panels show the evolution of

---

[3] Note that the equivalent latitude transformation negates the possibility that the difference in vortex strengths is due to differences in the displacement or distortion of the vortex, as might be the case if a straightforward zonal average were taken.
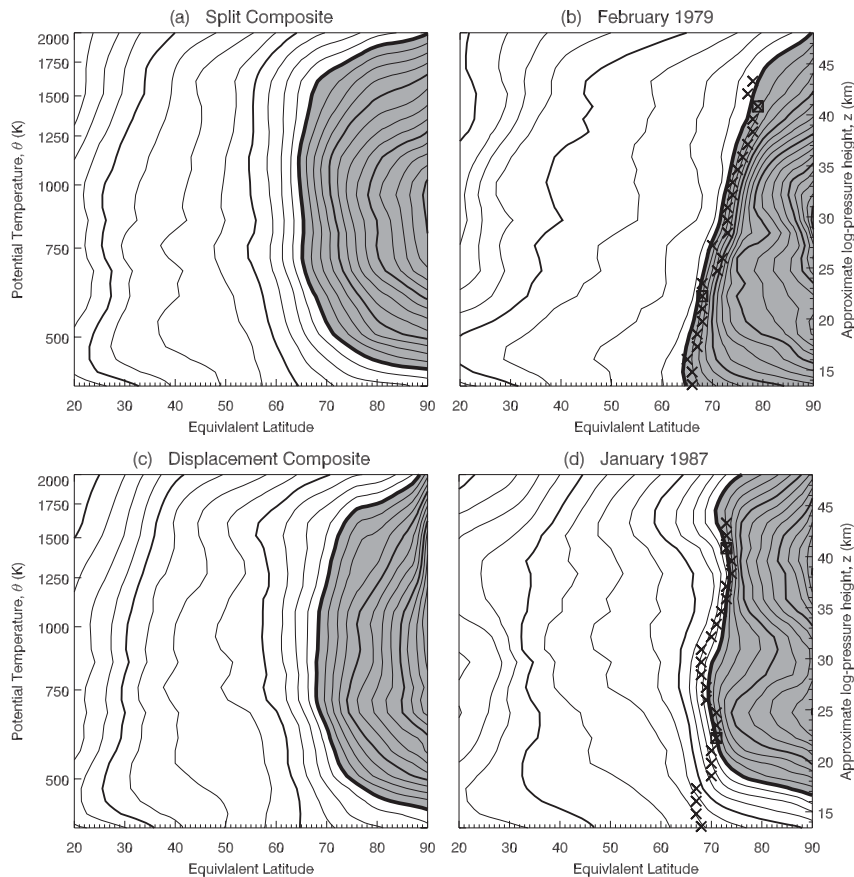
FIG. 2. (a) Composite vertically weighted potential vorticity $\bar{Q}$ at time lag $\tau = -9$ days, as a function of equivalent latitude and potential temperature, for the vortex split events. The contour interval is 2.93 PVU, and every fourth contour is grayed. The thick black contour dividing the shaded (vortex) region from the white (background) plots the vortex edge value $\bar{Q} = \bar{Q}^* = 38$ PVU. (b) As in (a), but for the vertically weighted PV Q at 0600 UTC 12 Feb 1979 ($\tau = -9$ days before the calculated onset date of the February 1979 event). Contour interval is 3.19 PVU; $Q^* = 27.5$ PVU. The black $\times$ signs show the location $\phi_m(\theta)$ of the maximum of the gradient $|\partial Q/\partial\phi_e|$, used in determining the vortex edge. (c) As in (a), but for $\bar{Q}$ at $\tau = -9$ days for the vortex-displacement events. Contour interval is 2.78 PVU; $\bar{Q}^* = 38$ PVU. (d) As in (b), but for Q at 0000 UTC 23 Jan 1987. Contour interval is 4.51 PVU; $\bar{Q}^* = 54.5$ PVU.

the vortex during the composite SSW, illustrated by the isosurface $\bar{Q} = \bar{Q}^*$ defined above, whereas the right panels show the corresponding pictures of the isosurface $Q = Q^*$ for the February 1979 event. Snapshots of the isosurfaces at time lags $\tau = -9, -2, 0,$ and $+5$ days are shown to illustrate the vortex behavior prior to and during the SSW. For comparison, Fig. 4 shows the equivalent ellipses on the 505-, 715-, 1010-, and 1425-K isentropic surfaces at the same time lags. Note that for visualization purposes the equivalent ellipses are not shown at their full size (see Fig. 4 caption for scaling details).

Figures 3 and 4 reveal that the composite SSW and February 1979 SSW share many key features. At time

lag $\tau = -9$ days, that is, 9 days before the onset time when the vortex edge contour on the 850-K isentrope splits, the composite vortex has a relatively circular cross section at all levels. The composite vortex centroid location ranges between (84°N, 75°E; 450 K) and (80°N, 40°W; 1425 K) . This location is somewhat poleward and eastward of the climatological values for most of the winter, as seen in Waugh and Randel (1999, see their Figs. 4d,f).

As the SSW onset time approaches, the vortex begins to slowly elongate and, for both the composite and February 1979 events, the major axis of the vortex rotates cyclonically. The February 1979 vortex rotates through approximately 60° between $\tau = -9$ and -2 days. By

Composite   February 1979
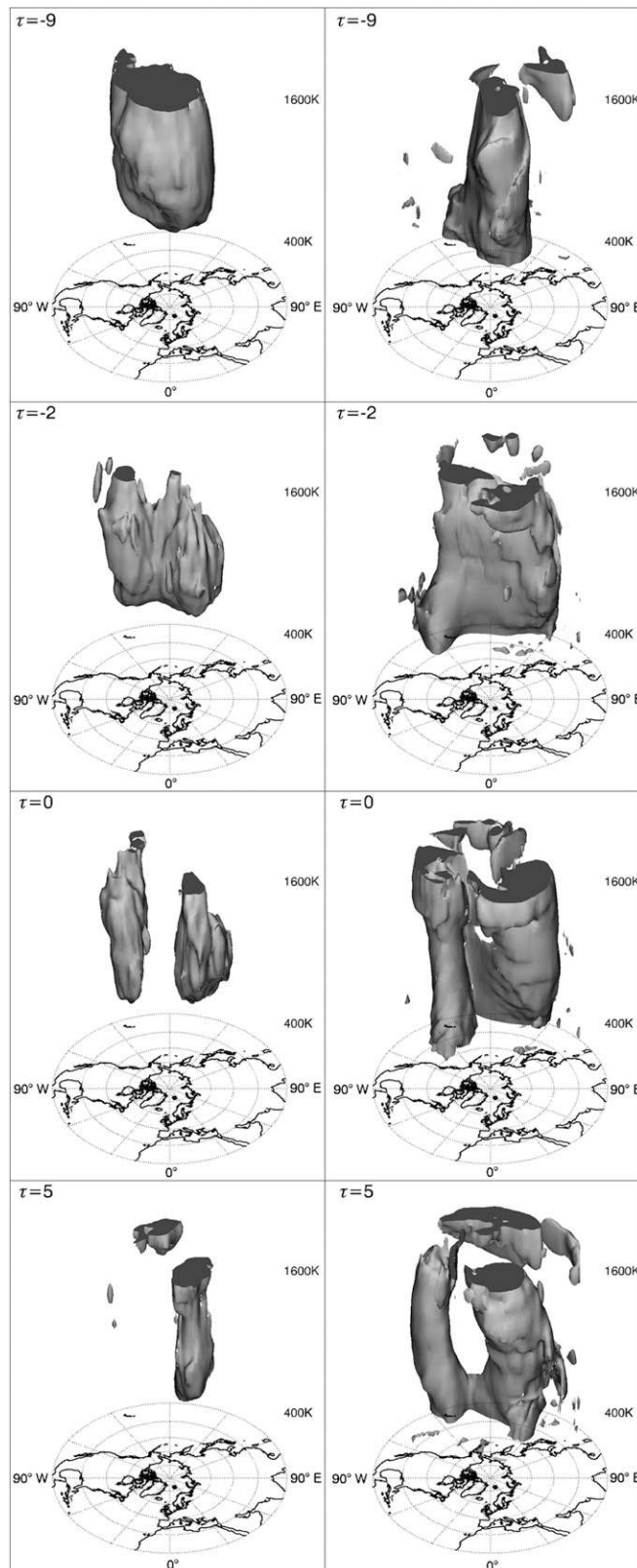
$\tau = -2$ days the cyclonic rotation slows and the vortex elongates rapidly along an axis approximately parallel to the 80°E–100°W longitude circle. Both the composite and February 1979 vortices have by this stage become "pinched," and begin to roll-up into two distinct centers located on either side of the original centroid.

By $\tau = 0$, the polar vortex has split completely and two distinct "daughter" vortices have formed. At $\tau = 0$ the centroids of the daughter vortices at 850 K are located approximately 4500 km apart, and are approximately located at 65°N, 90°W (Hudson Bay, northeast Canada) and 65°N, 60°E (Siberia), respectively. The two daughter vortices initially have a relatively "barotropic" vertical structure, although each has a slight westward tilt with height. As the SSWs develop further, the two vortices begin to rotate in retrograde motion around their (weighted) mean position. The rotation rate varies strongly with altitude, with stronger rotation in the upper stratosphere, as can be seen in Fig. 4 at $\tau = 5$ days, when the vortices at 1425 K are at 10° and 170°E, respectively.

The differential rotation inevitably leads to the breakdown of at least one of the vortices. For both the composite and February 1979 event, it is clear that it is the Siberian vortex that is the dominant one. By $\tau = 5$ days the Canadian vortex is no longer detectable in Fig. 3, except at high levels. In contrast to at earlier times, however, the details at this stage vary significantly between individual events. Hence, the composite at later times is no longer a meaningful guide to typical behavior.

The most striking aspect of Figs. 3 and 4 is the similarities between the composite and February 1979 SSWs. Compositing or otherwise averaging observational data can easily obscure important information. However Figs. 3 and 4 reveal the following:

- The composite vortex undergoes a vortex split resembling the individual events without any filtering of the data other than the PV transformation described above. This is because the vortex-splitting SSWs generally occur at the same orientation in relation to the earth's surface. At a given time lag, τ the vortex (or vortices) has a fixed characteristic orientation relative to the earth's surface, and this orientation evolves systematically during the life cycle of the SSW.
- The composite vortex has a life cycle of elongation and cyclonic rotation, followed by splitting and then

retrograde rotation of the two resulting vortices about their mean position.
- During the composite SSW, the vortex elongation and split occur near-simultaneously at all altitudes, and both vortices remain coherent for several days.
- For both the composite and February 1979 SSWs it is the Siberian daughter vortex that dominates. A few days after the SSW onset, the Canadian vortex has become sheared out at upper levels ($>$1010 K).

Despite the similarities detailed above, there remain some notable differences between the composite and the February 1979 SSW. First, the pair of vortices that result from the February 1979 vortex split have a significantly larger cross-sectional area than those of the composite. The explanation is that, as will be seen below, the composite vortices are unrepresentatively small because they are an average over an ensemble of vortices with slightly different locations. Second, the February 1979 vortex extends much lower in the stratosphere, as discussed in connection with Fig. 2c. In fact, there is considerable variability in the vortex structure at low levels between SSWs. Finally, the vortex is substantially more elongated prior to the February 1979 SSW, and the vortex major axis rotates cyclonically much more rapidly prior to the warming (through 60°, as opposed to 10° for the composite).

The near-simultaneous splitting of the vortex at all heights can be highlighted further by examining how the vortex-splitting parameter, or "excess kurtosis", κ_4 evolves during the composite and February 1979 SSWs. Figure 5 shows a contour plot of κ_4 as a function of altitude and time lag τ in each case. Dark shaded regions indicate where $\kappa_4 < -0.1$, that is, where a vortex split is diagnosed to have taken place. In each case, it is clear from Fig. 5 that the split occurs near simultaneously over a wide-altitude range, with the split criterion $\kappa_4 < -0.1$ being widely satisfied by $\tau = -1$ day. One unexpected feature of the February 1979 event is the relatively high values of κ_4 in the altitude range 600–750 K throughout the period of the SSW. These high values can be explained by the large difference in the magnitude of the daughter vortices at these levels, with the Siberian vortex being around 3 times greater in magnitude than the Canadian vortex. If the asymmetry between the vortices is sufficiently large, the excess

←
FIG. 3. (left) A three-dimensional isosurface of the composite vertically weighted PV $\bar{Q}^* = 38.0$ PVU, shown between isentropic surfaces of 400 and 1600 K at time lags of (top to bottom) $\tau = -9, -2, 0,$ and 5 days, for the composite vortex split SSW. (right) The isosurface $Q = Q^* = 27.5$ PVU for the February 1979 SSW at the same time lags, with τ relative to the SSW onset defined to be at 0600 UTC 21 Feb 1979.
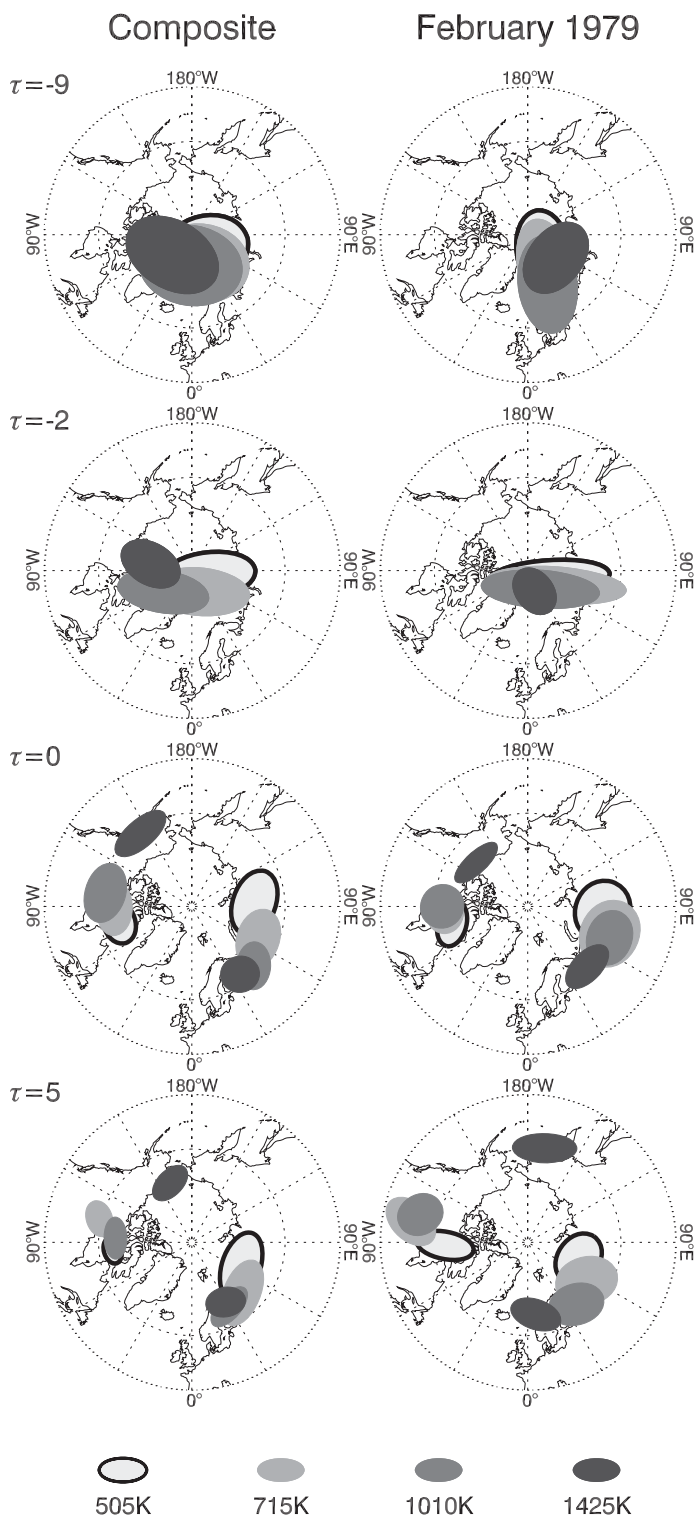
FIG. 4. The two-dimensional equivalent ellipses on isentropic levels of 505, 715, 1010, and 1425 K prior to and during the (left) composite SSW and (right) February 1979 SSW, at time lags of (top to bottom) $\tau = -9, -2, 0$, and 5 days measured relative to the SSW onset time defined in the text in each case. For visualization purposes, the equivalent ellipses are plotted at (left) 0.6 times and (right) 0.3 times their actual area, as defined in Eq. (10).
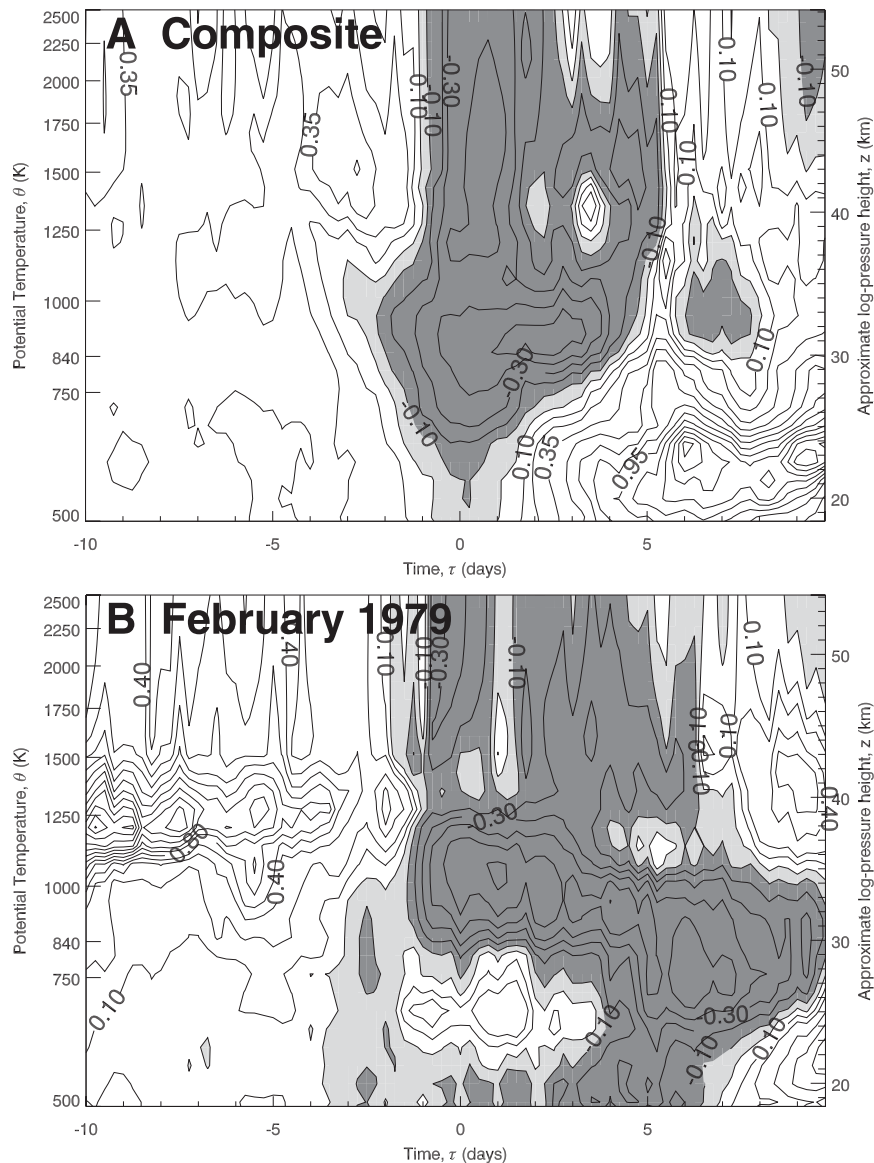
FIG. 5. Evolution in time and altitude–potential temperature of the excess kurtosis parameter κ_4, defined in Eq. (12). (a) The composite SSW and (b) the February 1979 event. The contour interval is 0.1 in each panel. Values below zero are lightly shaded and values below -0.1 are heavily shaded.

kurtosis parameter does not indicate that a split has taken place. In summary, the κ_4 parameter serves as a useful, objective definition of the onset of a vortex split, although some care is necessary in its interpretation.

To give an impression of the interevent variability, Fig. 6 shows the equivalent ellipses for all 13 splitting events, plus the composite (filled black ellipse) on the 600-K isentropic surface. Note that, as with Fig. 4, in order to emphasize differences between events the equivalent ellipses are scaled to be smaller than the actual

vortex size. To facilitate comparison with Fig. 4 the February 1979 event is also highlighted in Fig. 6 as a filled blue ellipse. Figure 6 reveals why the composite method has been successful in describing a typical splitting SSW: 9 out of the 13 events follow the sequence of rotation, elongation, and splitting described for the composite–February 1979 events above quite closely.

Four events can be singled out as exhibiting somewhat different behavior. These are the SSWs of January 1977 (thick blue ellipse), December 1987 (cyan), March

$\tau=-9$ 180°W   600K

$\tau=-2$ 180°W

$\tau=0$ 180°W

$\tau=5$ 180°W

Comp F79 J77 D87 M88 F99 Others
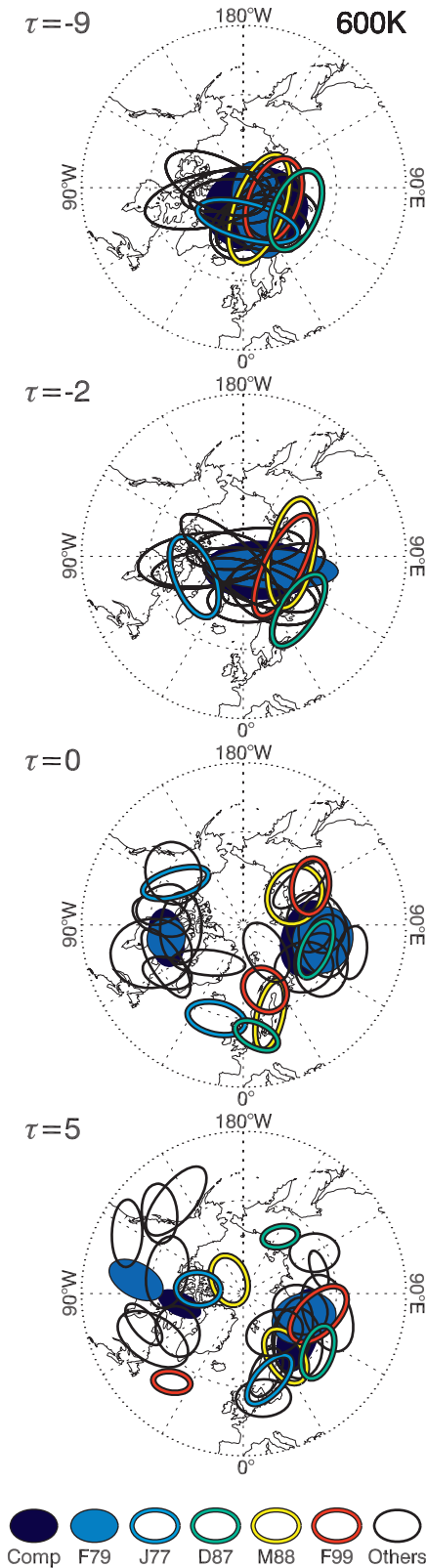
1988 (yellow), and February 1999 (red). The January 1977 event stands out among the four events, because the vortex split occurs in a direction parallel to the 160°W–20°E great circle. Notably, it is not selected as an SSW by the Part I algorithm when NCEP–NCAR reanalysis data are used in place of ERA-40. For the remaining three events, the split occurs approximately parallel to the 20°W–160°E great circle. Further examination of these latter three "outlier" events reveals the significant vertical structure of the daughter vortices compared to the composite "barotropic" splitting SSW. The European vortex in particular has a baroclinic structure and location that is reminiscent of the vortex structure during displacement SSWs, which will be described next.

Although not pursued here, further investigation of these hybrid characteristics may go some way to determining which SSW behavior is dominant during warming events that exhibit conditions favoring both types of SSW.

c. Vortex-displacement stratospheric sudden warmings

The evolution of the Arctic polar vortex during both the composite vortex displacement and the January 1987 SSW is shown in Figs. 7 and 8 . The time lags relative to the SSW onset time (taken to be 0000 UTC 23 January 1987 in our case study, following Part I) shown in Figs. 7 and 8 are $\tau = -10, -1, 1,$ and 4 days. As above, Fig. 7 shows the three-dimensional isosurface defined by $Q = Q^*$ (or $\bar{Q} = \bar{Q}^*$ for the composite), while Fig. 8 shows the equivalent ellipses of the polar vortex on a number of isentropic surfaces (450–1510 K). As above, the composite and January 1987 SSWs share many key features. At $\tau = -10$ days, that is, 10 days before the SSW onset date, the composite and January 1987 vortices each have a near-circular cross section in the lower stratosphere. The composite vortex centroid tilts westward and equatorward with altitude, ranging between (77.5°N, 70°E; 450 K) and (72°N, 15°E; 1425 K). The vortex is therefore significantly further equatorward, and has a substantially greater tilt with altitude prior to a displacement SSW than a splitting SSW. Also, the vortex is displaced further

←

FIG. 6. The equivalent ellipses for all 13 vortex-splitting SSW events at time lags of (top to bottom) $\tau = -9, -2, 0,$ and 5 days, on the 600-K isentropic surface. The composite SSW is shown as a filled black ellipse, and the February 1979 SSW as a filled blue ellipse. Four further events discussed in the text are highlighted as unfilled colored ellipses—January 1977 (blue), December 1987 (cyan), March 1988 (yellow), February 1999 (red)—and the remaining events are all unfilled black ellipses. The equivalent ellipses are plotted at 0.3 times their actual area (as in Fig. 4).

Composite   **January 1987**
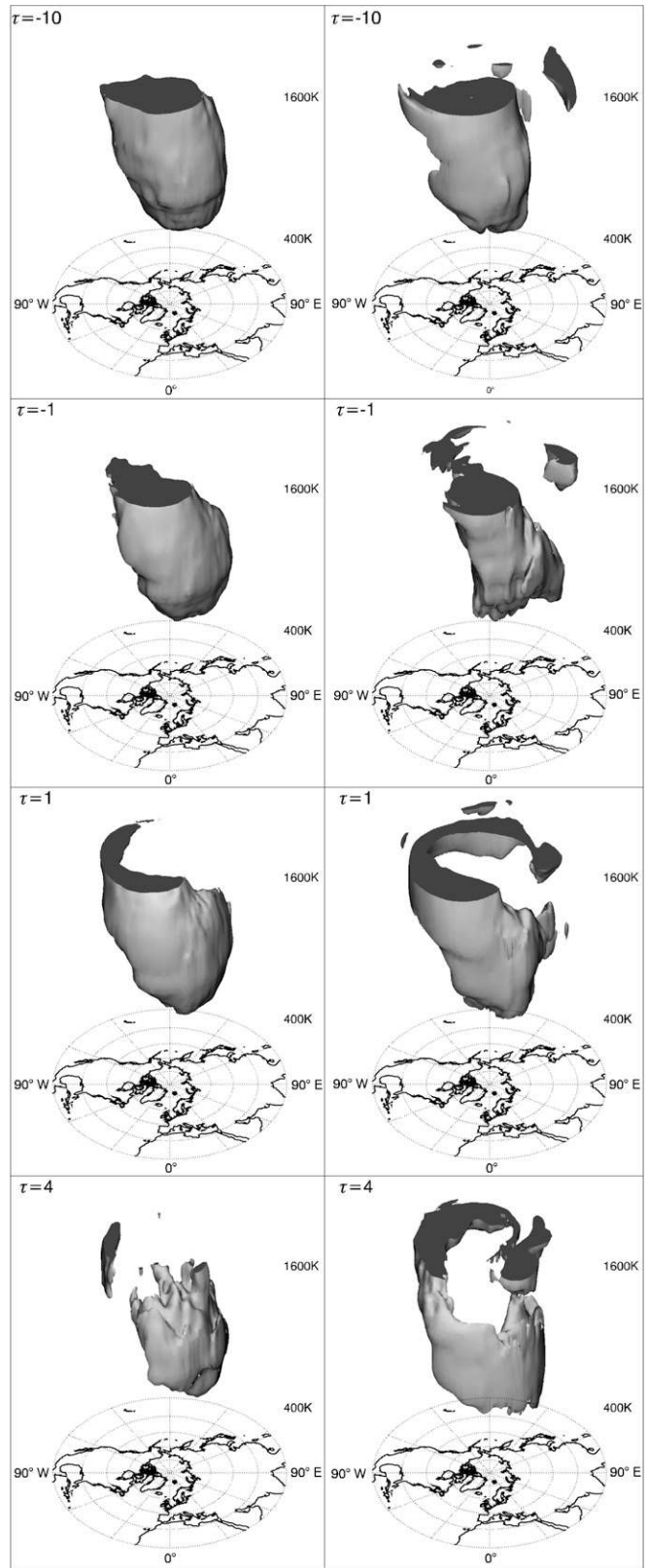
from the pole than the climatological displacement of 10°–14° (Waugh and Randel 1999, see their Figs. 4d,f), indicating that the vortex may be somewhat preconditioned prior to a displacement SSW (see, e.g., Labitzke 1981). Both the composite and January 1987 vortices are seen to be more elliptical at higher levels. The major axis of the ellipse is roughly parallel to the direction of the westward tilt along the 70°–80°N latitude band in each case.

As the onset time for the SSW draws closer, the westward tilt of the vortex with altitude increases, and the entire vortex simultaneously moves equatorward, so that by $\tau = -1$ day the composite vortex centroid lies within the 67°–74°N latitude band. The lower vortex moves slowly eastward while the upper vortex moves rapidly westward so that at $\tau = -1$ day the composite vortex centroid at 450 K is separated from that at 1425 K by 135° longitude. The process continues through the SSW onset itself, and by $\tau = 1$ day the composite vortex centroid position ranges between (75°N, 95°E; 450 K) and (68°N, 53°W; 1425 K), that is, the vortex spans nearly half the globe in longitude. During this period the upper vortex is strained by the large-scale flow, and by $\tau = 1$ day the vortex consequently has a crescent-like cross section. During the next few days, the deformation of the upper vortex continues until it is thinned and stirred into the background. By $\tau = 4$ days the averaging process used to create the composite acts to diffuse the upper-level composite vortex over a wide region; hence, the upper vortex is entirely absent from the $\tau = 4$ days three-dimensional representation. In the January 1987 event, the upper and lower parts of the vortex have become largely disconnected. At this stage, the life cycle of the displacement SSW is essentially complete. The remnants of the vortex in the lower stratosphere are located equatorward and eastward compared to the vortex at the beginning of the life cycle (at $\tau = -10$ days). This typical life cycle displayed by the composite vortex is in agreement with individual SSWs as discussed in Manney et al. (1999, 2005b), and references therein.

Apart from minor details of vortex orientation and timing, remarkably few differences are apparent between the January 1987 SSW and the composite in Figs. 7 and 8. The January 1987 event is revealed to be somewhat stronger than typical (Manney et al. 2005b), because the vortex at middle levels (850 K) is displaced as far equatorward as 55°N, compared to 67°N for the composite.

An impression of the interevent variability for all 15 displacement SSWs is given in Fig. 9. The vortex centroids at 450, 850, and 1425 K are plotted for each event at $\tau = -10, -1, 1$, and 4 days. The lines joining these three points intersect the vortex centroids at the intermediate levels. A remarkably consistent picture is revealed for all of the SSWs with the centroids quite tightly clustered at all levels at both early times and as the SSW develops. Differences in timing and vertical structure cause the largest variability in centroid position to occur for the middle-vortex level (850 K). For some events, the 850-K centroid remains close to the lower-vortex (450 K) centroid during the SSW onset, whereas for other events the 850-K centroid follows the upper vortex (1425 K) as it moves rapidly westward and equatorward. By $\tau = 4$ days the upper-vortex (1425 K) centroids also exhibit a large scatter. This is consistent with the lack of coherence of the upper vortex at this time, resulting from the effects of the straining flow thinning out and mixing the vortex into the background at upper levels. Overall, however, Fig. 9 reveals a remarkable degree of similarity between all of the displacement SSW events.

4. Proposed benchmarks

The results above suggest that a new set of modeling benchmarks may now be defined, against which the dynamical performance of GCMs may evaluated. These new benchmarks, in contrast to those proposed in section 7 of Part I, are primarily qualitative in character and relate to the evolving shape of the polar vortex during SSWs.

Specifically, after analyzing model output for vortex-splitting events, one might answer the following questions:

S-1) Does the vortex split near simultaneously, and with an approximate barotropic structure, over the 20–40-km altitude range?

S-2) Is the mean vortex orientation at the onset time of the vortex split such that the daughter vortices are approximately located over Canada and Siberia?

S-3) Is the SSW followed by a retrograde rotation of the daughter vortices about a common centroid,

←

FIG. 7. (left) A three-dimensional isosurface of the composite vertically weighted PV $\bar{Q} = \bar{Q}^* = 38$ PVU, shown between isentropic surfaces 400 and 1600 K at time lags of (top to bottom) $\tau = -10, -1, 1$, and 4 days, for the composite vortex split SSW. (right) The isosurface $Q = Q^* = 54.5$ PVU for the January 1987 SSW at the same time lags, this time relative to the SSW onset defined to be at 0000 UTC 23 Jan 1987.
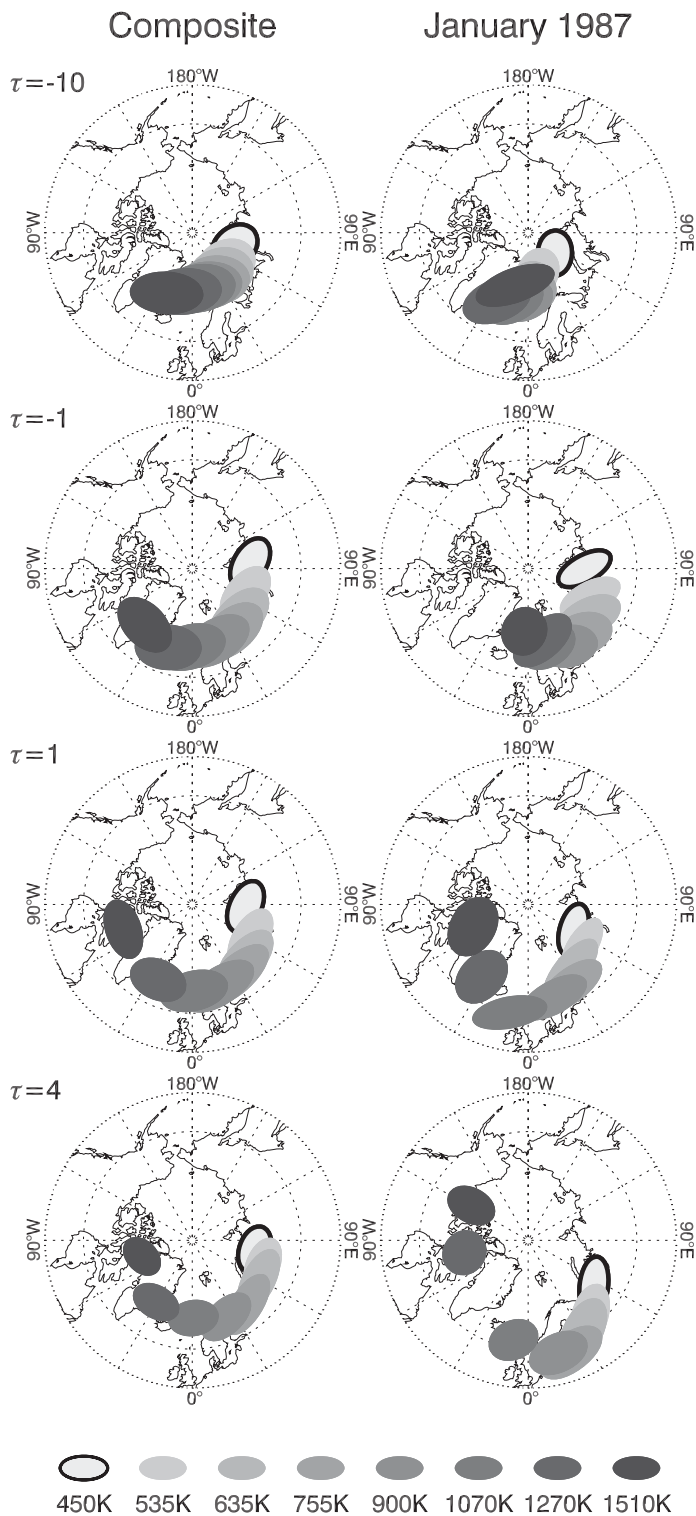
FIG. 8. The two-dimensional equivalent ellipses on eight isentropic levels between 450 and 1510 K prior to and during the (left) composite vortex-displacement SSW and (right) January 1987 SSW. Time lags of (top to bottom) $\tau = -10, -1, 1$, and 4 days are relative to the SSW onset time defined in the text in each case. For visualization purposes, the equivalent ellipses are plotted at (left) 0.2 times and (right) 0.1 times their actual area, as defined in Eq. (10).
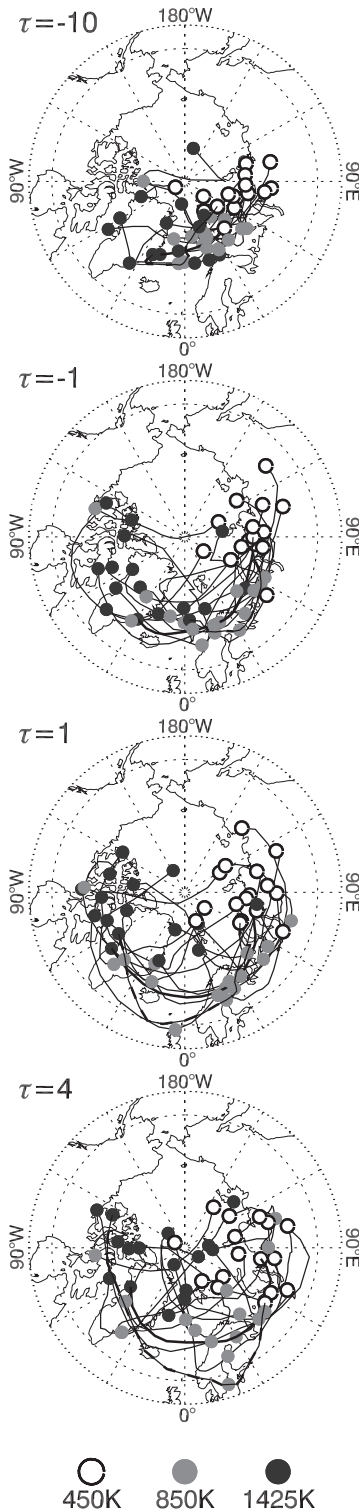
FIG. 9. The location of the vortex centroid (circles) at different isentropic levels (white circle, 450 K; gray circle, 850 K; black circle, 1425 K) for all 15 vortex-displacement SSWs, at (top to bottom) $\tau = -10$, -1, 1, and 4 days. The lines joining the points trace out the centroid location at intermediate levels.

and is this rotation more pronounced in the upper stratosphere?

S-4) Is the Siberian daughter vortex typically larger than its Canadian counterpart?

For the displacement SSWs, different questions must be asked.

D-1) Is the vortex centroid located in the 90°E–0° sector prior to the warming, with a distinct westward tilt with height?

D-2) Is the vortex evolution primarily baroclinic, with the westward tilt with height increasing rapidly as the onset time approaches and an accompanying equatorward displacement of the vortex centroid at all heights?

D-3) At the end of the SSW, does the longitude of the vortex centroid vary strongly with altitude, with up to 180° separating the centroid position in the lower stratosphere (450 K) from that in the upper stratosphere (1425 K)?

D-4) At the end of the SSW is the polar vortex almost entirely contained within the Atlantic hemisphere (90°E–90°W)?

It is proposed that straightforward application of the methodology of section 2 to absolute vorticity GCM output fields (or PV output fields if available) should be sufficient to reveal the bulk behavioral properties of the vortex during a SSW, which can then be compared with the benchmarks above.

5. Conclusions

The current investigation has focused on the evolution of the Arctic polar vortex during major midwinter stratospheric sudden warmings (SSWs). Time-lag composites and vortex moment diagnostics have been used to investigate the distinctive behavior of the polar vortex during both displacement and splitting SSW events. The main results, answering, as far as possible, the questions posed in the introduction, can be summarized as follows:

• The Arctic polar vortex has a highly characteristic vertical structure during the two different types of SSW. During vortex-splitting events, the vortex remains nearly barotropic, with the vortex split occurring near-simultaneously over a large altitude range (at least 20–40 km, see Figs. 3 and 4). In contrast, during vortex-displacement events, the vortex displacement off the pole increases with altitude above 30 km, as does the accompanying elongation of the vortex during the SSW, as can be seen in Figs. 7 and 8.

• For both types of SSW, there is relatively little variation between individual events in the orientation of

the developing vortex relative to the underlying topography, that is, SSWs of each type are to a large extent fixed in relation to the earth's surface. For displacement events the vortex is located in the 90°E–0° sector prior to the warming, and by the end of the warming spans the entire Atlantic hemisphere from 90°E (lower vortex) to 90°W (upper vortex). There is somewhat more variability in the case of vortex-splitting events (see Fig. 6), but the dominant direction for the splits is found to be parallel to the 60°E–120°W great circle.

- Each type of SSW exhibits a distinct life cycle, in the sense that the centroid position, orientation, and elongation of the Arctic polar vortex evolves in a similar fashion during the onset and occurrence of individual splitting events, with a separate behavior during displacement events. Splitting events are characterized by a rapid increase of the aspect ratio of the vortex a few days prior to the SSW (τ = from -4 to -2 days), followed by the roll-up of the elongated vortex into two distinct daughter vortices (τ = from -2 to 0 days). The daughter vortices propagate rapidly apart to a distance of up to 5000 km by $\tau = 1$ day, after which they experience a retrograde rotation around their common centroid, usually leading to the destruction of the weaker Canadian vortex and reformation of the main vortex around the stronger Siberian vortex. For vortex-displacement events, the separation in longitude between the lower vortex (450 K) and the upper vortex (1425 K) increases uniformly during the SSW from 85° longitude (at $\tau = -10$ days) to 150° longitude (by $\tau = 1$ day). By the end of the displacement SSW life cycle (at $\tau = 4$ days) the upper part of the vortex is destroyed by deformation, and the vortex subsequently reforms around its remnants in the lower stratosphere.

In addition to providing the modeling benchmarks listed above, it is intended that the results of this study will be of use to researchers seeking to differentiate between alternative dynamical theories of SSWs. The vertical structure of the vortex is necessarily a key diagnostic of the vertical propagation and nonlinear behavior of Rossby waves in the polar stratosphere. Following Matsuno (1971), SSWs of either type are often described as resulting from the vertical propagation and breaking of stationary planetary waves generated in the troposphere. Alternatively, it has been suggested (O'Neill and Pope 1988; Scott and Dritschel 2006) that the interaction between the polar vortex and Aleutian anticyclone may be important in the development of SSWs. Further, Tung and Lindzen (1979) and Plumb (1981) have suggested that SSWs might occur following the resonant excitation of a normal mode of the

stratosphere, an idea recently developed by Esler and Scott (2005) and Esler et al. (2006) who have suggested that the "barotropic" mode of the polar vortex is the relevant mode in the case of vortex-splitting events. Note that if the resonant excitation of the vortex normal modes in the above theories is due to stationary waves, then the resulting SSWs will be fixed in relation to the earth's surface as observed in this study. The vortex dynamics view of SSWs may be of particular relevance to splitting events, because the vortex split occurs rapidly and in a manner strongly reminiscent of Love's instability of Kirchhoff's elliptical vortex in two-dimensional hydrodynamics (Dritschel 1986; Love 1893), an aspect that will be further investigated. The climatology presented here, of the vertical structure of the Arctic polar vortex during observed SSWs, should be of considerable assistance in quantitatively distinguishing between the above theories.

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1985: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Black, R. X., and B. A. McDaniel, 2007: The dynamics of Northern Hemisphere stratospheric final warming events. *J. Atmos. Sci.,* **64,** 2932–2946.

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

Charlton, A. J., and L. M. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate,* **20,** 449–469.

——, and Coauthors, 2007: A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations. *J. Climate,* **20,** 470–488.

Dritschel, D. G., 1986: The nonlinear evolution of rotating configurations of uniform vorticity. *J. Fluid Mech.,* **172,** 157–182.

——, 1993: A fast contour dynamics method for many-vortex calculations in two-dimensional flows. *Phys. Fluids,* **25A,** 173–186.

Dunkerton, T. J., and D. P. Delisi, 1986: Evolution of potential vorticity in the winter stratosphere of January–February 1979. *J. Geophys. Res.,* **91,** 1199–1208.

Esler, J. G., and R. K. Scott, 2005: Excitation of transient Rossby waves on the stratospheric polar vortex and the barotropic sudden warming. *J. Atmos. Sci.,* **62,** 3661–3682.

——, L. M. Polvani, and R. K. Scott, 2006: The Antarctic strato-spheric sudden warming of 2002: A self-tuned resonance? *Geophys. Res. Lett.,* **33,** L12804, doi:10.1029/2006GL026034.

Jung, J. H., C. S. Konor, C. R. Mechoso, and A. Arakawa, 2001: A study of the stratospheric major warming and subsequent flow recovery during the winter of 1979 with an isentropic vertical coordinate model. *J. Atmos. Sci.,* **58,** 2630–2649.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Labitzke, K., 1981: The amplification of height wave 1 in January 1979: A characteristic precondition for the major warming in February. *Mon. Wea. Rev.,* **109,** 983–989.

Lait, L. R., 1994: An alternative for potential vorticity. *J. Atmos. Sci.,* **51,** 1754–1759.

Legras, B., and D. G. Dritschel, 1993: A comparison of the contour surgery and pseudo-spectral methods. *J. Comput. Phys.,* **104,** 287–302.

Liberato, M. L. R., J. M. Castanheira, L. de la Torre, C. C. DaCamara, and L. Gimeno, 2007: Wave energy associated with the variability of the stratospheric polar vortex. *J. Atmos. Sci.,* **64,** 2683–2694.

Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann, 2004: The life cycle of the Northern Hemisphere sudden strato-spheric warmings. *J. Climate,* **17,** 2584–2596.

Love, A. E. H., 1893: On the stability of certain vortex motion. *Proc. Roy. Soc. London,* **23,** 18–42.

Manney, G. L., J. D. Farrara, and C. R. Mechoso, 1994: Simula-tions of the February 1979 stratospheric sudden warming: Model comparisons and three-dimensional evolution. *Mon. Wea. Rev.,* **122,** 1115–1140.

——, W. A. Lahoz, R. Swinbank, A. O'Neill, P. M. Connew, and R. W. Zurek, 1999: Simulation of the December 1998 strato-spheric major warming. *Geophys. Res. Lett.,* **26,** 2733–2736.

——, and Coauthors, 2005a: Diagnostic comparison of meteoro-logical analyses during the 2002 Antarctic winter. *Mon. Wea. Rev.,* **133,** 1261–1278.

——, K. Krüger, J. L. Sabutis, S. A. Sena, and S. Pawson, 2005b: The remarkable 2003–2004 winter and other recent warm winters in the Arctic stratosphere since the late 1990s. *J. Geophys. Res.,* **110,** D04107, doi:10.1029/2004JD005367.

——, and Coauthors, 2005c: Simulations of dynamics and transport during the September 2002 Antarctic major warming. *J. At-mos. Sci.,* **62,** 690–707.

Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.,* **28,** 1479–1494.

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

Melander, M. V., N. J. Zabusky, and A. S. Styczek, 1986: A mo-ment model for vortex interactions of the two-dimensional Euler equations. Part 1. Computational validation of a Ham-iltonian elliptical representation. *J. Fluid Mech.,* **167,** 95–115.

Müller, R., and G. Günther, 2003: A generalized form of Lait's modified potential vorticity. *J. Atmos. Sci.,* **60,** 2229–2237.

Nash, E. R., P. A. Newman, J. E. Rosenfield, and M. R. Schoeberl, 1996: An objective determination of the polar vortex using Ertel's potential vorticity. *J. Geophys. Res.,* **101,** 9471–9478.

O'Neill, A., 2003: Stratospheric sudden warmings. *Encyclopedia of Atmospheric Sciences*, J. R. Holton, J. A. Pyle, and J. A. Curry, Eds., Elsevier, 1342–1353.

——, and V. D. Pope, 1988: Simulations of linear and non-linear disturbances in the polar stratosphere. *Quart. J. Roy. Meteor. Soc.,* **114,** 1063–1110.

Palmer, T. N., 1981: Diagnostic study of a wavenumber-2 strato-spheric sudden warming in a transformed Eulerian-mean formalism. *J. Atmos. Sci.,* **38,** 844–855.

Plumb, R. A., 1981: Instability of the distorted polar night vortex: A theory of stratospheric warmings. *J. Atmos. Sci.,* **38,** 2514–2531.

Scott, R. K., and D. G. Dritschel, 2006: Vortex–vortex interactions in the winter stratosphere. *J. Atmos. Sci.,* **63,** 726–740.

Simmons, A., M. Hortal, G. Kelly, A. McNally, A. Untch, and S. Uppala, 2005: ECMWF analyses and forecasts of strato-spheric winter polar vortex breakup: September 2002 in the Southern Hemisphere and related events. *J. Atmos. Sci.,* **62,** 668–689.

Tung, K. K., and R. S. Lindzen, 1979: A theory of stationary long waves. Part I: A simple theory of blocking. *Mon. Wea. Rev.,* **107,** 714–734.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Waugh, D. W., 1997: Elliptical diagnostics of stratospheric polar vortices. *Quart. J. Roy. Meteor. Soc.,* **123,** 1725–1748.

——, and W. J. Randel, 1999: Climatology of Arctic and Antarctic polar vortices using elliptic diagnostics. *J. Atmos. Sci.,* **56,** 1594–1613.