ABSTRACT

Inspection of the potential vorticity (PV) distribution on isentropic surfaces in the lowermost stratosphere reveals the ubiquitous presence of numerous subsynoptic positive PV anomalies. To examine the space–time characteristics of these anomalies, a combined “identification and tracking” tool is developed that can catalog each individual anomaly's effective amplitude, location, overall spatial structure, and movement from genesis to lysis. A 10-yr winter climatology of such anomalies in the Northern Hemisphere is derived for the period 1991–2001 based upon the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). The climatology indicates that the anomalies are frequently evident above high topography and in a quasi-annular band at about 70°N, are long lived (days to weeks), and that their effective amplitude is typically 2 PV units (PVU) higher than that of the ambient environment. In addition, the derived climatologies and associated composites pose questions regarding the origin of the anomalies, detail their life cycle, and shed light on their dynamics and role as long-lived precursors of surface cyclogenesis.

1. Introduction

This study is concerned specifically with the occurrence and climatology of subsynoptic positive potential vorticity (PV) anomalies in the lowermost stratosphere. The nature of the anomalies is illustrated in the two panels of Fig. 1 that display a typical wintertime distribution of PV on the 320-K isentropic surface and a vertical cross section through one particular anomaly, respectively.

Inspection of the isentropic pattern (Fig. 1a) reveals a comparatively uniform background field of PV within the stratosphere, a sharp PV gradient at the stratosphere–troposphere interface [essentially the neighborhood of the 2 PV unit (PVU) isoline], and numerous PV anomalies embedded within the stratospheric portion of the domain. In addition, the cross section (Fig. 1b) indicates that such anomalies can possess a vertically coherent PV structure with significant accompanying thermal and flow signatures, both in the near and the far-field. In effect these PV anomalies are notable for their amplitude, number, spatial coherence, and distinctive horizontal and vertical spatial scales. The nature of the possible impact of the anomalies upon the dynamics of the tropospheric and stratospheric flow will depend upon the aforementioned factors, and also upon the location of the anomalies relative to the stratosphere–troposphere interface.

In this context, we comment on four possible dynamic effects of the anomalies. First note that some anomalies (labeled \(J_1\) and \(J_2\) in Fig. 1a) are located close to the 2-PVU isoline that demarks the location of the stratosphere–troposphere interface and the extratropical jet. The proximity of the anomalies to the jet indicates that they can initiate Rossby waves trapped to, and propagating along, the waveguide formed by the intense PV gradient that accompanies the jet (Schwierz et al. 2004). In turn, the resulting waves can and do influence deep baroclinic wave development that extends down to the surface.

Second, a strong and vertically coherent anomaly either well within the stratosphere or near the jet itself play a direct role in instigating surface cyclogenesis (Hoskins et al. 1985). Note that the anomaly in Fig. 1b has an amplitude in excess of 6.5 PVU relative to the ambient environment, extends from about 295 K (380 hPa) to 330 K (310 hPa), and possesses a strong vorticity signal that has a maximum just below the PV maximum and penetrates down beyond the 500-hPa level. Note that,
from the perspective of flow dynamics, a local PV anomaly equates to an in situ upper-level cyclonic vorticity maximum, and that its presence near a jet connotes a jet streak and/or a small-scale trough (e.g., Reiter 1969; Velden and Mills 1990). Hence, a suitably located PV anomaly can be related to, and form an integral feature of, the traditional type B class of cyclogenesis (Petterssen and Smebye 1971), and thereby it can serve as a possible precursor signal to surface cyclogenesis. In particular, it is anomalies that move over a region of lower-level baroclinicity that are more likely to influence surface development (cf. Hoskins et al. 1985; Sinclair 1994). Numerous case studies of the role of such precursors have been undertaken (e.g., Kleinschmidt 1950; Davis and Emanuel 1991; Davis 1992; Reed et al. 1992; Huo et al. 1999; Fehlmann and Davies 1999).

Third, the finescaled filamentary structure of the PV anomalies evident on isentropic surfaces, either at the stratosphere–troposphere interface or sequestered into the troposphere (see the examples labeled $S_1$ and $S_2$ in Fig. 1a), are intrusions of stratospheric air down into the troposphere. These filaments, often referred to as PV streamers (Appenzeller et al. 1996), are features of the mature phase of baroclinic development (Davies et al. 1991; Thorncroft et al. 1993; Martius et al. 2007), can serve as precursors to heavy precipitation events (Massacand et al. 1998), and are also often indicative of the early phase of localized irreversible stratosphere–troposphere exchange of mass and chemical constituents (Appenzeller et al. 1996; Stohl et al. 2003; Sprenger et al. 2007).

Fourth, the anomalies located well within the stratosphere serve, via their associated velocity fields, as coherent flow entities that stir the ambient atmosphere of the lowermost stratosphere. In effect the number, amplitude, and distribution of the anomalies portrayed in Fig. 1a point to their pervasive influence upon isentropic mixing in the lowermost stratosphere, and concomitantly pose questions regarding the origin of these predominantly positive anomalies (cf. Hakim and Canavan 2005).

The foregoing remarks highlight the multifaceted contribution of these PV anomalies to the dynamics of the lowermost stratosphere and troposphere. The overall aim of the present study is to develop and examine the climatology for these anomalies. There already exist some upper-level climatologies that pertain directly or indirectly to PV anomalies. These include climatologies of coherent structures in the vicinity of the tropopause (e.g., Hakim 2000; Hakim and Canavan 2005), upper-level vorticity maxima (particularly at the 500-hPa level), stratospheric streamers and cutoffs detected on isentropic surfaces (Wernli and Sprenger 2007), and stratospheric intrusions (Cai 2003). The link between these other climatologies and PV anomalies will be commented upon later. Here it suffices to assert that it is the PV pattern of the lowermost stratosphere that best captures the essence of the in situ PV anomalies.
Viewed in an isentropic framework, there is no obvious structural distinction between the PV anomalies embedded well within the stratosphere and those close to the stratosphere–troposphere interface. It is therefore reasonable to speculate that many of the PV anomalies have a preexistence deep within the stratospheric interior prior to approaching the jet and serve as cyclogenesis precursors (e.g., the anomaly featured in Fig. 1b). It follows that to capture the evolution and life cycle of PV anomalies, it is necessary to take account of their entire Lagrangian history.

In the present study, a tool is set out that can identify a PV anomaly on an isentropic surface and can track its evolution from genesis to lysis. Compilation of a climatology and composite of such anomalies over an extended period provides information on their amplitude, frequency, preferred regions of occurrence, and their movement. Repeating the calculation for different isentropic surfaces yields information on their distribution with height within the lowermost stratosphere.

The paper is organized as follows. In section 2, a description is provided of the dataset and the anomaly identification and tracking tool. This is followed in section 3 by illustrative examples of the tool’s capabilities. The subsequent sections detail the results based on a 10-yr winter climatology and include the lifetime frequency distribution (section 4), geographic distribution of occurrence on isentropic surfaces (section 5), characteristics of composite life cycles (section 6), and also the distinctiveness of intense anomalies (section 7). In the final sections a comparison is made with the results of previous studies and consideration is given to the more general implications of the results.

2. Datasets and methodology

a. Dataset

The dataset used is the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). This set provides global reanalysis data on a 6-h temporal resolution (0000, 0600, 1200, and 1800 UTC), has a spectral resolution corresponding to triangular truncation at T159, equivalent to about 125 km in the horizontal, and 60 hybrid vertical levels from the surface to 1 hPa. The vertical resolution near the tropopause is approximately 32 hPa. Full details of this dataset can be found in Uppala et al. (2005).

Here we operate with a decade of winter months (i.e., 1991–2001). Variables relevant for the climatology are computed from the ERA-40 primary variables and interpolated onto isentropic surfaces on a 1° × 1° longitude–latitude grid. The 3D PV field is calculated using a finite-difference scheme for the hydrostatic version of the full Ertel PV. Differences between PV calculated with this method and the fields provided by the ECMWF were found to be essentially minor (not shown).

b. Methodology

1) Feature identification

We choose to identify and define the anomalies (such as the one shown in Fig. 1b) by their “effective PV amplitude” (i.e., their amplitude relative to that of the ambient isentropic environment). This approach, in comparison with defining an anomaly relative to a climatological mean, running mean, or exceedance of a fixed threshold, is appropriate and desirable on several grounds:

1) The “effective amplitude” is a key factor in determining the dynamical significance of an anomaly relative to its environment. Indeed if the PV field were to be decomposed into the anomaly and its local backgrownd, then the principle of invertibility (Hoskins et al. 1985) can be invoked to provide a direct interpretation of the contemporaneous influence of the anomaly.

2) It is suitable to apply this relative measure within the lower stratosphere because the large-scale isentropic PV gradient is relatively weak, and the PV anomalies are themselves quasi-conserved material entities. This stands in contrast to the task of identifying isobaric local vorticity maxima and/or geopotential minima (Sinclair 1994).

3) The presence of diabatic processes should, in principle, be detectable via changes of the anomaly’s PV on an isentropic surface. This would not apply for an analysis conducted on isobaric surface, where temporal changes might merely be attributable to an air parcel’s adiabatic translation to another pressure level.

The identification procedure itself involves two steps. First, the procedure selects a localized isentropic PV structure with a maximum that exceeds 2 PVU, thereby focusing on stratospheric features. Figure 2 provides a schematic of a single maximum, along with an illustration of some parameters used in the identification procedure. The anomaly’s outer boundary is defined by a closed PV contour (referred to as the “edge PV”) that encloses an area of preassigned value. The identification of a finite domain of fixed size facilitates an evaluation of the anomaly’s shape and is integral to the tracking technique. The core/target area is set at 17 × 104 km2 (mesoscale), and this corresponds to the characteristic physical dimensions of analyzed anomalies. Moreover it is large enough to capture the form of the features, yet small enough for most isolated maxima to remain separated.
To find the edge PV, a search is performed around the local maximum in steps of 0.01 PVU. First an evaluation is made of the areas enclosed by the PV contours in the anomaly’s neighborhood. Then the contour enclosing the area closest in size to the target area is selected. The step size $\Delta PV$ is selected to be small enough that any further decrease would make no significant change to the area of the majority of (isolated) identified features. Note that, on the one hand, this technique excludes the possibility of monitoring the size evolution of an isolated anomaly and that, on the other hand, such an evolution would itself be dependent upon diabatic effects. [In this context we note that observations suggest a slight increase in the size of near-tropopause coherent structures with amplitude (see Hakim 2000).] An illustration of the sensitivity of the identification procedure to the area criterion can be found in Kew (2007).

In the second step, the amplitude of a given anomaly relative to its ambient environment is established by determining the local “background PV.” The approach chosen to determine the latter is motivated by the observation that the isentropic PV fields are relatively uniform on a synoptic-planetary scale. In the same way as above, we search around the anomaly for a contour to define the background that encloses a pre-specified area, this time of planetary scale. The area threshold is set large enough so that, because of the weak large-scale PV gradients within the lower stratosphere, even a moderate increase in the specified area would make insignificant difference to the PV value of the contour (see Fig. 2). In practice, the threshold value for the area was set to $10 \times 10^7$ km$^2$.

The amplitude of an anomaly is defined as the difference in PV between the local maximum and the local background. In principle, an amplitude criterion could also be introduced at this stage (cf. Bell and Bosart 1989; Dean and Bosart 1996) so that only strong anomalies are retained for examination. However, we do not apply such a constraint universally but require that the amplitude of an anomaly merely exceeds a specified threshold at some phase of its life cycle.

A clustering of the Northern Hemispheric isentropic PV field into domains with PV exceeding 2 PVU enables the main stratospheric body (the largest cluster) to be distinguished from stratospheric cutoffs located in the troposphere (the remaining clusters). In line with our earlier remarks, the identified structures are categorized into the following three classes based upon their proximity to the dynamical tropopause:

- **polar class** (class I)—well inside the stratospheric body,
- **near-tropopause class** (class II)—inside the stratospheric body but close to the stratosphere–troposphere boundary,
- **cut-off class** (class III)—outside the main stratospheric body.

If any point of an identified structure in the main stratospheric body is within a distance equivalent to $3^\circ$ latitude ($3 \times 110$ km) from the troposphere, the structure is classified to be “near the stratosphere–troposphere boundary” and interaction between the structure and the strong PV gradient at the boundary is deemed likely. The chosen distance is of the same order as the $\varepsilon$-folding distance of the vorticity field that we found to be associated with a typical PV anomaly.

It is anticipated that an anomaly’s impact and evolution will depend significantly upon the classification. Anomalies in the immediate vicinity of the jet will be affected by the jet’s horizontal and vertical shear. In contrast, the cutoff anomalies that are far removed from the tropopause should be less influenced by the shear and, because of their isolation, might be expected to have longer lifetimes, although they could be subject to radiative growth or decay (Forster and Wirth 2000; Cavallo and Hakim 2009).

### 2) Feature Tracking

An objective tracking technique is developed based upon the recognition that, to a first order, the flow in the lowermost stratosphere is quasi-adiabatic so that there is quasi-material advection of PV on isentropes. The procedure delivers a catalog of information including a register of merging and/or splitting, genesis and lysis, the geometric mean coordinates, the maximum PV, amplitude, isentropic wind velocity at the anomaly center, isotropy, and class (i.e., location relative to the stratosphere–troposphere interface). For the most part, only anomalies that persist for at least 1 day are considered, so as to eliminate transient features indicative of high-frequency variations or artifacts of the analysis process.
Here we set out the process that delivers this information. On the basis set out earlier, a PV anomaly is identified at time $t_n$ and advected forward for 6 h on an isentropic surface, using the analyzed winds. This provides the first estimate of the anomaly’s shape and location at the next time, $t_{n+1}$. The advection is undertaken in two time steps each of 3 h. The wind at the intermediate time step $t_{n+1/2}$ is calculated by averaging the $t_n$ and $t_{n+1}$ isentropic wind fields. Shape determination involves (i) compaction that excludes isolated points of the advected anomaly that lie at a radial distance of more than 2.5 standard deviations from the cluster center, and (ii) assigning to the anomaly a boundary (the convex hull) that encompasses the remaining points. (This process is akin to that of contour surgery, common to many contour advection schemes.)

An advected anomaly at $t_n$ is retained at the subsequent time step $t_{n+1}$, provided a significant overlap (10%, chosen after undertaking a comparison of objectively tracked structures with those tracked subjectively by inspection of the PV field) exists between its predicted spatial domain and that of an anomaly identified from the analysis fields at $t_{n+1}$ (Fig. 3a). Lysis is deemed to have occurred when more than one identified analysis-based anomaly at $t_{n+1}$ experiences significant overlap with a single advected feature (with the 10% overlap ratio threshold required for each identified fragment individually).

An indicator of an anomaly’s shape is constructed based upon its isotropy, and is given by the squared ratio of the diameter of a circle possessing the same area as the anomaly, to the anomaly’s longest linear dimension (great circle distance). For an elliptical form, this equates to the ratio of the minor to major axis length (i.e., the aspect ratio). For a nonelliptical shape, the measure is biased toward lower isotropy, since the major axis always takes the maximum length scale, whereas the effective minor axis length will be an average, rather than a maximum, of the shape’s extension in the perpendicular direction. Nevertheless, this method is capable of detecting any trend in aspect ratio over an anomaly’s lifetime.

3) Construction of climatology

The identification and tracking algorithms are run over the entire winter seasons covering the period 1 December to 28 February, for the 10 winters 1991/92–2000/01. The distributions for the individual years are combined to form the 10-yr climatology. These climatologies were compiled for four isentropic surfaces (310, 320, 330, and 340 K) that are located within the lowermost stratosphere in polar regions, and these surfaces typically intersect the tropopause in midlatitudes during winter.

3. Illustrative examples

To examine and illustrate the attributes and effectiveness of the identification and tracked procedures, and to underline the physical character of the associated
anomalies, we show the results of applying the method to two consecutive time steps of the ERA-40 dataset (at 1200 UTC and 6 h later at 1800 UTC 12 October 2000).

Figure 4 presents the 320-K PV fields for the two times (top panels) and the corresponding identified PV anomalies (bottom panels). Also superimposed in Fig. 4d are the anomalies advected from 1200 to 1800 UTC. During the 6-h period, the number of identified anomalies increased from 43 to 44, 11 structures underwent genesis (with 5 due to splitting), 10 anomalies experienced lysis, and there were 2 mergers. The period provides examples of most of the cataloged events (e.g., merger, splitting, pure genesis, lysis, and change of class), and these are noted briefly below.

- A splitting event is provided by the structure situated at about 80°N straddling the date line at 1200 UTC. At 1800 UTC, a splitting (labeled “S” in Fig. 4d) is detected, and is illustrated by the same advected contour encompassing the two identified fragments.
- A merger (“M”) is captured by two advected contours intercepting a single identified structure near 60°N, 80°W in Fig. 4d.
- A new structure (“N”) is identified at the tip of a streamer near 50°N, 160°W in Fig. 4d, in the absence of an advected contour. Inspection of the PV fields supports this identification that is associated with the evolution of a PV streamer with no local maximum at 1200 UTC, but with such a feature at 1800 UTC.
A lysis event (labeled “L” in Fig. 4d) occurs near 45°N, 130°E, where an advected contour is present but no identified structure.

A class change, from class III to class II, is illustrated by the engulfment of a previous cutoff back into the main stratospheric body near 70°N, 30°E (“C”)—see Fig. 4d.

Note that the scale of the identified structures matches the scale of the observed isentropic PV features and that Fig. 4d illustrates the close agreement in location and form between many of the advected and identified structures, including those close to the jet (the 2 PVU contour). This correspondence lends credence both to the isentropic advection technique and its application for the detection of splitting and merger. Clearly there can also be some events that involve very narrow filamentary structures that might not be successfully tracked with the procedure.

To illustrate the tracking procedure’s ability to trace the evolution of a single PV structure, consider an anomaly (Fig. 5) that played a significant role in instigating a surface cyclogenesis (Dirren 2002). The procedure identifies this precursor 5 days before cyclogenesis, located at 82°N, 130°W. It was generated at this location by the splitting of one structure into two components (event S of Figs. 4b,d) and then tracked for a total of 24 time steps (6 days).

Initially in the stratospheric interior (class I), the structure approached the tropopause (class II) and remained close to it until the cyclone it initiated reached maturity. During the class transition, the background PV decreased, and thereby contributed to an increase in the anomaly’s effective amplitude. This happened as the anomaly entered and intensified a planetary-scale trough and the local planetary-scale region defining the background PV successively split, resulting in the detection of a lower
background PV. The kink in the anomaly’s track from 1200 to 1800 UTC 17 October and a simultaneous decrease in the background speed occur when it approaches closely over the underlying pressure minimum of the surface cyclone. Entering the final phase of the life cycle, the anomaly recedes away from the stratosphere–troposphere interface (reassuming a class I categorization). At the time of its lysis, there were significant diabatic effects associated with the underlying occluded cyclone, and the anomaly experienced a strong decrease in amplitude.

The selected “shape” measure indicated that the anomaly underwent two complete cycles in isotropy during its life cycle. In harmony with this measure, inspection of the isentropic PV time series indicates that the anomaly first emerged as an elongated feature from a splitting event. It then sequentially became compact, elongated in the zonal direction on approaching the jet, compacted at the front of the trough taking the form of a head with an extended tail, and thereafter the head alone was retained by the identification procedure. The vacillation had a period of about 2.5 days.

4. Lifetimes

Figure 6 displays the frequency distribution of anomaly lifetimes on the 320-K isentropic surface. The curve’s slope corresponds reasonably well to a constant exponent for lifetimes in the time frame of 2–9 days. This indicates that within this time window the probability of survival is constant per unit time, and the linear regression \( R^2 = 0.949 \) shown in the figure gives a 6-h survival probability of 80.0%. For short life times, there is a lower survival rate—approximately 27% of structures on 320 K survive longer than 1 day. More pointedly, there are typically 10 anomalies per winter with a lifetime of 4 days. It is the existence of these isolated longer-lived anomalies that are of particular importance for medium-range forecasting.

The distributions for the other isentropic surfaces are qualitatively similar. However, features on the lowest level (310 K) have greater survival probabilities once they have survived for over a day. The longest-lived features were tracked on the 330- and 310-K surfaces and had lifetimes of 18.75 and 17 days, respectively.

5. Isentropic geographic climatologies

In this section, climatologies are presented of event and track densities of anomalies on the 320-K isentropic surface. Event density at a particular point is defined as the total number of events (identified anomalies) registered within a great circle distance equivalent to 5° latitude (560 km) from that point. Track density is defined as the number of tracks passing within a great circle distance corresponding to 5° latitude of a point.

Track density differs from the event density in that counts from slow-moving systems are reduced (a track is counted just once per grid point). In effect, the difference between the event and track density distributions is an indication of the mobility of the tracked systems, and the track density relates to the degree of consistency in track direction. Maxima in the track density distribution will therefore tend to occur closer to the jet than maxima in the event density distribution.

Only anomalies that exhibit a lifetime of greater than 1 day and that attain an amplitude of at least 1 PVU during their lifetime are considered here. In addition, to avoid geometric artifacts near the Pole, only the total population event density is portrayed north of 84°N. Results are presented successively for the total population, then genesis, lysis, merger and splitting events, and finally a comparison between the distributions on the different isentropic surfaces is given.

a. Total population statistics

The geographical distributions of event and track densities for anomaly events are shown in Fig. 7. A total of 118 334 events and 16 129 tracks are featured for the 10 winter seasons on 320 K.

Event density maxima form a quasi-annular band at about 70°N comprising primary separate maxima over northern Greenland and the Canadian Arctic and Alaska, and with secondary maxima over Siberia and the Barents Sea. Additional maxima are located between Mongolia and the Sea of Okhotsk at the entrance to the Pacific storm track and also over the southern part of the Kamchatka Peninsula. A significant number of these maxima occur over high terrain with the largest densities recorded over Greenland. In contrast, a significant
minimum occurs over the Arctic Ocean, and also low event frequencies prevail within the annular band at locations to the lee of the Rockies over Canada.

The track density pattern (Fig. 7b) resembles the form of the planetary-scale isentropic wintertime PV pattern (see e.g., Martius 2005) with the exception of the strong minimum over the Arctic Ocean. This suggests, in general, that the anomalies are advected with the mean flow. Again the high concentration of tracks north of the entrance regions of the Atlantic and Pacific jets suggests that the associated mobile anomalies are linked to or implicated in the creation of the surface storm tracks. Elsewhere over North America and in a band extending from Scandinavia into central Russia, the tracks are organized in a northwest–southeast orientation. This is in harmony with the observation (Lackmann et al. 1997) that many vorticity-related disturbances detected in the troposphere over the United States prior to rapid cyclogenesis take the form of coherent structures that move toward the coast within a northwesterly flow.

b. Genesis and lysis

Genesis and lysis event densities are presented in Fig. 8. The former are most prevalent over the Greenland landmass (particularly to the northwest), the Canadian Arctic, and Alaska, and there are secondary maxima near the entrance to the Pacific storm track and over the Norwegian Sea and Scandinavia.

Lysis regions coincide to a measure with the genesis regions, but are slightly more evenly distributed with longitude. There are also high lysis densities over northern Quebec, southern Greenland, and over northern Canada between Alaska and Hudson Bay. The overall lysis distribution bears some resemblance to the total population track density distribution, suggesting that either mobile anomalies are more amenable to lysis, or merely that lysis events are spread out downstream from the genesis regions by the steering of the mean flow.

The “genesis minus lysis” frequency distribution (not shown) shows that genesis tends to dominate over the eastern part of the Pacific and Atlantic Oceans and western sides of the American and European continents (i.e., at the end of the storm tracks), and contrariwise, lysis tends to dominate over the eastern side of the American and Asian continents and at the beginning of the storm tracks. The end of the storm tracks and upstream of orographic barriers are regions of divergent flow, frequented by streamers (Martius et al. 2008). Wave breaking and the redistribution of PV by lower-level diabatic processes can result in the meridional elongation and splitting of PV structures (Morgenstern and Davies 1999). Where splitting (production of two or more entities from a single structure) dominates merger, genesis is likely to dominate lysis.

c. Splitting and merger

The subsets of genesis events due to splitting and to merger are shown in Figs. 9a,b, respectively. Both distributions exhibit high frequencies over Greenland, the east Canadian Arctic, and Alaska, and there are secondary maxima across northern Europe and Asia. Merger events are more confined to regions of high topography, whereas their frequency is noticeably reduced over the ocean basins in the east and over large bodies of water (e.g., Hudson Bay, Davis Strait, and the Sea of Okhotsk).
There is a notable difference, particularly over North America and the Atlantic, between the orientation of tracks generated by splitting as opposed to merger events (Figs. 9c,d). Tracks initiated by merger exhibit stronger curvature over eastern Canada and Europe and tend to turn to the north over Greenland whereas, in general, the tracks resulting from splitting cross the Atlantic at the latitude of Iceland. The track density for splitting shows two prominent maxima corresponding to the entrance of the storm tracks, and this is consistent with the predilection for splitting to occur in regions of strong shear.

d. Interlevel comparison

Inspection of the event densities on different isentropic levels reveals the following:

- the Arctic maximum to the north of Greenland that is visible on 320 K in Fig. 7 becomes the most dominant feature at higher elevations;
- on 310 K, the meridional extent of the stratosphere is limited and a polar maximum becomes dominant;
- preferred regions poleward of 60°N are similar to those on 320 K but expand with height, whereas new maxima appear at lower latitudes over central Europe and close to the stratospheric-tropospheric interface over Asia;
- on 340 K, the maximum over the Canadian Arctic is no longer a prominent feature.

Inspection of the track densities indicate that from 330 to 340 K, the track alignment over central Asia shifts from being a zonal band near 60°N to a zonal band 20° farther south.

Genesis minus lysis frequency distributions exhibit significant variations with height. These include the following:

- on the 310-K surface the patterns resembling those on the 320-K surface, except that the Greenland tip is the most dominant genesis (lysis) location on the 310-(320-) K surface;
- on the 330-K surface the genesis region over western America becomes stronger and extends farther south;
- on the 340-K surface a pattern of alternating genesis and lysis dominance extends from North Africa to eastern China close to 40°N.

For mergers, the frequency distributions above 320 K show a noteworthy weakening of the Greenland maximum seen at 320 K, an increase in frequency over the central United States on the 330-K surface that gives way to a strong band of high frequencies extending from Alaska southeastward into the central United States and across the Rockies on the 340-K surface. For splitting, the frequency distribution above 320 K shows, relative to the 320-K feature, a reduction of the Greenland and Alaskan maxima, the emergence of a strong zonally oriented maximum over the central United States, increased frequencies over central Europe, and the appearance of a zonal band at 40°N emanating in northwest China and terminating over the Pacific storm track.

6. Composites

Composites are constructed of several characteristics of the anomalies tracked on isentropic surfaces. In particular, the effective PV amplitude, anomaly PV, the
isotropy, and background speed were examined as a function of absolute age since genesis (not shown) and also as a function of the age normalized by lifetime (Fig. 10). The interest here is on longer-lived anomalies and therefore a lifetime threshold was set at 4 days.

The composite for effective amplitude reveals a clear cycle. Maximum amplitudes tend to be reached during the first half of the lifetime corresponding to 2–3 days after genesis on the lower levels, and later for longer-lived features on the highest levels. On 320 K, the composite amplitude varies between 1.5 and 2.5 PVU with the initial amplitude being noticeably larger than the final amplitude. This is attributable in part to a substantial proportion of genesis events arising from merger and splitting, so that the newly created individual anomalies retain the significant amplitude of the parent anomalies. The maximum–minimum range of the composite amplitude over the lifetime is about 1 PVU at all levels.

The composite isotropy, viewed in absolute age, exhibits a small positive trend, and this is more pronounced on the lower 310- and 320-K isentropic surfaces. The composite, based upon normalized age, indicates that a reduction in isotropy tends to occur toward the end of the lifetime. In the mean, the isotropy is about 0.32, and this corresponds well to the aspect ratio of the anomalies seen in the case study.

Composites of the background speed (the speed at the location of the vortex) show a local minimum in background speed about 2 days after genesis, and this is coincident with the time of maximum effective amplitude. This suggests that an anomaly can exert a sustained impact upon its immediate surroundings at the time when the amplitude is largest. This inverse correlation between

---

**FIG. 9.** Splitting and merger. (a) Splitting event density at genesis, (b) merger event density at genesis, and track densities for tracks initiated by (c) splitting and (d) merger. Lifetime (1 day) and amplitude thresholds (exceeds 1 PVU within lifetime) apply. Densities show 10-yr winter-season totals over circular areas of 5° radius centered on each grid point. The mean dynamical tropopause is overlaid (thick black line).
speed and amplitude is stronger for changes in effective amplitude than changes in absolute PV. An anomaly would undergo strong increase in its effective amplitude as it approaches the stratosphere–troposphere interface or becomes embedded within a PV streamer due to the reduction in the ambient PV. In the first setting, the inverse correlation would prevail if the anomaly became the tropopause-level component of deep cyclogenesis resulting in a quasi-stationary cyclone, while for the second setting the inverse correlation is in harmony with the frequently observed quasi-stationarity of meridionally aligned streamers (e.g., Massacand et al. 1998; Hoinka et al. 2003).

7. Intense anomalies

a. Geographical distribution

An event climatology of anomalies attaining an amplitude of least 3.5 PVU within their lifetimes and surviving at least 1 day is shown in Fig. 11. Intense anomalies are most frequently found over the Mediterranean and Turkey, central United States, Greenland, the polar region, and aligned in a nearly zonal band across the Atlantic from Newfoundland to the United Kingdom. In contrast, intense events are comparatively rare over the Siberian Arctic and Asia, and are not as frequent over the Pacific storm track. For an even higher threshold value, the most favored locations are the central-west United States and the central Mediterranean. These locations coincide with the upstream flank of the Atlantic storm track and the location of Mediterranean and Alpine lee cyclogenesis, respectively.

There are some variations with elevation. On higher isentropic surfaces (340 K), the favored locations are the polar region, and again the Mediterranean and the central United States, and there remains a nearly zonal band between Newfoundland and the United Kingdom. On 310 K, the favored regions include Greenland, eastern Europe, and the polar domain. There is a modest increase in the frequency over Siberia and the entrance...
of the Pacific storm track, but these values remain less than those at the entrance to the Atlantic storm track. This difference at the entrance to the two storm tracks is a notable feature. Although the mean intensities on the 320-K surface (not shown) do have a local maximum over the Sea of Japan (>1.5 PVU), it is much lower than the maxima over the United States and Mediterranean (3 and 2.5 PVU, respectively).

b. Growth and decay

The effective amplitude of an anomaly can change because of a decrease in the ambient PV and/or an increase in that of the anomaly. Temporal composites of the amplitude, anomaly PV, and background PV are compared (Fig. 12) for tracks on 320 K passing through the aforementioned favored regions, to provide insight into which processes are responsible for the intensification. The major features are the following:

- over the Pole and Greenland, the gradual growth and decay in amplitude are highly correlated with changes in PV, while the background remains relatively constant;
- over the Mediterranean and the United States, growth is primarily associated with a reduction in the background PV, and the steady decay is primarily due to changes of the anomaly’s PV.

The initial reduction in the background PV for the Mediterranean and U.S. composites is consistent with the occurrence of strong troughs in these particular regions, and indeed the Mediterranean is a favored climatological region for the PV streamers (Martius et al. 2008). Moreover it can be seen (Fig. 11a) that the U.S. and Mediterranean intensity maxima lie very close to the mean position of the stratosphere–troposphere interface, and temporal composites for the anomaly class (not shown) indicate that the U.S. and Mediterranean anomalies do indeed tend to move toward the interface (class changes from I to II, and in the United States some also to class III) at the beginning of their lives.

The overall evolution of amplitude for the U.S. composite is markedly different to the evolution in the other composites. The peak amplitude tends to be reached later in the lifetime (70%, 3.5 days) and, although the growth rate is comparable, the decay is more rapid. The U.S. composite speed (not shown) exceeds that of the other regions in the latter half of the lifetime and there is a simultaneous decrease in isotropy. Many of these anomalies end their lives near the Atlantic storm track, where they possibly undergo radiative decay (especially if cutoffs), become sheared out (as suggested by the decreased isotropy) or are subjected to diabatic interactions with the synoptic storms they induce.

Finally, the influence of anomaly PV on the trend in amplitude for the U.S., Greenland, and polar composites suggest either that diabatic processes (e.g., radiative cooling) play a role in the local intensification (Cavallo and Hakim 2009) or that the changes in PV are non-physical results (a shortcoming) of data assimilation.

8. Relationship to other climatologies

It was noted earlier that several existing climatologies relate somewhat to the present climatology of PV anomalies in the lowermost stratosphere. These studies include a tropopause-level climatology of potential temperature anomalies (Hakim and Canavan 2005; Cavallo and Hakim 2009), and 500-hPa climatologies of vorticity maxima.
(Sanders 1988; Bell and Bosart 1989; Lefevre and Nielsen-Gammon 1995; Dean and Bosart 1996; Hakim 2000). A detailed intercomparision is given in Kew (2007). Here we indicate briefly the nature of the relationship of these climatologies to the present study.

There is a direct dynamical link because a PV anomaly in the lowermost stratosphere is associated at the tropopause level with a negative potential temperature anomaly, and with a vorticity maximum that can extend down to the midtroposphere (Fig. 1b). Hence, the aforementioned alternative climatologies can capture aspects of the associated PV climatology. For example:

- All show favored regions near the beginning of the Pacific storm track, near the Canadian Arctic (with the exception of Lefevre and Nielsen-Gammon 1995), and locally reduced frequencies over northwest Canada. None however reproduce an event maximum over Alaska.
- The 5-day composite in amplitude of cyclonic tropopause potential temperature anomalies of Hakim and Canavan (2005) is qualitatively similar, also peaking 2–3 days after genesis.
- Both Hakim and Canavan (2005) and Lefevre and Nielsen-Gammon (1995), report fixed 6-h survival probabilities (0.9 and 0.93, respectively). Our slightly lower survival rate of 0.8 is likely related to the smaller size and greater number of our identified structures.

However, there are caveats to this linkage. Identification of potential temperature anomalies at the tropopause or a related iso-PV surface is hampered by the existence of tropopause folds and spheroidal-like PV structures at the tropopause. Likewise, an approach based upon vorticity maxima on the 500-hPa surface, although avoiding the complications that arise from examining vorticity maxima located close to or embedded within jet-related large shear regions at the tropopause, could contain contributions attributable to more than one stratospheric PV anomaly and also to the tropospheric PV anomalies induced by diabatic effects.

A further caveat is related to airflow over broadscale topography. Thus, for example, quasi-steady airflow over such terrain will result in a decrease in the separation of isentropic surfaces and in the relative vorticity of air parcels surmounting the obstacle, and a possible increase of the same fields downstream due to the orographic generation of a Rossby wave. The key point is that these aforementioned major changes decay with height and need not be accompanied or associated with a PV anomaly aloft, and indeed the changes in stratification and vorticity will compensate to leave the Ertel PV of the air parcels unchanged on isentropic surfaces. In effect, the anomalies in the 500-hPa vorticity field need not always be related to the presence of PV anomalies in the lowermost stratosphere. A comparison of vorticity-maxima frequencies at 500 hPa over Greenland with the frequency of PV anomalies starkly underlines this difference: the aforementioned 500-hPa climatologies show that vortices tend to be absent over major orographic barriers or located downstream, whereas at the tropopause level, they tend to be vertically collocated with high topography.

A more comparable study to the present is the National Centers for Environmental Prediction (NCEP) winter-season climatology of positive PV anomalies on 330 K by Hoskins and Hodges (2002). However, they do not restrict PV to stratospheric values, and furthermore they remove the planetary wave background field. Consequently, the detection of anomalies is favored farther equatorward in the vicinity of jets and storm tracks. Common features include frequency maxima near Newfoundland, north Greenland, and the Pacific storm track entrance and a weak band of higher frequencies extending from Newfoundland toward Greenland. Greater intensities are likewise observed over the Mediterranean and North America and reduced intensities are observed.
over Asia. There are also locally raised intensities in the vicinity of Greenland and over the Pole.

9. Further remarks

In this paper, a combined identification and tracking tool has been developed and utilized to study the distribution, character, and evolution of subsynoptic PV anomalies on isentropic surfaces in the lowermost stratosphere. The tool identifies an anomaly as local maximum, and its effective amplitude is defined relative to the surrounding “background” PV. The tracking component of the tool involves first isentropically advecting the anomaly with the analyzed wind field, and then employing a multiple criteria procedure to associate the resulting advected anomaly at a later time either directly with an anomaly identified in the analysis field at that later time or categorizing the matching process as indicating the demise, merger, or splitting of the original anomaly.

The results derived with the tool, when applied to 10 successive winter seasons, provide climatologies of event and track densities, genesis and lysis, merger, and splitting on isentropic levels spanning 310–340 K in the Northern Hemisphere, as well as yielding information on the composite characteristics of anomalies. The significant results include the following:

- the predilection for the anomalies to occur within a quasi-annular band at around 70°N, and over high topography;
- qualitatively similar patterns for merger and splitting events, with merger events being more confined to high topography and less prevalent over oceans;
- tracks initiated by merger events possessing higher curvature than the tracks initiated by splitting events, at the location of the climatological planetary-scale troughs;
- lifetimes for anomalies that last more than 1 day exhibiting a constant subsequent 6-h survival probability of about 80% independent of lifetime, with approximately 10 anomalies per season surviving for 4 days;
- along-track composites show that the anomalies reach their largest effective amplitude during the first half of their lifetimes with the PV decreasing by typically 0.2 PVU day⁻¹ after 3–4 days, and isotropy showing a small positive trend over the life cycle;
- intense anomalies (− effective amplitudes > 3.5 PVU) found most frequently over the Mediterranean, Greenland, the Pole, and in a band across the United States and the Atlantic, with significantly higher frequency at the entrance of the Atlantic, as opposed to the Pacific, storm track.

The ubiquity, longevity, and effective amplitude of the PV anomalies in the lowermost stratosphere serve to underline their significance both for the isentropic mixing of this region and for their potential potency as instigators of Rossby waves at the stratosphere–troposphere interface and as precursors of surface cyclogenesis. Their extended lifetime allied to their role as wave instigators and as precursors of cyclogenesis, helps extend back in time the phenomenological causal change that culminates in major flow development (i.e., Rossby wave breaking and cyclogenesis). Examination and exploitation of such a causal chain is aligned to the objectives of The Observing System Research and Predictability Experiment (THORPEX) program. Future work could focus on interlevel interactions and coherent structures that have a large vertical impact.

Acknowledgments. The authors appreciate the access to the ERA-40 dataset granted by MeteoSwiss and thank Greg Hakim and one anonymous reviewer as well as Ryan Teuling for helpful comments on the manuscript.

REFERENCES


