

Methods of Calculating Transport across the Polar Vortex Edge

A. H. SOBEL AND R. A. PLUMB

Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, Massachusetts

D. W. WAUGH

Cooperative Research Center for Southern Hemisphere Meteorology, Monash University, Victoria, Australia

(Manuscript received 5 July 1996, in final form 15 January 1997)

ABSTRACT

Existing quantitative calculations of material transport across the stratospheric polar vortex edge are difficult to interpret. This is because what is actually calculated has not been clearly shown to be irreversible transport, because of ambiguities inherent in defining the vortex edge, and (relatedly) because the uncertainties in the various sorts of calculations have not been quantified. The authors discuss some of the conceptual and technical difficulties involved in such calculations. These typically use a tracer coordinate, so that an air parcel's "position" is defined as a function of some tracer that it carries. Also examined is the sensitivity to noise of a method that has been used in several prior studies, which the authors call the "contour crossing" method. When contour crossing is implemented with no explicit threshold to discriminate noise from signal, a realistic amount of noise in the tracer data can cause apparent transports across the vortex edge in the range of ten percent to several tens of percent of the vortex area per month, even if the true transport is zero. Moreover, contour crossing does not discriminate between dynamically driven transport and that due to large-scale nonconservative effects acting upon the tracer used to define the coordinate. The authors introduce a new method, which is called the "local gradient reversal" method, for estimating the dynamically driven component of the transport. This method is conceptually somewhat similar to contour surgery but applies to gridded fields rather than material contours. Like contour crossing, it can thus be used in conjunction with the reverse domain filling advection technique, while contour surgery is used with contour advection or contour dynamics. Local gradient reversal is shown to be less sensitive to noise than contour crossing.

1. Introduction

A number of studies in recent years have attempted to determine directly, through kinematical calculations, the degree to which the stratospheric polar vortices are isolated from their environments. These have used winds and tracers from both observational analyses (Vaughn et al. 1994; Plumb et al. 1994; Dahlberg and Bowman 1994; Manney et al. 1994; Norton and Chipperfield 1995) and general circulation models (Eluszkiewicz et al. 1995) and have not yielded uniform results.

Unless otherwise specified, from this point on our mentions of these studies will refer to results on or near the 450-K isentropic surface, in the Arctic. Dahlberg and Bowman (1994, hereafter DB94) obtained relatively large values for both inward and outward transport, whereas Vaughn et al. (1994, hereafter W94), Plumb et al. (1994, hereafter P94), and Eluszkiewicz et al. (1995,

hereafter E95) obtained considerably lower values for both of these but especially for the inward transport. The values of Manney et al. (1994, hereafter M94) are at times closer to one and at other times closer to the other of these extremes (where their results translated into percent per month were obtained from E95). W94, P94, and DB94 all used National Meteorological Center (NMC, now the National Centers for Environmental Prediction) analyzed winds from the same periods, yet the first two obtained quite different results from the third. This suggests that differences between different datasets are not the source of the disagreement. It is our view that a large part of the difficulty derives from an ambiguity as to how transport across an interface such as the vortex edge can best be defined. This difficulty is worth addressing directly, particularly since essentially the same one arises if one wishes to study transport across the subtropical edge of the stratospheric surf zone (Polvani et al. 1995), the midlatitude tropopause (Danielsen 1968; Wei 1987; Lamarque and Hess 1994), or other atmospheric or oceanic "transport barriers."

In accord with our belief that some of the most critical difficulties involved in such transport calculations are theoretical ones, the next section is devoted to a dis-

Corresponding author address: Adam Sobel, University of Washington, Department of Atmospheric Science, Box 351640, Seattle, WA 98195-1640.
E-mail: sobel@rossby.mit.edu

discussion of theoretical issues. Section 3 contains a description of the advection algorithms used in this study, and section 4 applies these algorithms to the particular data used throughout the paper, in the process providing an overview of the vortex evolution during the month-long period studied. In section 5 we discuss what we call the “contour crossing” method, used by DB94, M94, and E95. This method is a diagnostic for use with “domain filling” parcel advection calculations, either forward or reverse. It is straightforward to implement and is defensible on physical grounds. However, we test its sensitivity to noise in the analyzed tracer fields and find it to be large. This sensitivity manifests itself by the appearance of spurious transport, so that the effect of noise is always to increase the apparent transport, both inward and outward, at the vortex edge. We use an idealized example to illuminate the mechanism by which this occurs and crudely estimate its effect for an amount of noise that one may expect the analyses to exhibit. The resulting estimate is in agreement with the result of our sensitivity test.

In section 6 we introduce the “local gradient reversal” method of performing transport calculations. This is a diagnostic for use in conjunction with reverse domain filling (RDF). Local gradient reversal and RDF together may be thought of as analogous (though different in some basic respects) to contour advection with surgery (CAS) with the surgery implemented only at the end of the calculation in order to calculate net transport, as done by W94, P94, Polvani et al. (1995), and Waugh (1996). The phrase “the CAS technique” will be used hereafter to mean this method of obtaining quantitative transport estimates, as opposed to the basic CAS method itself (in which surgery may be applied during the run as well as at the end). The primary difference between local gradient reversal and the CAS technique is that the former combination applies to two-dimensional gridded fields, while the latter applies to one-dimensional material contours in a two-dimensional space. Other differences follow from this. Both methods have the advantage (compared to contour crossing) that the analyzed tracer fields are used only as initial conditions and not thereafter, substantially reducing the sensitivity of the techniques to noise in those fields. We show this by testing local gradient reversal on the same case on which we tested contour crossing. In section 7 we summarize and discuss our results.

2. The problem

It is by now almost universally agreed that the edge of the polar vortex forms some kind of barrier to transport, so that the air inside is to some extent isolated. The question we wish to address is how we can best determine, given the accuracy of the available data, just how isolated it is. Especially when working with coarse meteorological data, it seems most reasonable to define a vortex edge *region* of finite width, rather than an in-

finitesimally thin interface separating vortex from surf zone air. A sensible scheme for doing this has recently been developed by Nash et al. (1996). Hereafter we shall use a definition roughly consistent with theirs: the words “vortex edge region” should hereafter be interpreted as meaning the belt of high potential vorticity (PV) gradients in the polar region. The transition from high to low PV gradient is usually quite marked so that this is fairly unambiguous. However, in what follows, sufficient information will be provided so that the reader can interpret our calculations without accepting our definition, or indeed any definition, of the vortex edge.

If we do define a vortex edge region, the distinction between vortex and surf zone then subdivides into one between inner vortex, edge region, and surf zone. Ideally, we would like to know all the details of the transport of air within and between these three regions. More generally, if we can use a coordinate such that the vortex is represented by a particular region of “space,” then we would like to quantitatively determine the transport of material from each point in that coordinate to each other point. One could then objectively identify transport barriers as regions or surfaces across which transport is particularly inhibited.

The primary difficulty is that, in the case of the polar vortex, traditional Eulerian coordinates, in which transport is easily defined and evaluated, are inappropriate. This is because, despite the difficulties that arise when one tries to discuss the process in precise language, intuition (combined with observational evidence) tells us that the entire vortex may move around while remaining coherent and without exchanging fluid with its exterior. We do not wish to call this transport. Hence it is common in stratospheric studies to define the vortex in a Lagrangian sense. That is, we associate the vortex with a certain set of “tagged” air parcels. In practice this is typically (and in this paper) done by picking a tracer, such as PV or a chemical species, and defining a coordinate based on that tracer. The tracer contours are then coordinate surfaces, and the vortex is made up of all parcels carrying values of the tracer within a certain range. The problem is then that if the tracer is truly conserved following the motion, a parcel once defined as in the vortex will always be defined as in the vortex, no matter where it goes, and “transport” cannot occur. We may get around this by associating transport with nonconservative processes acting on air parcels, changing their tracer values from “vortex” to “nonvortex” values and vice versa. The recent theoretical work of Nakamura (1995) makes this idea more precise.

In a follow-up study, Nakamura (1996) posits an explicit diffusion that can change parcels’ tracer values and hence allow them to cross the tracer contours. The resultant transport may still be considered dynamical in nature since dynamical processes act to produce small-scale features, hence large gradients, which increase the effectiveness of the diffusion. This approach is attractive for its elegance and rigor, but using it to

obtain quantitative results requires the specification of a small-scale diffusivity. At present, it is unclear how best to choose this number. Perhaps more to the point, it seems to us that for the purpose of diagnosing the mass transport at the vortex edge in a wave-breaking event, one ought not need to know the small-scale diffusivity (assuming it is small), since the transport is driven by large-scale dynamics. This dissonance with our intuition has led us to search for an alternate way of formulating the problem.

The critical problem is to separate reversible from irreversible transport. The power of Nakamura's method lies at least partly in that the explicit diffusion causes his irreversibility to be *thermodynamic*, hence rigorously defensible on very basic physical grounds. However, the results of Carnevale (1982) constitute a powerful defense for the statement of Pierrehumbert (1991) that the "chaotic mixing" of vorticity is "macroscopically irreversible despite the reversible Euler equations in *precisely* the same sense that molecular diffusion is irreversible despite the reversibility of Newtonian mechanics" (italics in original). The argument applies to vorticity or PV, despite the fact that these are dynamically active quantities that affect the flow mixing them, because of the scale selectivity of the inversion relating the PV distribution to the flow (e.g., Hoskins et al. 1985). This ensures that sufficiently small vorticity or PV features (compared to the largest ones present) are effectively as passive as a chemical tracer. This passivity is the reason narrow vortex filaments do not tend to roll up into secondary vortices when they are near the parent vortex from which they were shed (Dritschel 1989b; Waugh and Dritschel 1991; Dritschel and Polvani 1992).

On these grounds, it should be possible to separate reversible from irreversible transport using a *fluid-dynamical* criterion. This can be done by taking the formation of PV features smaller than some cutoff scale to represent irreversible transport with respect to a set of PV contours from which those features have somehow been removed. The technique of contour surgery (Dritschel 1989a) performs this removal in a very literal sense. In local gradient reversal, we are proposing a technique for performing an analogous separation, less literally, on a gridded field. We assume that diffusion or some other dissipative process acts at scales smaller than we wish to or are able to resolve. We then make an additional assumption that, since the scale cascade is one-way, we can predict in advance which features will eventually be dissipated. As long as the dissipation acts only at some scale well below the cutoff scale, its nature is otherwise irrelevant.

Thus, a sufficiently large secondary vortex will be considered to remain part of the main vortex until such time as it is stretched to below the cutoff scale. This is tantamount to defining transport to occur not at the time of detachment from the main vortex, nor at the later time when it has been stretched to filaments small enough for dissipation to act upon them (Nakamura's

approach), but at the intermediate time of cascade below the cutoff scale. By the classic statistical-mechanical definition of irreversibility, as noted in the pertinent discussion of McIntyre and Palmer (1985), any secondary vortex formation is an irreversible event, suggesting the time of detachment as the key moment. However, we have no way of predicting, at the time of detachment, the ultimate fate of a large secondary vortex. Our definition is not, however, entirely inconsistent with Nakamura's but simply involves the additional assumption of a one-way cascade. If this assumption is correct, our formulation and Nakamura's should clearly give the same results for the total transport due to a given wave-breaking event.

This definition of irreversibility does not allow for a reverse cascade in the form of the roll-up of filaments below the cutoff scale into vortices of larger size. However, at least in the context of transport across the vortex edge, any secondary vortex with circulation much smaller than that of the main vortex will be stretched back into filaments rather than merge with the main vortex (Dritschel and Waugh 1992). It seems inconceivable that filaments thinner than any reasonable cutoff scale could roll up into vortices of sufficient size to survive such an interaction. In the examples shown below, secondary vortices of appreciable size do not form, and so these issues do not arise.

Large-scale nonconservative processes, such as radiation in the case of PV, can also transport material in a tracer coordinate. Contour crossing does not distinguish this component from the dynamically induced component discussed above, while local gradient reversal and the CAS technique estimate only the latter. It is certainly desirable to estimate both components, but there is justification for separating them. In general, they are simply different physical processes and one ought to understand them as such. For an example, consider a case in which the vortex size is statistically steady over some period, and the radiatively induced PV tendency is nonnegative throughout the extratropics. Nash et al. (1996) present a two-dimensional equation for the evolution of the area of PV contours that shows that, if friction and horizontal flow divergence are negligible, the dominant balance in this case must be between stripping of material from the vortex edge—the dynamical transport—and radiation. Edge stripping removes material from within PV contours in the edge region, while radiative forcing increases PV (by assumption) just outside those contours, allowing the contours to move outward so that the material is replaced. In this steady-state situation, if we find a transport barrier where the dynamical transport is very small, we know that the radiative transport must be small in the region of that barrier as well, so knowing one implies knowledge of both. If we do not distinguish the two components, however, we cannot tell whether the steady state is due to the near vanishing of both terms or whether both terms are large but canceling.

3. Advection algorithms

Kinematic, or trajectory-based, transport calculations have two steps. First, one performs an advection calculation. Contour advection (CA) and reverse domain filling (RDF) are two algorithms that serve this purpose. Both take some representation of a tracer field and cause it to evolve in time due to the effect of advection by a prescribed wind field. An additional calculation is then required to obtain a transport estimate from the time-evolved field. Contour crossing, local gradient reversal, and contour surgery based techniques are examples that serve this second purpose.

In this section we briefly review the contour advection and reverse domain filling algorithms. The advection algorithms generate fine-scale features from initially low-resolution tracer data, using winds which are also low resolution. The justification for doing this comes from a numerical study (Waugh and Plumb 1994) and from the agreement that such calculations have obtained with high-resolution aircraft tracer measurements (Waugh et al. 1994; Plumb et al. 1994; Newman et al. 1996). Each technique requires as input a tracer field at a single initial time, and a temporal sequence of wind fields covering the entire period covered by the calculation. We will call these tracer and wind fields the “analyzed” fields, as distinct from the tracer field produced from them by the advection calculation. Each method is here taken to be two-dimensional, though the domain filling techniques are straightforwardly extended to three dimensions.

In the domain filling technique (Fisher et al. 1993; O’Neill et al. 1994), the tracer field is initially represented by its value at a set of regularly spaced points throughout the domain, usually with significantly higher resolution than that of the analyzed tracer field. At each point a hypothetical parcel is placed. Each parcel is then advected according to

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t), \quad (1)$$

where \mathbf{x} is the parcel’s position and $\mathbf{u}(\mathbf{x}, t)$ is the prescribed wind field. The tracer is assumed to be conserved over the duration of the calculation. Hence, at the end one has a new arrangement of the parcels. Assuming that the tracer is conserved following the motion, this may be interpreted as a new configuration of the tracer field. The difficulty is that the parcels will no longer lie on a regular grid, making further analysis awkward. The RDF technique (Sutton 1994; Newman and Schoeberl 1995; Schoeberl and Newman 1995) addresses this problem by placing the parcels on the regular grid at the final time, and then integrating (1) *backwards* in time from the final to the initial time. The parcels are then labeled by values of the tracer field interpolated to their computed positions at the initial time. Again these tracer values are assumed conserved throughout the advection period. Hence, assigning these

values to the same parcels at their positions on the regular grid at the final time gives a high-resolution tracer field on a regular grid at that time.

In the CA technique (Waugh and Plumb 1994; Norton 1994; Dritschel 1989a), the initial tracer field is represented by a finite number of isopleths, or *contours* of constant tracer. These contours are then in turn represented by a large number of points, which are then advected according to (1). New points are added and removed as needed by an interpolation scheme, so that the contour continues to be well represented as it is stretched and folded. A technique known as “contour surgery” (Dritschel 1989a) may be used to remove fine-scale features, though this technique has not been used in any calculations in this paper.

The primary difference between the CA and RDF methods is that the former represents a two-dimensional field by a set of one-dimensional curves, whereas the latter maintains the two-dimensionality by using the more standard grid representation. Discussions of the close relationship between the two techniques may be found in Yang (1995) and Schoeberl and Newman (1995). This paper focuses on techniques of obtaining transport estimates from the results of RDF calculations. We focus on this because we feel that gridded fields are easier to work with and discuss than contours, because they do not require a decision in advance about which PV contours are important (since the entire field is represented) and because methods associated with domain filling techniques can be more easily extended to three dimensions. We do not implement the latter extension here, but one may anticipate that it will be desirable to do so in the future.

4. Case study

The techniques examined in this study will be applied to a case study of the arctic polar vortex during a single month, broken into two consecutive 15-day periods. Before performing any quantitative transport calculations, we present RDF and CA calculations in this section to illustrate the vortex dynamics during these periods. Though we will use only RDF for our transport calculations, we also use CA here for the very clear qualitative picture that it can provide of the Lagrangian kinematics.

We have chosen the 15-day periods of 16–31 January and 1–16 February 1993, on the 450-K isentrope. Our calculations are two-dimensional and assume that the fluid motions are isentropic. Most recent calculations of radiative heating rates yield cooling of no more than 1 K day⁻¹ in the regions and times of interest here (e.g., Rosenfield et al. 1994; Strahan et al. 1994; Eluszkiewicz et al. 1996); hence we expect cooling of 15 K or less over the duration of each calculation. The 3D analysis fields from which our isentropic 2D fields were obtained suggest only small variations in the (horizontally) large-scale wind field on a vertical scale corresponding to 15

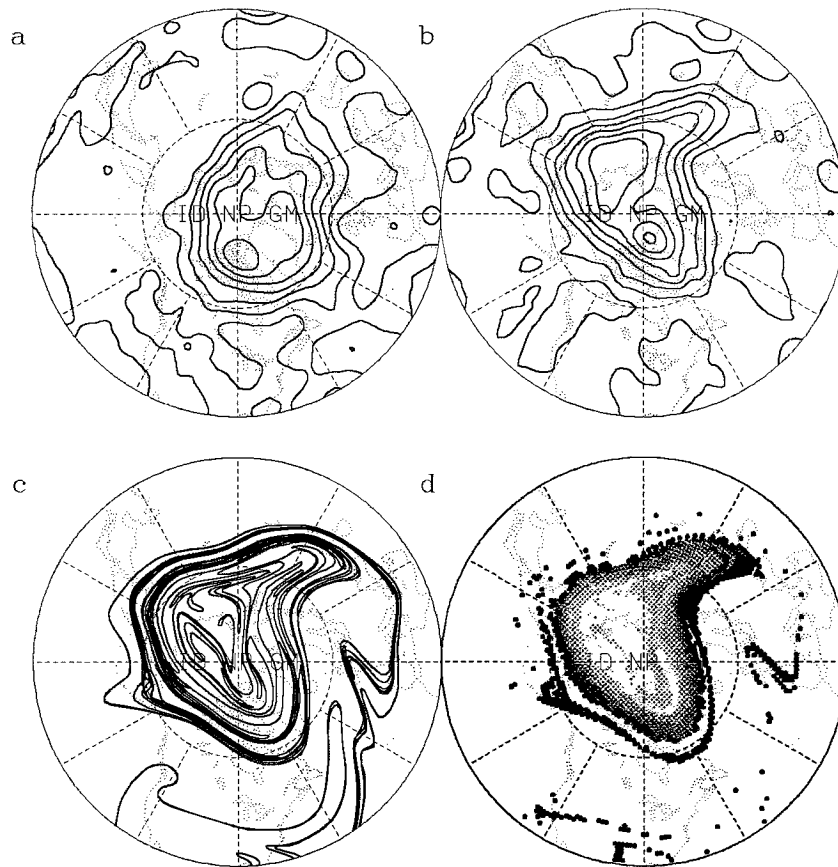


FIG. 1. (a) NMC analysis of PV on the 450-K isentropic surface from 16 January 1993. Contour interval is 5 PVU, 1 PVU = 10^{-6} K m² s⁻¹ kg⁻¹. (b) As in (a) but for 31 January 1993. (c) Result of a contour advection calculation initialized on 16 January 1993 and run for 15 days so that it is valid on 31 January 1993. Potential vorticity contours used are those from (a) excluding those with PV less than 20 PVU [the largest obviously closed contour in (a)] and greater than 40 PVU, so that the region represented is the region of the largest PV gradients. NMC-analyzed winds were used to advect the contours. (d) Reverse domain filling (RDF) calculation for the same period as (c) and using the same winds and PV. Grayscale color scheme delineates the same PV intervals as in (c); note that white is double valued, representing PV greater than 40 PVU and less than 20 PVU. All plots are valid at 1200 UTC on the given day.

K. Thus the adiabatic approximation is reasonable, if the calculations are considered as representing the motion of a layer of fluid of finite depth. In discussions of DB94, M94, and E95, we will refer only to their conclusions as they apply to the horizontal (i.e., isentropic) component of the transport, though the latter two performed calculations in three dimensions. The NMC observational analyses of wind and PV have been used here, whose grid spacing is 2° latitude by 5° longitude. The NMC analyses of temperature are described by Nagatani (1990) and Newman et al. (1990). NMC winds are derived using the method of Randel (1987), and PV is then derived from the winds and temperatures.

Figures 1a and 1b show the NMC analyses of PV on the 450-K isentropic surface on 16 and 31 January 1993, respectively, while Figs. 2a and 2b show the same analyses for 1 and 16 February 1993. Figures 1c and 2c show the results of 15-day CA calculations ending on

31 January and 16 February. The CA code is the same one described in Waugh and Plumb (1994). All runs were done without surgery. The large-scale patterns at the end of each run are in good agreement with the analyses for the corresponding days, as was seen to be the case in the comparisons of W94.

In the first period, the vortex is very stable. Although it is not obvious from this figure, the filament extending well into the surf zone consists only of the outermost contour [20 potential vorticity units (PVU), 1 PVU = 10^{-6} K m² s⁻¹ kg⁻¹]. The 25-PVU contour (corresponding to roughly 64° equivalent latitude) is drawn only into an extremely thin filament that wraps tightly around the edge. The 30-PVU contour (68° equivalent latitude) forms an even tinier filament at the end of the period, which remains confined by the 25-PVU contour, the latter not being itself drawn into this filament. These CA results form a strong, if subjective, basis for the

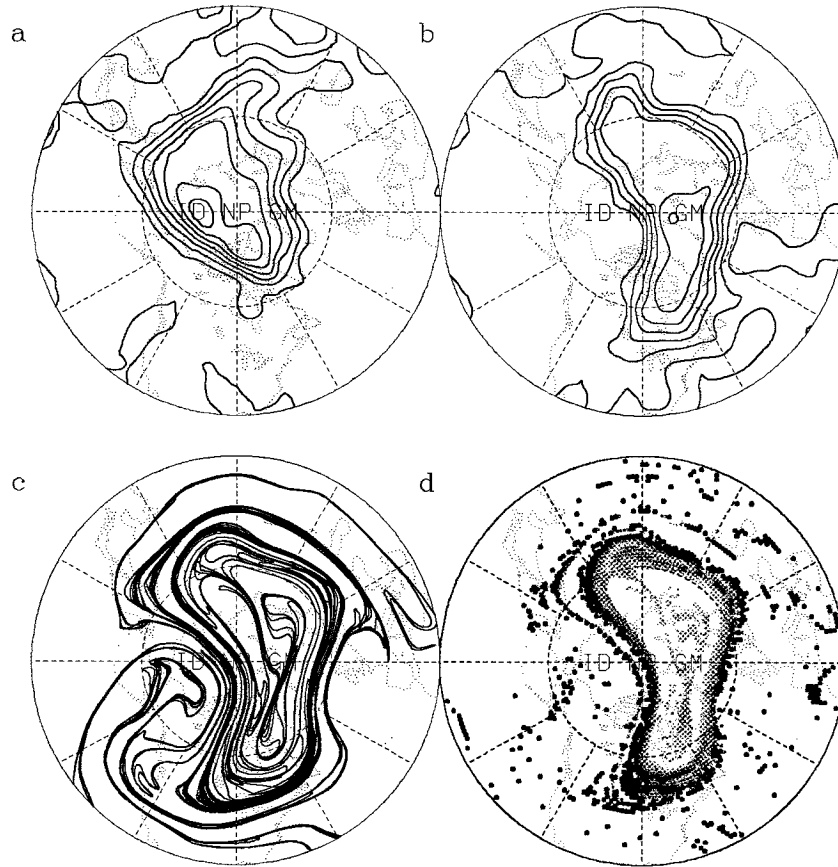


FIG. 2. Same as Fig. 1 but for the period from 1 to 16 February 1993.

statement that transport of material with PV greater than 25 PVU into the surf zone during this period is zero or very nearly so and transport of surf zone air into the vortex is zero. The situation is more complicated in the more disturbed second period. Among other things there is a clearly visible, if relatively small, intrusion into the vortex in the second half of the period, so in this case we would say that there has been some transport into as well as out of the vortex.

Figures 1d and 2d show the results of RDF calculations performed for the same periods, and with the shading of the points demarcating the same PV contours, as the CA calculations in Figs. 1c and 2c. Note that the color white is double valued to improve the contrast. The patterns in the CA and RDF results are identical down to the resolution of the RDF grid, which here is equal area with spacing in both longitude and latitude equivalent to 1° latitude (about 110 km). This should not be surprising, since the same winds and tracer fields were used and the two methods of calculation are essentially similar (Yang 1995; Schoeberl and Newman 1996). In this case exactly the same algorithm, a fourth-order Runge–Kutta scheme, is used to solve (1) in both codes.

5. Contour crossing method

a. Algorithm

The contour crossing method is used to obtain estimates of the transport across various tracer contours from the output of either a forward or reverse domain filling calculation. In addition to requiring the initial tracer field and sequence of wind fields as described above, in order to use the contour crossing method one also needs an analysis of the tracer field at the final time. Then at the final time of the calculation, each particle has *two* tracer values, which in general will be different. The first is the tracer value that the particle had at the initial time, which has simply been carried by the particle throughout the calculation. Call this the “advected” value. The second is the value obtained by interpolating the analyzed tracer field at the final time to the particle’s final position. Call this the “analyzed” value. The contour crossing technique uses essentially the following line of reasoning. At the final time, the advected value is that which the particle would have had if the tracer were truly conserved following the particle’s motion. The analyzed value is its true value (this implies a high degree of trust in the analyzed fields). The dif-

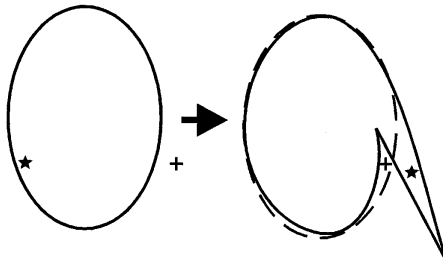


FIG. 3. Schematic diagram illustrating the principle behind the contour crossing method. The oval on the left represents a contour of constant tracer from the analyzed tracer field at the initial time. The tracer field is mapped forward to the final time using an advection algorithm, resulting in the configuration indicated by the solid black contour on the right. The dashed contour on the right represents the analyzed field at the final time, which contour crossing interprets as the "true" field. The parcel represented by the star is thus considered to have been transported out of the contour, while that represented by the plus is considered transported into it. Summing the areas of all such parcels gives estimates of the total outward and inward transport, respectively.

ference must represent the action of nonconservative processes on the air parcel that the particle represents. Since the particle's tracer value has changed by the action of these processes, we may say that it has "crossed" the tracer isopleths (contours) whose values are between the advected and analyzed values. This is illustrated schematically in Fig. 3. Counting the number of parcels that cross a given contour then provides an estimate of the transport of material across that contour. The transports "into" and "out of" the contour, as opposed to just the net transport, are represented separately by this technique.

The coordinate can be most clearly labeled by converting the tracer value of each contour to its "equivalent latitude," which is the latitude of the zonal circle that encloses the same area as the contour (Butchart and Remsberg 1986; Nakamura 1995). That is, if a parcel's tracer value is q , then its equivalent latitude is

$$\phi_e = \sin^{-1} \left(1 - \frac{A(q)}{2\pi r_e^2} \right), \quad (2)$$

where r_e is the radius of the earth and $A(q)$ is the area enclosed by the contour of constant tracer whose value is q . Note that if the contour is multiply connected, the area is computed by summing over all "islands." The "vortex following coordinate" of Norton (1994) is closely related but not identical.

In practice, the parcels in a domain filling or reverse domain filling calculation need not be acted upon by any true nonconservative processes in order to appear to cross contours. This is because the analyzed tracer fields have finite, and often low, resolution. If, as is typical, features with scales too fine to be represented in the analyses are generated in the advected fields, then at least some of the particles making up those features will appear to change their tracer values, even though in reality they may not. In addition, apparent tracer nonconservation may be caused by the presence of noise in the analyzed tracer fields. The consequences of these effects will be discussed further in what follows.

b. Results: Scatterplots and probability densities

The contour crossing method allows one to calculate not only how much material crosses each contour, but by how far, in equivalent latitude space. Manney et al. (1994) call this PV dispersion, while DB94 display essentially the same thing as "exchange flux curves." Here the same information is displayed in a particularly compact form. Neither DB94 nor M94 converted their results into equivalent latitude space, instead using PV directly as a coordinate. In such a coordinate, large-scale diabatic effects can cause parcels to move and the poles are not fixed; for these reasons, among others, we prefer to convert results explicitly to equivalent latitude space.

First, in Figs. 4a and 4b, scatterplots are shown of

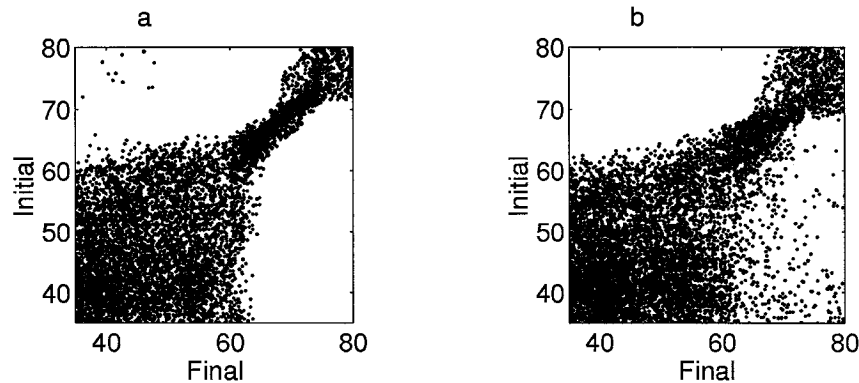


FIG. 4. (a) Scatterplot of initial vs final equivalent latitude for the particles in the RDF calculation in Fig. 1d, where the NMC-analyzed PV is used to construct the equivalent latitude coordinates on the initial and final days. (b) As in upper left, but for the RDF calculation in Fig. 2d.

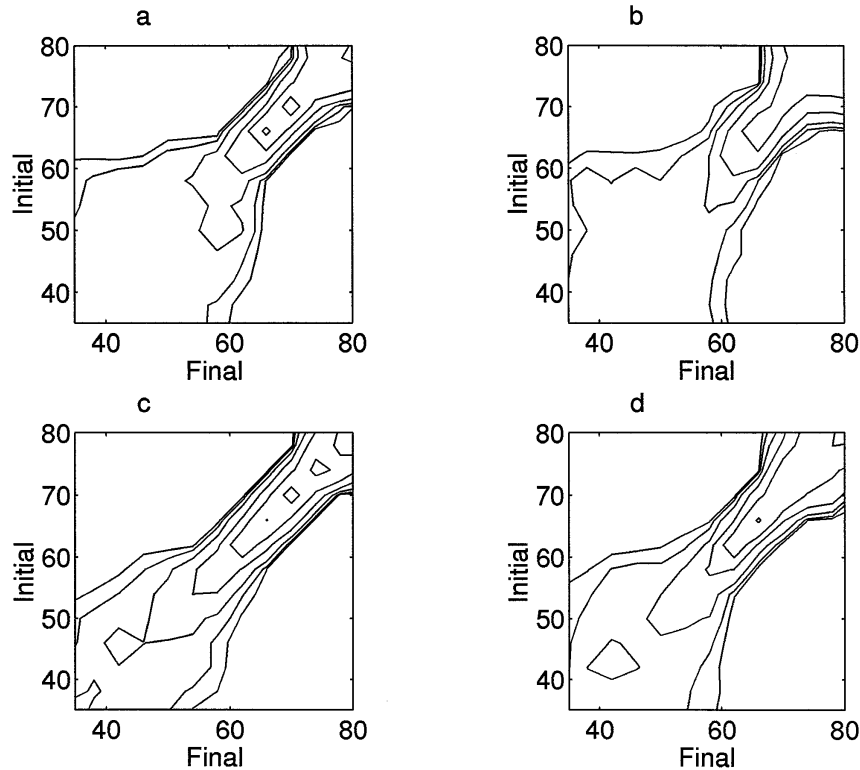


FIG. 5. Probability density fields constructed from the data in Fig. 4 (see text for details). The value of the contour farthest from the diagonal is $0.5 \times 10^{-14} \text{ m}^{-2}$, and succeeding contour values increase by consecutive factors of 2. (a) Probability density field constructed from the data in Fig. 4a. (b) Probability density field constructed from the data in Fig. 4b. (c) As in (a) except that the RDF calculation was run for only 1 day as opposed to 15 days. (d) As in (b) except that the RDF calculation was run for only 1 day as opposed to 15 days.

initial equivalent latitude versus final equivalent latitude for the RDF calculations shown in Figs. 1d and 2d, derived as described above. That is, the maps between PV and equivalent latitude on the initial and final days were obtained from Eq. (2) using the analyzed PV fields on the initial and final days, respectively. The vortex edge region is clearly defined by the narrow “neck” in the cloud of particles. Particles initially in this region, around $65^\circ\text{--}70^\circ$, do not travel far, and parcels initially on either side tend to stay on the same side. There appears to be considerable mixing within the regions to either side of the edge. This sort of plot contains much useful information, but there are so many particles in the picture that one cannot extract it all by eye.

Figures 5a and 5b show probability density fields constructed from the same data as in Figs. 4a and 4b. The probability P_{if} that a particle whose final location is equivalent latitude ϕ_f will have come from a location between equivalent latitudes ϕ_i and $\phi_i + \Delta\phi_i$ is

$$P_{if} = Q_{if} 2\pi r_e^2 \Delta\phi_i \cos\phi_i,$$

where Q_{if} is the probability density and A_e is the radius of the earth. The fields here were computed using bins 4° wide in equivalent latitude; a smaller bin size might be appropriate in principle but results in noisier plots.

This field is closely related to the “transilient matrix” of Stull (1984; see also Stull 1993 and references therein).

Note again the neck corresponding to the vortex edge and the apparent mixing on either side, especially on the equatorward side. Our results aside, a large number of prior studies lead us to expect strong mixing in this “surf zone.” However, our coordinate is also particularly ill defined there. The large-scale gradients of PV are very weak, and the PV field is dominated by noise and ill-resolved small-scale structure. So while the strong mixing in the surf zone is reflected in the calculations, it is in part because the mixing causes the analyses to be noisy and the contour crossing method interprets noise as transport, rather than because the method can, in the presence of noise, faithfully capture the transport process.

It is possible to use the information in these plots to calculate a net transport across any given contour; one simply adds up the areas of all parcels that have crossed from one side of a given contour to the other, that is, the parcels off the diagonal in Figs. 4a and 4b. This is equivalent to the method used by DB94 and M94. This can be divided by the area enclosed by the contour to get a value in percent per month. Using this method we

obtain minimum transport rates between 20% and 40% per month (i.e., across no contour is the transport less than this value) from a typical 15-day calculation, based on the calculations shown here as well as others performed for the same winter and the preceding one. This figure is not an average of some events having near zero transport and others having transport of more than 40%; virtually no calculation, no matter how quiet the period, exhibits transport of much less than 10%, which must be multiplied by 2 to give percentage per month. One may reduce the computed transport by using a threshold, that is, by requiring that the parcels in Fig. 4a or Fig. 4b be more than a certain distance away from the diagonal to be considered transported. The transport will then, in general, depend on what the threshold is. This will be discussed further below.

c. Sensitivity to duration of calculation

In this section, we examine the sensitivity of the amount of transport exhibited in our calculations to the length of time over which the RDF calculations are carried out. Figures 5c and 5d show probability density matrices for RDF calculations of duration 1 day, at 450 K, terminating on 31 January and on 16 February. The degree of mixing across the vortex edge is not much less than that in the 15-day calculations. Reducing the bin size does not change this result. A single day is clearly too short a time for substantial mixing to occur by any large-scale advective process. Simple scaling arguments, or visual inspection of a time sequence of plots from a CA run, can be used to demonstrate that the timescale for Rossby wave breaking is on the order of at least a few days. Hence we must interpret the apparent single-day mixing evinced by Figs. 5c and 5d as an artifact of the calculation. The primary cause of this artifact is that the analyses do not exactly conserve PV even on nominally resolvable spatial scales (since one day is not long enough for the advection to generate scales noticeably smaller than are initially present) and a 1-day timescale. Since a change in PV is equivalent to transport in a PV-based equivalent latitude coordinate, this erroneous nonconservation leads to spurious apparent transport. The extent of this failing appears minor to the eye when analyses and RDF calculations are compared, as in Figs. 1b and 1d and Figs. 2b and 2d. However, the errors can add up to cause a significant problem for transport calculations using contour crossing.

d. Effect of noise in analyses: Illustration

An analytical example for an idealized case can be used to demonstrate that a realistic amount of noise in the PV analyses may be expected to lead to transport errors of the magnitude shown here. Consider a perfectly circular patch of conserved tracer centered on the pole in a two-dimensional atmosphere undergoing pure zonal

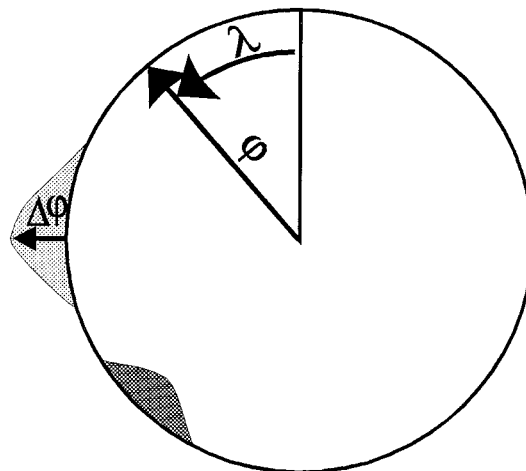


FIG. 6. Schematic illustration of setup for analytical example illustrating the effects of tracer analysis noise on contour crossing calculations (see text). Here ϕ is latitude, λ longitude, ϕ_0 is the latitude of a (zonally symmetric) tracer contour, and $\Delta\phi$ is the error in the analyzed contour placement.

flow on a sphere. A tracer isopleth at the edge of the patch will be defined by the curve $\phi = \phi_0$, where ϕ represents latitude and ϕ_0 is a constant. Clearly no transport of any kind occurs in this case, however defined, since there is no change of any quantity in time due either to motion or any other process. Imagine that the tracer field is diagnosed imperfectly from observations, so that the circle appears slightly “wiggly”; the analyzed tracer boundary is $\phi_a = \phi_0 + \Delta\phi(\lambda)$, where λ is longitude and $\Delta\phi(\lambda)$ is always small compared to $90 - \phi_0$. A polar stereographic view of this situation is given in Fig. 6.

Now consider a calculation in which one attempts to estimate the transport across a contour at the vortex edge, using the contour crossing method and using the imperfect observations of the tracer field as the analyses. Consider first the initial condition: in places where $\Delta\phi$ is negative, so that the wiggle is outward, parcels will be marked as inside the vortex that really are not. Conversely, when $\Delta\phi$ is positive, parcels will be marked as outside the vortex that really are inside.

To find the amount of spurious transport caused by the observational errors, first consider the simpler (if unlikely) case in which the tracer field is somehow known exactly correctly at the final time, even though it was not so known at the initial time, that is, $\Delta\phi = 0$ at the end of the calculation though it was nonzero at the beginning. The amount of apparent inward transport then corresponds simply to the number of parcels that were falsely marked as outside initially and the apparent outward transport to those falsely marked as inside. If we assume that these two transport rates are equal (i.e., that the zonal average of $\Delta\phi$ is zero), then it may easily be verified that they are both equal to

$$T \equiv dA/A = \frac{1}{2} \sum_i^M \left| \int_{\lambda_i}^{\lambda_i + \Delta\lambda_i} \frac{\sin(\phi_0 + \Delta\phi) - \sin(\phi_0)}{1 - \sin(\phi_0)} d\lambda \right|,$$

where the λ_i are the zeros of $\Delta\phi$ for which $\Delta\phi(\lambda_i + \epsilon)$ is either positive or negative, depending on whether the inward or outward transport is being calculated; $\Delta\lambda_i$ is the distance between λ_i and the next zero of $\Delta\phi$; there are M such λ_i for $0 \leq \lambda < 360$; and ϵ is an infinitesimally small increment of longitude. That is, the positive and negative areas under the curve $\Delta\phi$ are calculated separately, as is appropriate. For small $\Delta\phi$, we may approximate T by

$$T = \frac{\cos(\phi_0)}{2[1 - \sin(\phi_0)]} \overline{|\Delta\phi|},$$

where the overbar indicates a zonal average.

Now consider the more probable case in which, at the end of the calculation, one again has only an imperfectly analyzed tracer field to decide which parcels are in and which out, and that the error term $\Delta\phi$ has the same mean absolute value at the end as at the beginning. For convenience, consider first the special case in which the atmosphere is at rest. In principle, then, the erroneously calculated inward and outward transport could be anywhere from zero to $2T$, depending on the degree of correlation between $\Delta\phi_f$ and $\Delta\phi_i$. If $\Delta\phi_i$ and $\Delta\phi_f$ are uncorrelated, and if the function $\Delta\phi(\lambda)$ is such that its rms value $[(\Delta\phi^2)]^{1/2}$ is a good estimate of the mean of its absolute value (e.g., the rms of a sinusoid is $\sqrt{2}/2$, whereas the mean of its absolute value is $2/\pi$, about a 10% difference), then it can be easily shown that the inward and outward transport will each be equal to $\sqrt{2}T$. This result also holds for arbitrary zonal flows as long as the analyzed edge displacements at the final time overlap in an uncorrelated way with those of the contour generated by the advection of the initial analysis. The advection of the initial contour by the flow, so that it may become multiple valued in latitude, simply makes the mathematical expression of this condition more complex.

Using these results, one finds that for a vortex edge located at 65° the false inward and outward transport are each equal to 5% for $|\Delta\phi|$ equal to 1° , with that value increased to 7% for a vortex edge at 70° . It is hard to imagine that any existing observational dataset can allow one to place PV contours closer, on average, than one or two grid spacings (typically one to several hundred kilometers) of their true locations. This can be expected to cause false transports in the neighborhood of 10%–20% per calculation. If the calculations are of a fraction of a month duration, then the estimate of transport per month will be this percentage divided by that fraction. By this spurious mechanism, one quickly obtains estimates of transport rates much higher than the lowest ones available in the literature, say, for a quiet month with no major disturbances to the vortex.

Though this example is highly idealized, the result does not appear to depend very much on any of the unrealistic assumptions made in order to simplify the analysis. The extension to more complex vortex shapes and zonally asymmetric flows makes analytical progress difficult, but we see nothing that would cause the errors associated with a given amount of noise in the tracer analyses to be significantly smaller than in our idealized case. This view is supported by the magnitude of the false transport in our 1-day contour crossing calculations.

It is worth emphasizing two aspects of the false transport that can appear by this mechanism.

- It contributes equally to inward and outward transport, as long as the noise in the analyzed PV has zero mean. Thus the calculated transport will be exaggerated more in the direction in which the true transport is smaller.
- If we assume that the noise in the analyzed PV is uncorrelated from one day to the next, then the amount of spurious transport is not dependent on the duration of the calculation. Thus if one divides the spurious transport by the duration to get a flux, the magnitude of the flux will be inversely proportional to the duration.

The transport calculated by contour crossing can be reduced by choosing an appropriate threshold, that is, requiring that a parcel's initial and final tracer values change by more than some specified finite amount in order for it to be counted as having been transported in equivalent latitude space. While DB94 did not specify a particular threshold, they did provide a complete accounting of how many of the parcels transported across the vortex edge had changed their PV by every specific amount. The reader is thus free to choose his or her own threshold, as the authors suggest may be appropriate; by choosing points far enough off the peaks of the curves in their Figs. 5, 10, and 12, one can reduce their estimates greatly. One can consider applying an analogous interpretation to Fig. 7 of M94. We note that E95 did use a small threshold, but their transport rates are so small that the results are not particularly sensitive to the precise value.

If one wishes to use contour crossing with a threshold, a good choice would be the error in equivalent latitude that corresponds to the rms amplitude of the noise in PV. This corresponds to the average spatial distance by which PV contours are misplaced. Lack of knowledge of the exact PV field makes evaluating this difficult, but a reasonable, if crude, guess is something on the order of a grid spacing or two in the analyses.

6. Local gradient reversal

a. Concept

Local gradient reversal differs fundamentally from contour crossing but resembles the CAS technique in

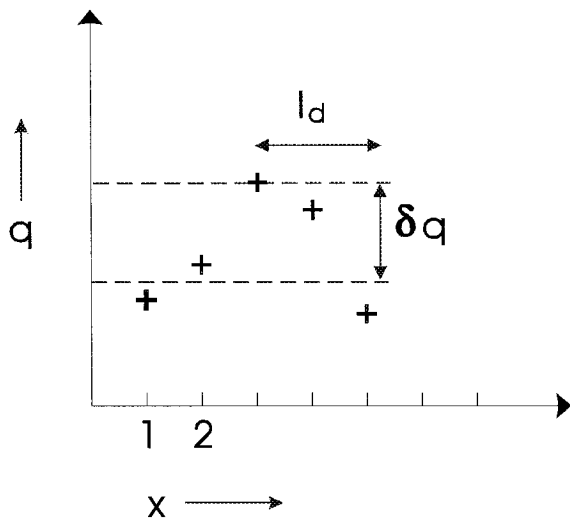


FIG. 7. Schematic illustration of the inequalities used in the local gradient reversal algorithm. The figure depicts a one-dimensional example, showing several points in a grid representation of a tracer field, q , as a function of spatial coordinate x . Each point has an index i , indicated by the numbers under the x axis, such that its location is $i\Delta x$ with Δx the grid spacing. With the threshold parameter δq set as shown, the point with the largest value for q satisfies the criterion for outward transport (assuming that q increases towards the pole) if the value $n = 2$ is used in the inequalities, so that the scale $l_d \sim 2\Delta x$.

that it makes no use of the tracer analysis on the final day of the calculation. Since the analyzed PV is noisier (due to the extra spatial derivative) than the analyzed winds, from a practical perspective it is desirable that the PV enter the calculation as infrequently as possible.

The approach of local gradient reversal is to perform an advection calculation using a tracer analysis as an initial condition, and then, as in the CAS technique, to identify small-scale features (e.g., filaments) in the result. The material contained in these features is then considered irreversibly transported. Local gradient reversal is essentially a simple pattern recognition algorithm for identifying small-scale features. The algorithm defines these as regions where the tracer gradient is both large and changes sign. This is why we call it the “local gradient reversal” method. A similar idea has been used previously by Baldwin and Holton (1988) to define regions of Rossby wave breaking in analyzed PV fields.

Following the ideas in section 2, and the fundamental assumption on which the technique rests is that there is some scale, l_d , below which structures are dynamically passive. Thus any structure smaller in at least one dimension than l_d is destined to be rapidly stretched to still smaller scales by the large-scale flow. At some small scale, we assume that the structure will be dissipated, but we do not need to know the details of this as long as we know no other outcome is possible, that is, that the feature cannot undergo an inverse cascade back upscale. This assumption must be checked in each case, by examining the advection results to ensure that filament roll-up followed by vortex remerger has not

occurred. The assumption is also made that the tracer analyses to be used as initial conditions are smooth, containing no structures of scale less than some scale that is larger, possibly much larger than l_d . In practice we always start with low-resolution tracer data, so the scale separation between the smallest scale in the analyses and l_d is enforced by the choice of l_d (if high-resolution, accurate tracer data were available, one might imagine intentionally smoothing them at the outset, but this is not a relevant problem here). The rate of irreversible transport occurring during the course of the advection calculation is then estimated by adding up the areas of features with scales less than l_d at the calculation’s end.

We might expect that our choice of l_d will affect the estimate of transport. Although we do not resolve the issue completely in this study, results shown below suggest that for currently plausible resolution in the tracer analyses, and a reasonable choice of duration time for the advection (presumably this should be of the same order as the dynamical timescale for Rossby wave breaking), there is a “good” range of l_d that both can be justified on physical grounds and within which the transport is fairly insensitive to the precise l_d . Also, a threshold parameter δq , which prescribes the minimum distance in equivalent latitude that a parcel must travel in order for the algorithm to count it as having been transported, plays an important role. Sensitivity to these parameters is discussed below.

b. Numerical algorithm

Here, the numerical algorithm used to implement the local gradient reversal method is introduced. First, perform an RDF calculation (Sutton et al. 1994; Schoeberl and Newman 1995) on an equal area grid. This provides a high-resolution tracer field, gridded in such a way that each grid point represents an equivalent amount of mass if we neglect the variation of isentropic density along the isentrope. This latter approximation is shown by Nakamura (1995) to be a good one in the circumstances of interest here. For simplicity, we first consider a Cartesian domain with coordinates x , y , and grid points evenly spaced so that the grid spacings along both dimensions are equal, $\Delta x = \Delta y$. Then the tracer field $q(x, y)$ is represented discretely as $q(i, j)$ where i and j are the indices of the two coordinates, $x = i\Delta x$, $y = j\Delta y$. Then, count the points that satisfy either

$$q(i, j) > q(i - n, j) + \delta q$$

and

$$q(i, j) > q(i + n, j) + \delta q$$

or

$$q(i, j) > q(i, j - n) + \delta q$$

and

$$q(i, j) > q(i, j + n) + \delta q,$$

where δq is a threshold that must be chosen and n is an integer. For small n , these points are thus small-scale local maxima in at least one spatial direction and so are considered as being part of small-scale features that have been transported outward across tracer contours whose values are between $q(i, j)$ and $q(i, j) - \delta q$, for a tracer that increases towards the pole (as PV does in the Northern Hemisphere). To measure inward transport, count the points satisfying either

$$q(i, j) < q(i - n, j) - \delta q$$

and

$$q(i, j) < q(i + n, j) - \delta q$$

or

$$q(i, j) < q(i, j - n) - \delta q$$

and

$$q(i, j) < q(i, j + n) - \delta q.$$

This gives the inward transport across contours with values between $q(i, j)$ and $q(i, j) + \delta q$. Figure 7 depicts the meaning of the inequalities through a schematic one-dimensional example.

We neglect the pathological case of a saddle point, at which $q(i, j)$ is a local maximum in one direction and a local minimum in the perpendicular direction, as this is a very rare occurrence in realistic tracer fields. The criteria for inward and outward transport should obviously be reversed for a tracer that decreases towards the pole.

We divide the entire range of q present in the domain into bins, which then represent regions of space in an equivalent latitude coordinate ϕ_e , where the transformation is given by Eq. (2). Then the fractional outward transport from a given equivalent latitude bin, T_o , is defined as the number of points in that bin satisfying the outward breaking criteria divided by the total number of points in the bin. In other words, if a certain fraction of the material initially lying in a given range of equivalent latitude is drawn out into small-scale features according to the definitions given in the inequalities above, then we say that that fraction has been transported to a different equivalent latitude. The distance that the material has been transported, which we may call $\delta\phi_e$, is that which corresponds to δq via Eq. (2). In words, the distance traveled by the material in the small-scale feature is directly related to the difference between its tracer value and that of its environment, with which it is assumed to be destined to mix. The fractional inward transport T_i is defined analogously but using the inward-breaking criteria.

c. Choice of l_d

The integer n is the concrete representation of the scale l_d in the numerical algorithm. In the results to follow, only $n = 1$ and $n = 2$ have been used. Recall

that a grid spacing in our RDF calculations corresponds to 1° of latitude, or about 110 km. This means that we have chosen to call the formation of features roughly between 100 and 300 km wide irreversible. This choice for l_d is not provably correct, but it is at least reasonable. Anything much larger begins to approach, in a scaling sense, the Rossby radius of deformation. It seems untenable to assume that the velocity field will be insensitive to PV features at this scale or larger. At the smaller-scale end, our choice is consistent with the recent results of Bacmeister et al. (1996). They analyzed aircraft data from the stratosphere over a wide geographical and temporal range and showed that below the 100-km scale (the largest in their analysis) any possible contribution to the kinetic energy spectrum from a 2D enstrophy cascade was invisible, due to the dominating influence of gravity waves. The unbalanced nature of the flow below the 100-km scale is further evidence of the dynamical irrelevance of PV features at this scale.

It can easily be seen that if saddles are excluded, performing the tests along only two perpendicular axes leads to an uncertainty in the effective l_d of, at most, a factor $\sqrt{2}$. In the results shown below, an equal area grid has been used, in which the nodes are arranged in zonal circles but the spacing in longitude is inversely proportional to the cosine of latitude. Hence the nodes do not lie neatly along the meridians. This is dealt with by performing three sets of tests rather than two, since each node has three pairs of adjacent nodes with which it is approximately collinear.

Given a choice for l_d of some multiple of the grid scale, the threshold δq measures the distance, in equivalent latitude, that a feature has been transported, through the direct relationship between q and ϕ_e . Hence, setting δq higher than the maximum variance of q present within the surf zone allows us to look at entrainment of vortex air into the surf zone without the picture's being complicated by the in situ stretching of blobs of slightly high PV that originate in the surf zone. Both will produce filaments with higher PV than their environments, but only at the vortex edge are the net PV drops across finite regions large enough to produce PV drops of, say, 10 PVU at 450 K across a single grid spacing at the end of a 15-day run. Figure 8 illustrates this point schematically.

The same choice for l_d should be applicable for different tracers, as long as δq is adjusted accordingly and as long as the initial tracer fields have the same resolution. Since different long-lived tracers tend to be highly correlated with one another, the δq values for two different tracers should simply be adjusted so that they correspond to the same difference in equivalent latitude, and then the results should be independent of which tracer is used to the extent that the correlation holds.

d. Local gradient reversal results

Figures 9a and 9b show results of a local gradient reversal analysis carried out for the same RDF calcu-

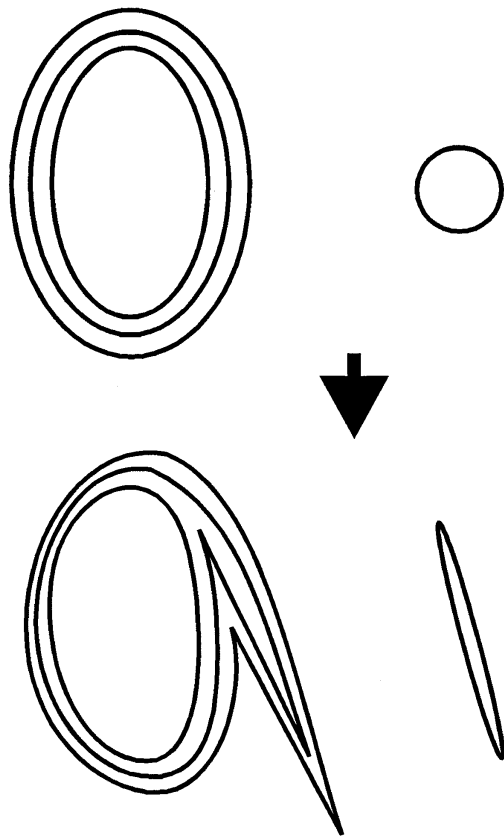


FIG. 8. Schematic illustration of the role of the threshold parameter δq in distinguishing transport out of the vortex from effects of surf zone "noise." The top portion of the figure represents a contour map of a PV field to be used as an initial condition for an advection calculation. The field contains a main vortex and a small secondary vortex. After the passage of some time (denoted by the arrow) the field is distorted by the flow into the configuration shown in the bottom portion of the figure, in which a filament has been partly peeled off the main vortex and the secondary vortex has been sheared out into another filament. If δq is chosen to correspond to twice the contour interval, no parcels within the filament formed from the secondary vortex can ever satisfy the local gradient reversal criterion, while some of those within the filament peeled off the main vortex can do so.

lations used to generate Figs. 1c, 2c, 4a–b, and 5a–b. The quantities plotted are T_o and T_i as defined above. The PV bin spacing is chosen to be equal to 2° in equivalent latitude, where the map between equivalent latitude and PV is constructed directly from the RDF output. The threshold used is $\delta q = 10$ PVU. In the vortex edge region 10 PVU corresponds to roughly 7° in equivalent latitude, so that the quantity plotted is the fraction of the air initially at the equivalent latitude given on the x axis (within 2°) that has moved, over the period of the calculation, at least 7° from that location. Alternatively, one can note that during this period, the PV drop from the innermost vortex to the surf zone is around 20–25 PVU (see Figs. 1a, 1b, 2a, and 2b). This means that the quantity plotted is the fraction of the air

initially at each ϕ_o , which crosses roughly half of the vortex edge region during the calculation.

In both cases, we see more transport for $n = 2$ than for $n = 1$. This is required since, by definition, if the local gradient reverses over one grid spacing, it will do so as well at two grid spacings (except in the unrealistic case that there are alternating filaments of high and low PV of a single grid spacing in width), while the converse is not true. However, the difference is modest, indicating that within this range, the results are not particularly sensitive to the choice of this scale. The outward curves are small everywhere, but in particular are indistinguishable from zero in the region around 65° – 70° , so that this region forms a transport barrier. The peaks just poleward of 60° are either outside the vortex completely or barely in the outer vortex edge, depending on the definition. Thus these peaks correspond to stripping of the outer edge, as opposed to transport of true vortex air to the surf zone. The spike at twice grid scale right near the pole corresponds to stirring in the deep interior of the vortex; the single bin at which T_o is nonzero corresponds in any case to a very small area. The inward transport is zero everywhere except for a tiny spike poleward of 70° corresponding to slight inward entrainment of inner edge air into the innermost vortex.

In the later period, we see somewhat more vigorous mixing than in the earlier period, as we expect. In particular, a small but not insignificant amount of material is shown to be detrained from the surf zone into the vortex when $n = 2$; examination of the CA and RDF results (Figs. 1c, 1d, 2c, and 2d) clearly reveals an intrusion occurring during this period. Outward transport is increased as well, though it should be recognized that the inward and outward transport are not completely independent. If a filament of vortex air extruded into the surf zone then wraps tightly around the vortex edge and encloses some surf zone air, then those points that are enclosed will be counted as having been transported inward, as the PV on either side of them will be higher than their own. This mechanism works the same way for intruding filaments enclosing vortex air. In the lower stratosphere, it appears common that filaments wrap tightly around the vortex edge, forming a laminar structure (W94; P94). It is a reasonable description of this situation to say that inward and outward transport are occurring in the same event.

e. Sensitivity to threshold

Figures 10a and 10b show the results of calculations identical to those in Figs. 9a and 9b with the exception that the threshold has been set to 5 PVU, or 3.5° of equivalent latitude in the vortex edge region. Unsurprisingly, the transport increases *almost* everywhere. However, especially in the first period it is important to note that there are still regions of zero or nearly zero (i.e., less than 3%) transport, both inward and outward, in the vortex edge region. These regions are more than

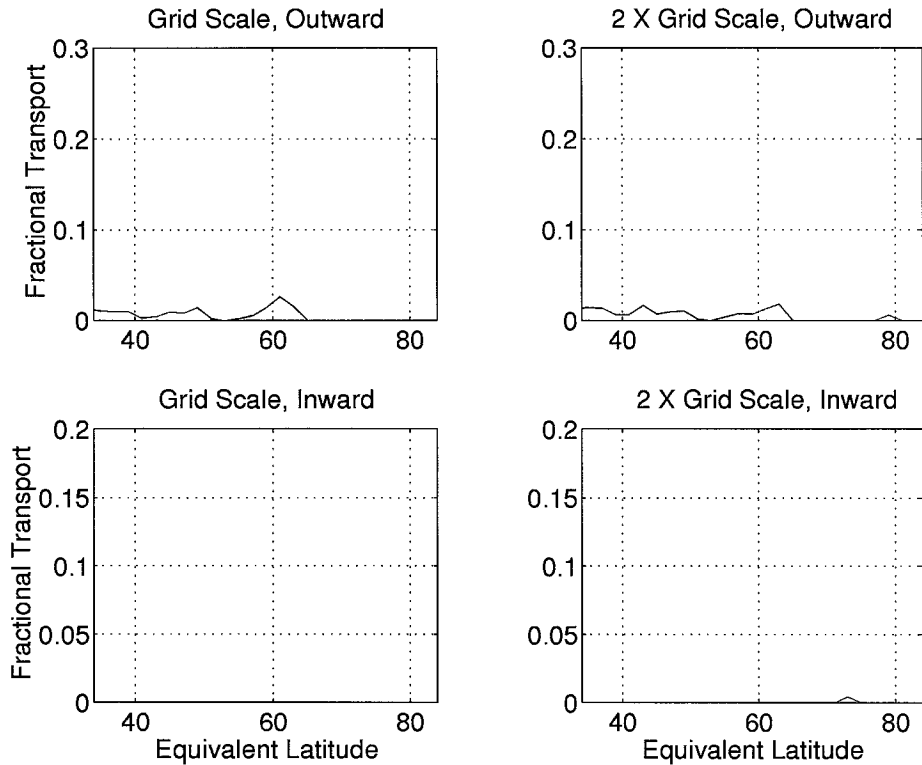


FIG. 9a. Results of local gradient reversal diagnostic (see text) applied to RDF calculations shown in Fig. 1d. The threshold parameter has been set to 10 PVU, corresponding to approximately 7° of equivalent latitude in the vortex edge region. Hence the quantity plotted in the “outward” (“inward”) case is the fraction of the parcels in each 2° equivalent latitude bin that has been transported 7° or more equatorward (poleward) in equivalent latitude. “Grid Scale” or “Twice Grid Scale” indicates $n = 1$ and $n = 2$, respectively, where n is the parameter used in the local gradient reversal algorithm (see text and Fig. 7).

3.5° wide, so it can still be sensibly said that virtually no material is going across them. Thus with a lower threshold we pick up more short-range stirring at the edges of the transport barrier, but the barrier itself is still clearly present as such. Close visual examination of the CA and RDF results in Figs. 1c and 1d suggests that zero, or very close to it, is the appropriate result for the contours in the sharpest part of the vortex edge.

An equivalent threshold in conjunction with the contour crossing calculation for this period gives a significantly nonzero result everywhere, both inward and outward, since the neck in Fig. 4a is around 6° wide (10 PVU) at its narrowest point and roughly symmetric about the diagonal. Noise-induced “lumps” in the PV contours can cause spurious minor filamentation that local gradient reversal will detect, and this phenomenon is probably contributing to the results in Figs. 10a and 10b. Our test here shows, though, that local gradient reversal is more robust to this noise than contour crossing, since a smaller threshold is needed for the former technique to show zero transport when one’s eye suggests that is the correct answer, as in our quiet first period (for contours well within the vortex edge that are not deformed into filaments). For the purpose of quantifying

transport where it is clearly nonzero, the issue of definition remains a tricky one. However, the question raised by the threshold is really just whether one wishes to look at long-range or short-range transport, as opposed to which threshold gives one the “true” transport.

If the threshold is increased to 15 PVU (results not shown), the transport becomes zero everywhere in the first period, in both the inward and outward directions, for both $n = 1$ and $n = 2$. In the second period, the inward transport becomes zero everywhere, and the outward transport becomes zero everywhere poleward of 70° and is nowhere more than 3%. Thus it is justified to say that no air crosses the entire vortex edge region, or even significantly more than half of it, in even this fairly disturbed 15-day period.

f. Sensitivity to analysis noise

In the special case of the idealized circular vortex in an atmosphere at rest given in section 5c, the local gradient reversal method clearly gives zero transport even for a very noisy analysis ($\Delta\phi$ with large rms), as long as the noise is at scales larger than l_d . In the case of the 1-day RDF calculations for which probability density

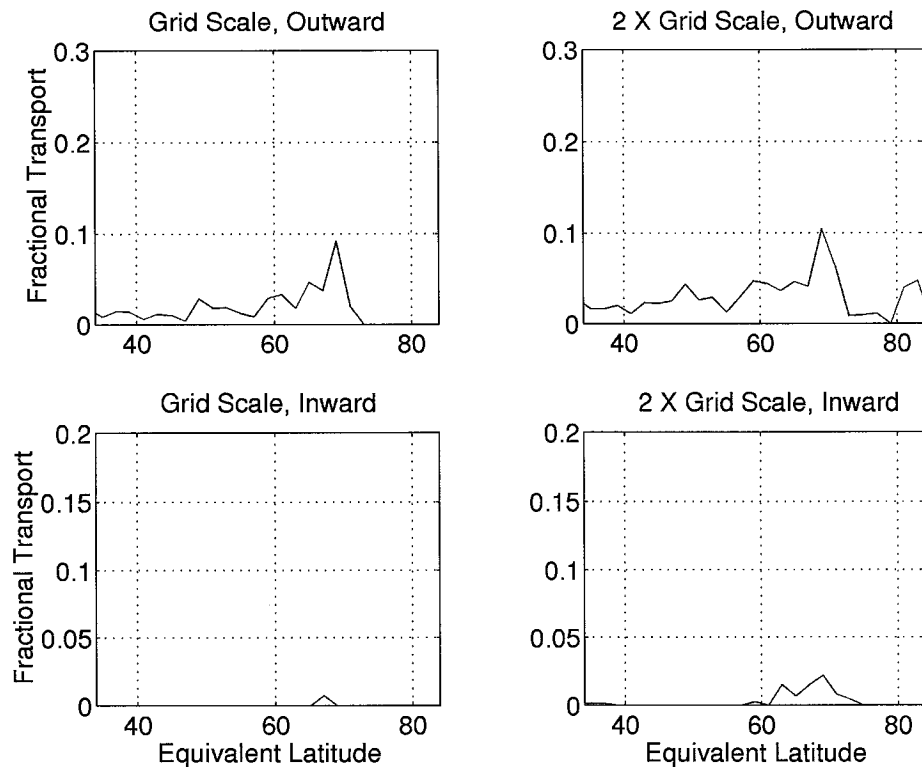


FIG. 9b. As in Fig. 9a but for RDF run shown in Fig. 2d.

matrices were shown in Fig. 5, the method with parameter settings as in Fig. 9 or 10 gives zero transport in both directions at both once and twice the grid scale. Zero is clearly the correct result in both of these cases. In situations where the correct result is not so obvious, a test of sensitivity to noise is more difficult. This is true for contour crossing as well. However, the zero-transport case is a useful one to understand if one is interested particularly in transport barriers.

g. Sensitivity to initial conditions

Figures 11a and 11b show the results of calculations identical to those in Fig. 9a and 9b, except that the wrong initial conditions were used. The parcels are advected exactly as before but marked at the initial time by PV fields from 1 day earlier than is correct. In the 15-day calculation terminating on 31 January, the parcels were marked on 16 January with PV from 15 January, and in the latter calculation the parcels were marked on 1 February with PV from 31 January. The PV fields used are similar to the correct ones but are significantly different in detail. It is important to realize that the changes in the initial conditions are not small-scale perturbations about the large-scale structure but small wholesale translations, rotations, and/or deformations of the large-scale structure itself. This is why sensitivity to initial conditions has been considered separately from sensitivity to analysis noise: the latter is

assumed to be negligible at the largest scales present in the analyzed tracer fields.

The values are somewhat larger in Figs. 11a and 11b than the corresponding values in Figs. 9a and 9b but not very much so. This reflects the fact that local gradient reversal is not particularly sensitive to small variations in initial conditions. If the initial conditions are varied to a greater extent, the diagnosed transport will increase correspondingly. This is not surprising; no method can work well if the tracer data are completely wrong.

7. Conclusions and discussion

We have shown that a realistic amount of noise in the tracer fields input to a contour crossing calculation can cause the transport to be significantly overestimated. The local gradient reversal method is less sensitive to this noise because it does not rely as heavily on the tracer fields. It must be emphasized that we have not shown that, in general, local gradient reversal gives "correct" results and contour crossing "incorrect" results; we have simply shown that the former is in one particular respect more robust than the latter. Other major questions remain unresolved, and for that reason we do not claim great quantitative accuracy for our transport results. The exception to this is that when our calculations show transport of a few percent per month or less across some region and that result is fairly insen-

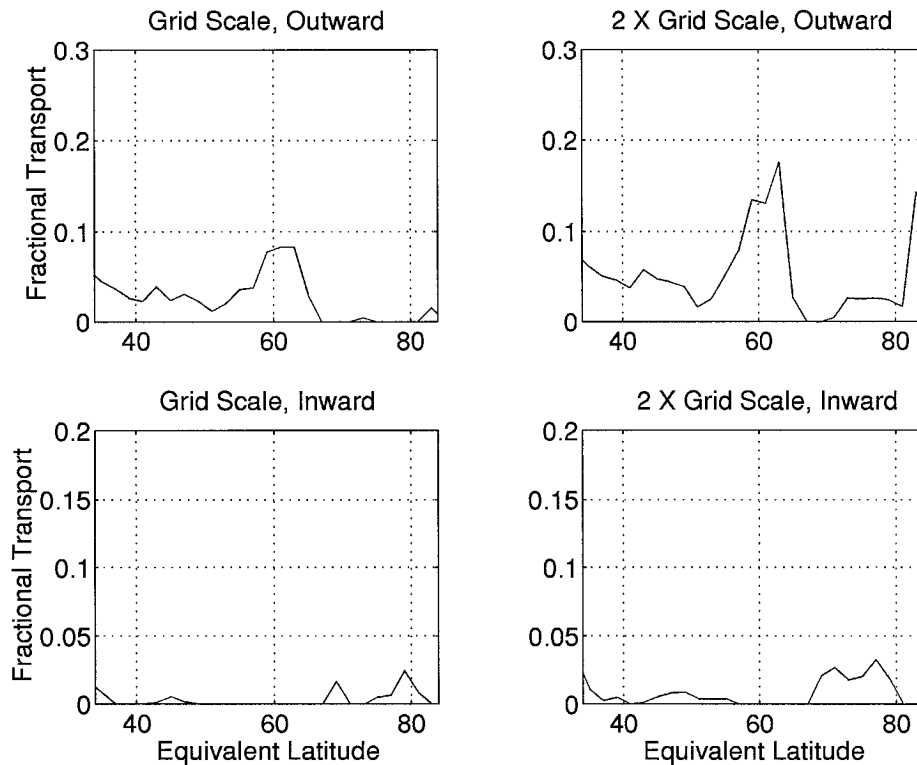


FIG. 10a. As in Fig. 9a but with the threshold parameter set to 5 PVU.

sitive to parameter variations, we feel confident in saying that the correct result is not many times larger than that, that is, not tens of percent per month.

Because of the unresolved issues, this study should be considered an introductory exploration. For example: from observational data we have no direct way of knowing what the “true” transport is, and this limits the extent to which we can validate any technique using such data. The algorithm as implemented here is fairly crude, and one can imagine refinements. The relationships between different parameters, in particular analysis resolution, l_d , and the duration of the calculation, have not been fully explored. One imagines that, for example, for lower-resolution analyses, a longer calculation will be needed to develop features of a given l_d . Another study, in which the local gradient reversal technique is subjected to further development and testing using numerical model output rather than observational analyses as input to the advection calculations, is currently under way and will shed further light on these issues. The use of numerical model output allows more accurate validation and allows the roles of the parameters to be examined somewhat more carefully, since the complicating effect of observational noise is removed.

One conclusion that is certain is that contour crossing calculations using observed tracer data should be viewed with care. Say one wished to compute the inward transport across the vortex edge by doing isentropic con-

tour crossing calculations of 10 days’ duration; this is an entirely reasonable choice (Schoeberl and Newman 1995). Using a longer duration would reduce the spurious transport on a per-month basis, but increasing it too much would increase the error due to the isentropic approximation; this error is not just a result of PV non-conservation but also of the three-dimensionality of the true motion (E95). We have shown that, for analyzed tracer contours with errors in their placement no worse than can be reasonably expected from any observational dataset, transports of as much as 10%–20% per calculation may be expected, even if the true transport is zero, if the contour crossing method with no threshold is used. This corresponds to 30%–60% per month. In the context of arguments about “containment vessels” versus “flowing processors,” it is important to be able to distinguish zero from 30%–60% per month. DB94 discussed, in a qualitative way, some of the sources of uncertainty inherent to contour crossing (though they did not call the method by that name) and stated that their estimates should be taken as upper bounds. Our results lead us to agree with this interpretation.

Though we have not shown the results here, we have performed contour crossing calculations using wind and PV data output from the high-resolution shallow water model of Polvani et al. (1995). These model fields are obviously free from observational noise, and in this case contour crossing gives reasonable results. A 1-day calculation yields almost zero transport everywhere, as it

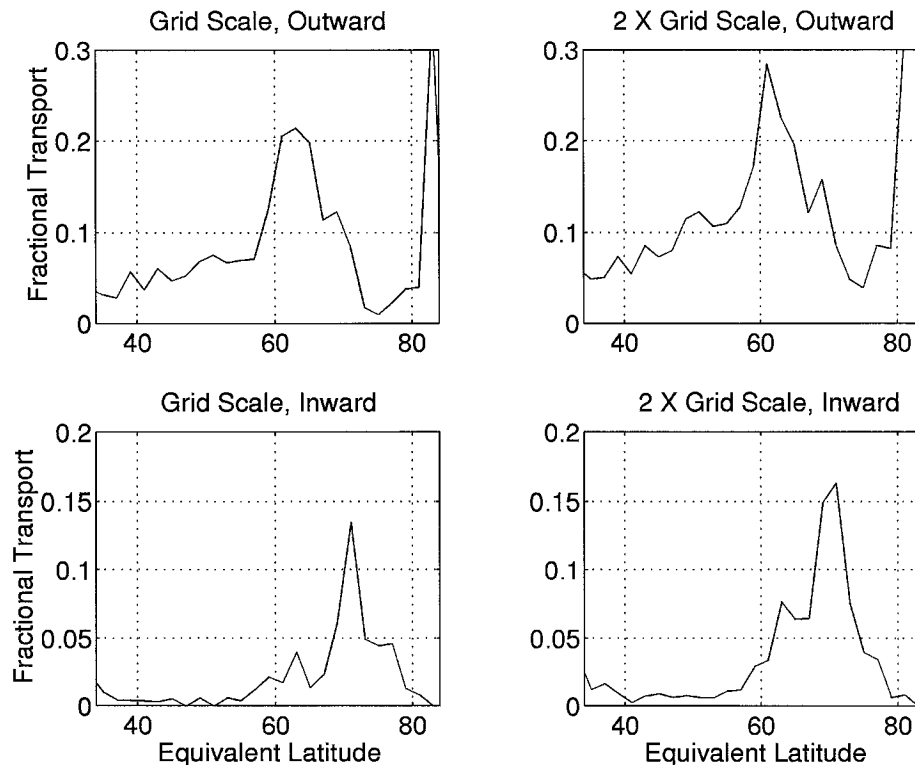


FIG. 10b. As in Fig. 9b but with the threshold parameter set to 5 PVU.

should. As the calculation's duration is increased, it yields increasingly more transport at the outermost vortex edge and within the surf zone, while the transport across PV contours well within the edge region remains indistinguishable from zero. This suggests that E95 may have obtained smaller vortex edge transports than DB94 or M94 simply because the former authors used winds and tracer fields from a model (in which the tracer is advected by the same wind fields E95 used) rather than from observational analyses, guaranteeing greater consistency between the winds and the tracer, hence less error to pollute the contour crossing calculations. Of course, some degree of genuine difference between the transport across the model vortex edge and that across the real one cannot be entirely ruled out.

Though this has been a paper about methods, in the process of demonstrating the local gradient reversal technique we have obtained estimates of transport in the vortex edge region for two 15-day periods of one arctic winter at 450 K. In order to tally the transport across some boundary in percent per month, one has to pick a threshold, decide whether once or twice the grid scale is a more appropriate choice for l_d , choose the equivalent latitude of the boundary, and then perform a weighted integral of appropriate parts of the appropriate curves in Fig. 9 or Fig. 10.

If we choose a boundary between 65° and 70° (since we are looking for the location of minimum transport, this is reasonable based on all the results in this paper,

by both methods) and take $\delta q = 10$ PVU, we find that both outward and inward transport proceed at a rate corresponding to no more than 1% or 2% per month for the period 16–31 January 1993. We get somewhat larger values for the more disturbed period 1–16 February 1993, perhaps up to 5%–10% per month outward and 2%–5% per month inward. Since our two periods are half a month long, the net figures for inward and outward transport for the entire month should be taken as averages of the two periods, giving 1%–4% per month inward and 3%–5% per month outward. However, it should be kept in mind that this transport is really inward and outward transport of material originating within the edge region.

Direct exchange of air in either direction between the deep vortex interior, or even the inner edge region, and the surf zone is not distinguishable from zero in our calculations. Visual investigation of the RDF results indicates that this is a robust result and not dependent (within reason) on our choice of irreversibility cutoff scale. The amount of material diagnosed as being irreversibly stripped from the edge and transported shorter distances is more sensitive to this choice, particularly at the very innermost and outermost edges. However, the meaning of the transport in these regions is questionable since our ability to diagnose the precise width of the vortex edge, for example, using the criterion of Nash et al. (1996), is limited by the resolution of the analyses. Some of the stripping of outer edge material

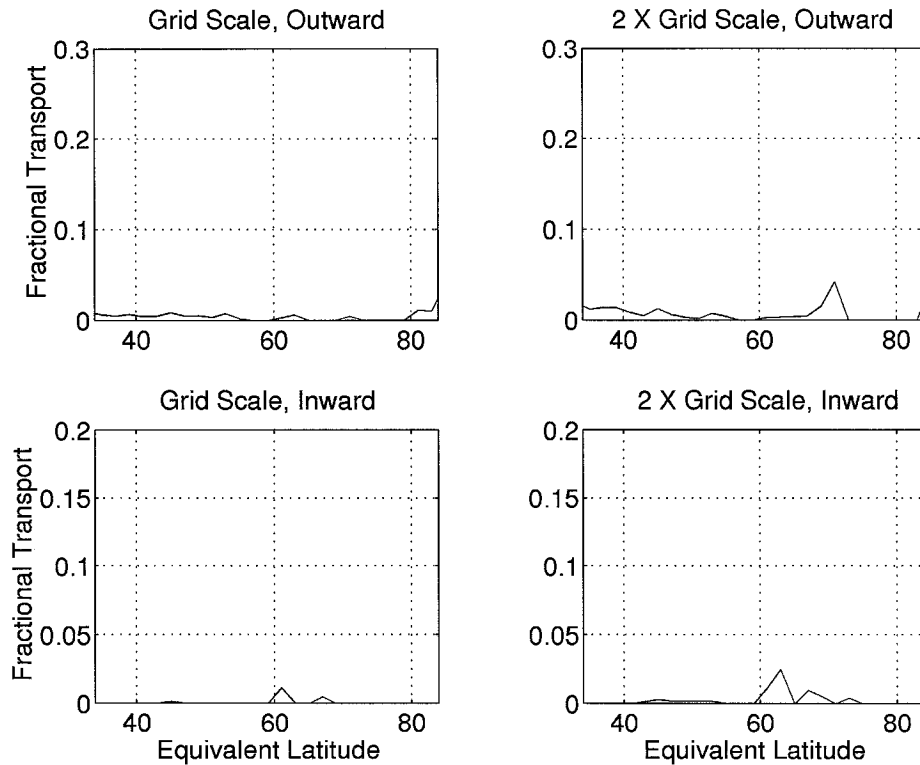


FIG. 11a. As in Fig. 9a but with incorrect initial conditions used for the RDF runs (see text).

into the surf zone may simply be due to initial misdiagnosis of surf zone material as outer edge material if, as is far from impossible, the true PV gradients at the vortex edge are sharper than the analyzed ones. The same process may affect the diagnosed entrainment of inner edge material into the inner vortex.

We have also performed calculations from the winter of 1991/92 for the purpose of comparing the results to those of W94, who used the CAS technique. For identical periods, and the same choice of PV contour for the vortex edge (really an outermost edge contour, about 62° equivalent latitude), and $\delta q = 10$ PVU as above, our transport values are generally smaller than theirs by 20%–50%. We think that this is well within the present uncertainties due to the choice of l_d in our case and the maximum surgery scale in theirs, parameters with analogous roles in their respective techniques. Because of these uncertainties, we have focused here on only two cases, rather than performing a large series of calculations over many years to define a climatology of transport. Such an exercise would provide some information on the relative interannual variability, but perhaps not much more than one could obtain by counting the number of major disturbances to the vortex in each of some set of winters using subjective or crude objective criteria.

We have not performed a full comparison here between local gradient reversal and the CAS technique. The two are similar in spirit, and part of our motivation

here has been a desire to extend ideas originally developed in the context of the contour-based methods to gridded fields, for reasons given in section 3. Based on these admittedly loose notions, and the comparison for the specific cases mentioned in the preceding paragraph, we imagine that in general results from the two methods should not differ too dramatically. In particular, for noisy tracer data, and excepting major vortex fragmentation events, we expect that the two will usually differ by less than either one will differ from the contour crossing method. At this point, though, we cannot justify any statement stronger than this.

A truly complete uncertainty analysis would, of course, have to take into account uncertainties in the wind analyses, an issue we have avoided entirely. Our justification for this is partly that, however uncertain the winds, the PV will always be worse, so it makes sense to examine the effects of errors in PV first. Also, however, the nature of the advection problem is such that accounting fully for the effects of wind errors is sure to be a daunting task.

The work presented here has been motivated by our interest in the polar vortex, and this has dictated our choice of examples. However, we reiterate that transport across both the subtropical edge of the surf zone and the midlatitude tropopause are similar in a number of respects to transport across the polar vortex edge. We imagine that some of the ideas and techniques discussed

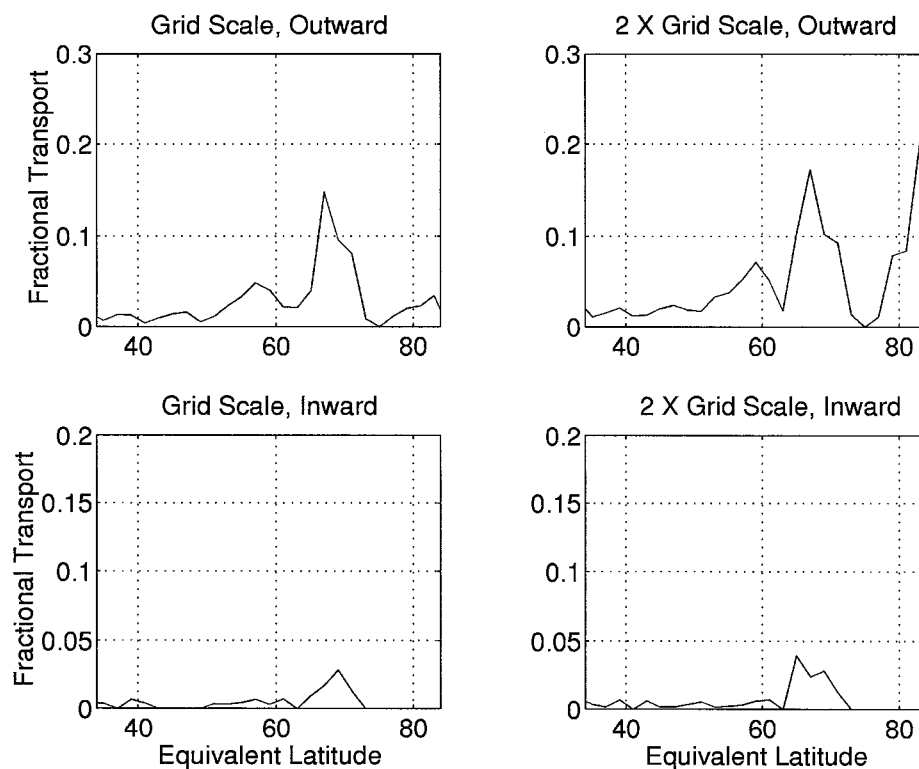


FIG. 11b. As in Fig. 9b but with incorrect initial conditions used for the RDF runs (see text).

here may be, perhaps with some modifications, relevant to the former two problems as well.

Acknowledgments. We thank Paul Newman for suggesting the use of scatterplots as in Fig. 4; Lorenzo Polvani for providing the shallow water model data; Bjorn Stevens for insightful comments on early versions of the manuscript; and Myles Allen, Janusz Eluszkiewicz, Paul Kushner, and Lorenzo Polvani for helpful discussions. Comments by two anonymous reviewers led to a number of improvements. This work has been partially supported by NASA Grant NAGW-1727 and by a NASA graduate student Global Change Research Fellowship.

REFERENCES

- Bacmeister, J. T., S. D. Eckermann, P. A. Newman, L. Lait, K. R. Chan, M. Loewenstein, M. H. Proffitt, and B. L. Gary, 1996: Stratospheric horizontal wavenumber spectra of winds, potential temperature, and atmospheric tracers observed by high-altitude aircraft. *J. Geophys. Res.*, **101**, 9441–9470.
- Baldwin, M. P., and J. R. Holton, 1988: Climatology of the stratospheric polar vortex and planetary wave breaking. *J. Atmos. Sci.*, **45**, 1123–1142.
- Butchart, N., and 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.
- Carnevale, G. F., 1982: Statistical features of the evolution of two-dimensional turbulence. *J. Fluid Mech.*, **122**, 143–153.
- Dahlberg, S. P., and K. P. Bowman, 1994: Climatology of large-scale isentropic mixing in the arctic winter stratosphere from analyzed winds. *J. Geophys. Res.*, **99**, 20 585–20 599.
- Danielsen, E. F., 1968: Stratospheric-tropospheric exchange based on radioactivity, ozone and potential vorticity. *J. Atmos. Sci.*, **25**, 502–518.
- Dritschel, D. G., 1989a: Contour dynamics and contour surgery: Numerical algorithms for extended, high-resolution modeling of vortex dynamics in two-dimensional, inviscid, incompressible flows. *Comp. Phys. Rep.*, **10**, 77–146.
- , 1989b: On the stabilization of a two-dimensional vortex strip by adverse shear. *J. Fluid Mech.*, **206**, 193–221.
- , and L. M. Polvani, 1992: The roll-up of vorticity strips on the surface of a sphere. *J. Fluid Mech.*, **234**, 47–69.
- , and D. W. Waugh, 1992: Quantification of inelastic interactions of vortices in two-dimensional vortex dynamics. *Phys. Fluids A*, **4**, 1737–1744.
- Eluszkiewicz, J., R. A. Plumb, and N. Nakamura, 1995: Dynamics of wintertime stratospheric transport in the Geophysical Fluid Dynamics Laboratory SKYHI general circulation model. *J. Geophys. Res.*, **100**, 20 883–20 900.
- , and Coauthors, 1996: Residual circulation in the stratosphere and lower mesosphere as diagnosed from microwave limb sounder data. *J. Atmos. Sci.*, **53**, 217–240.
- Fisher, M., A. O'Neill, and R. Sutton, 1993: Rapid descent of mesospheric air into the stratospheric polar vortex. *Geophys. Res. Lett.*, **20**, 1267–1270.
- Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.
- Lamarque, J.-F., and P. G. Hess, 1994: Cross-tropopause mass exchange and potential vorticity budget in a simulated tropopause folding. *J. Atmos. Sci.*, **51**, 2246–2269.
- Manney, G. L., R. W. Zurek, A. O'Neill, and R. Swinbank, 1994: On the motion of air through the stratospheric polar vortex. *J. Atmos. Sci.*, **51**, 2973–2994.

- McIntyre, M. E., and T. N. Palmer, 1985: A note on the general concept of wave breaking for Rossby and gravity waves. *Pure Appl. Phys.*, **123**, 964–975.
- Nagatani, R. M., A. J. Miller, M. E. Gelman, and P. A. Newman, 1990: A comparison of the Arctic lower stratospheric winter temperatures for 1988–89 with temperatures since 1964. *Geophys. Res. Lett.*, **17**, 333–336.
- Nakamura, N., 1995: Modified Lagrangian-mean diagnostics of the stratospheric polar vortices. Part I: Formulation and analysis of GFDL SKYHI GCM. *J. Atmos. Sci.*, **52**, 2096–2108.
- , 1996: Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate. *J. Atmos. Sci.*, **53**, 1524–1537.
- 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.
- Newman, P. A., and M. R. Schoeberl, 1995: A reinterpretation of the data from the NASA stratosphere-troposphere exchange project. *Geophys. Res. Lett.*, **22**, 2501–2504.
- , L. R. Lait, M. R. Schoeberl, and R. M. Nagatani, 1990: Stratospheric temperatures during the 88–89 Northern Hemisphere winter. *Geophys. Res. Lett.*, **17**, 329–332.
- , and Coauthors, 1996: Measurements of polar vortex air in midlatitudes. *J. Geophys. Res.*, **101**, 12 879–12 892.
- Norton, W. A., 1994: Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. *J. Atmos. Sci.*, **51**, 654–673.
- , and M. P. Chipperfield, 1995: Quantification of the transport of chemically activated air from the Northern Hemisphere polar vortex. *J. Geophys. Res.*, **100**, 25 817–25 850.
- O'Neill, A., W. L. Grose, V. D. Pope, H. McLean, and R. Swinbank, 1994: Evolution of the stratosphere during northern winter 1991–1992 as diagnosed from the UKMO analyses. *J. Atmos. Sci.*, **51**, 2800–2817.
- Pierrehumbert, R. T., 1991: Chaotic mixing of tracer and vorticity by modulated travelling Rossby waves. *Geophys. Astrophys. Fluid Dyn.*, **58**, 285–319.
- Plumb, R. A., D. W. Waugh, R. J. Atkinson, P. A. Newman, L. R. Lait, M. R. Schoeberl, E. V. Browell, A. J. Simmons, and M. Loewenstein, 1994: Intrusions into the lower stratospheric Arctic vortex during the winter of 1991–1992. *J. Geophys. Res.*, **99**, 1089–1105.
- Polvani, L. M., D. W. Waugh, and R. A. Plumb, 1995: On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.*, **52**, 1288–1309.
- Randel, W. J., 1987: The evaluation of winds from geopotential height data in the stratosphere. *J. Atmos. Sci.*, **44**, 3097–3120.
- Rosenfield, J. E., P. A. Newman, and M. R. Schoeberl, 1994: Computations of diabatic descent in the stratospheric polar vortex. *J. Geophys. Res.*, **99**, 16 677–16 689.
- Schoeberl, M. R., and P. A. Newman, 1995: A multiple-level trajectory analysis of vortex filaments. *J. Geophys. Res.*, **100**, 25 801–25 816.
- Strahan, S. E., J. E. Rosenfield, M. Loewenstein, J. R. Podolske, and A. Weaver, 1994: Evolution of the 1991–1992 Arctic vortex and comparison with the Geophysical Fluid Dynamics Laboratory SKYHI general circulation model. *J. Geophys. Res.*, **99**, 20 713–20 723.
- Stull, R. B., 1984: Transient turbulence theory. Part I: The concept of eddy-mixing across finite distances. *J. Atmos. Sci.*, **41**, 3351–3367.
- , 1994: Review of non-local mixing in turbulent atmospheres: Transient turbulence theory. *Bound.-Layer Meteor.*, **62**, 21–96.
- Sutton, R. T., H. MacLean, R. Swinbank, A. O'Neill, and F. W. Taylor, 1994: High-resolution stratospheric tracer fields estimated from satellite observations using Lagrangian trajectory calculations. *J. Atmos. Sci.*, **51**, 2995–3005.
- Waugh, D. W., 1996: Seasonal variation of isentropic transport out of the tropical stratosphere. *J. Geophys. Res.*, **101**, 4007–4023.
- , and D. G. Dritschel, 1991: The stability of filamentary vorticity in two-dimensional geophysical vortex-dynamics models. *J. Fluid Mech.*, **231**, 575–598.
- , and R. A. Plumb, 1994: Contour advection with surgery: A technique for investigating finescale structure in tracer transport. *J. Atmos. Sci.*, **51**, 530–540.
- , and Coauthors, 1994: Transport out of the lower stratospheric arctic vortex by Rossby wave breaking. *J. Geophys. Res.*, **99**, 1071–1088.
- Wei, M.-Y., 1987: A new formulation of the exchange of mass and trace constituents between the stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 3079–3086.
- Yang, H., 1995: Three-dimensional transport of the Ertel potential vorticity and N₂O in the GFDL SKYHI model. *J. Atmos. Sci.*, **52**, 1513–1528.