Quantitative Diagnostics of Mixing in a Shallow Water Model of the Stratosphere

ADAM H. SOBEL

University of Washington, Seattle, Washington

R. ALAN PLUMB

Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 15 January 1998, in final form 2 November 1998)

ABSTRACT

Two different approaches are applied to quantify mixing in a shallow water model of the stratosphere. These are modified Lagrangian mean (MLM) theory and a technique referred to as ''reverse domain filling with local gradient reversal'' (RDF-LGR). The latter is similar to a previously existing technique using contour advection and contour surgery.

It is first proved that in an inviscid shallow water atmosphere subject to mass sources and sinks, if the mass enclosed by a potential vorticity (PV) contour is steady in time, then the integral of the mass source over the area enclosed by the contour must be zero. Next, the MLM and RDF-LGR approaches are used to diagnose the time-averaged transport across PV contours in the model simulations.

The model includes a sixth-order hyperdiffusion on the vorticity field. Except in a thin outer ''entrainment zone,'' the hyperdiffusion term has only a very weak effect on the MLM mass budget of the polar vortex. In the entrainment zone, the hyperdiffusion term has a significant effect. The RDF-LGR results capture this behavior, providing good quantitative estimates of the hyperdiffusion term, which is equivalent to the degree of radiative disequilibrium at a PV contour. This agreement shows that the main role of the hyperdiffusion is to ''mop up'' the filaments that are produced by the essentially inviscid large-scale dynamics. All calculations are repeated for two values of the hyperdiffusion coefficient that differ by a factor of 50, with little difference in the results. This suggests that the amount of material entrained from the vortex edge into the surf zone does not depend on the details of the small-scale dissipation, as long as it is sufficiently weak and has some degree of scale selectivity.

1. Introduction

While much has been learned about stratospheric transport in the past decade, a fully quantitative understanding is still lacking. One difficulty is the quantitative description and diagnosis of the relatively rapid isentropic mixing that coexists with the slower meridional overturning (e.g., Andrews et al. 1987; Holton et al. 1995). Several approaches to quantifying this mixing using global-scale data, whether observational or model generated, have been developed, but each has its flaws. This study applies two apparently different approaches to output from a shallow water model of the stratosphere, namely, the model of Polvani et al. (1995, hereafter PWP). Both approaches use contours of quasi-conserved tracers as the relevant boundaries across which to compute transport. It is important that the tracer chosen for the coordinate not be absolutely conserved, since in this case the transport across the contours is trivially zero at all times.

The first approach uses modified Lagrangian mean (MLM) theory, originally suggested by McIntyre (1980) and recently developed by Nakamura (1995, 1996). In the shallow water system, MLM theory involves using isopleths of a tracer (two tracers, if the theory is applied in three dimensions) to define a partially Lagrangian coordinate. The primary quantity is the mass enclosed by a given contour. All ''transport'' in this coordinate is due to nonconservative processes acting on the tracer. This approach is theoretically rigorous, but limited in its applicability to observations because it requires arbitrarily high-resolution data and accurate knowledge of all significant nonconservative processes. When applied to numerical model output, however, the MLM approach is quite viable. Its application here enables some basic dynamical properties of the forced-dissipative, shallow water polar vortex to be elucidated. A new, exact analytical result concerning the nature of an inviscid steady state in a shallow water atmosphere subject to

Corresponding author address: Dr. Adam H. Sobel, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640.
E-mail: sobel@atmos.washington.edu

thermal forcing (a mass source) aids in the interpretation of the numerical results.

The second approach was originally developed in studies by Waugh et al. (1994) and Plumb et al. (1994). Those authors use the contour advection (CA) technique (Waugh and Plumb 1994; Norton 1994) to generate fine-scale vortex ''filaments'' from contour representations of initially smooth tracer fields. Contour surgery (Dritschel 1989a) is then applied repeatedly to remove the filaments, producing ''coarse-grained'' contours. The material contained in the filaments is then considered transported across the contour. We will call this technique ''contour advection with coarse graining'' (CA-CG). Here, we use a different technique, though one embodying the same approach. Reverse domain filling (RDF) (Sutton et al. 1994; Schoeberl and Newman 1995) replaces CA, and a refined version of the local gradient reversal (LGR) algorithm (Sobel et al. 1997, hereafter SPW) replaces coarse graining. We will call this ''reverse domain filling with local gradient reversal'' (RDF-LGR). These techniques do not require very high-resolution data nor any explicit knowledge of nonconservative processes, but they are at bottom ad hoc, lacking rigorous theoretical justification. They also involve one or more free parameters.

This second approach may, however, be viewed as consistent with MLM theory if the assumption is made that the filaments, once formed, can have no other fate than dissipation by some small-scale process, as suggested by SPW. If the additional assumption is made that the removal of the filaments is the primary effect of that small-scale process on the MLM mass budget, then the transport estimated by this approach may be compared to the net transport associated with the small-scale dissipation in the MLM mass budget. This comparison is performed here using the shallow water model output, and the two quantities are shown to be quite similar. This agreement is robust to modest variations of the free parameters in the RDF-LGR technique. That technique and CA-CG are also compared and shown to give very similar results. Taken together, these results indicate that the ad hoc approaches provide estimates of transport across tracer contours that are, in fact, robust and physically meaningful.

This conclusion holds because of the validity of the assumptions concerning the fate of filaments and the role of small-scale dissipation. That these assumptions hold implies also that *the details of the small-scale dissipation are irrelevant* to the net rate of entrainment of vortex material into the surf zone, as long as the dissipation is sufficiently weak and has some scale selectivity. The amount of material shed in the form of filaments is determined by large-scale dynamical processes that do not directly feel the dissipation. This is reminiscent of the famous result, originally assumed by Kolmogorov and later verified experimentally [for a good review and references to the original literature, see Frisch (1995)], that in homogeneous isotropic three-dimensional turbulence, the rate of energy dissipation is independent of the viscosity as the limit of infinite Reynolds number is approached.

These results, namely, the insensitivity of the net transport to the details of the dissipation and the ability of contour surgery-type methods to reproduce MLM results, may not necessarily generalize to dynamical situations very different from the one considered here. Here, a strong, coherent vortex dominates the flow at all times, providing strong shear and strain that suppress the instabilities that might otherwise cause filaments to roll up into secondary vortices (Dritschel 1989b; Waugh and Dritschel 1991; Dritschel and Polvani 1992), complicating the dynamics. Even within the winter stratospheric context, major disturbances, in which large, dynamically active vortex fragments are formed, may result in the violation of these assumptions. Other possible limitations on the generality of the results stem from the inherent limitations of single-layer models.

Nonetheless, our results suggest that it may be possible to converge toward a robust quantitative understanding of the transport that results from Rossby wave breaking at the polar vortex edge.

2. Outline

A brief outline of the study is as follows. Section 3 presents the model equations and details of the particular simulations to be diagnosed. Section 4 presents the MLM diagnostics as applied to the shallow water system, using potential vorticity (PV) to define the coordinate. We consider an inviscid steady state as defined by the MLM coordinate. This does not imply steady flow, but only that the amount of mass enclosed by a PV contour does not change with time. In such a steady state, with arbitrary mass source (the shallow water analog of diabatic heating) the integral of the mass source over the area enclosed by a PV contour must be zero. This may be viewed as the single-layer analog of radiative equilibrium. It means also that the net mass flux across the PV contour must also be zero. This result generalizes earlier results by Schneider (1987), Juckes (1987), and McIntyre and Norton (1990).

In section 5, we compute the contributions of different nonconservative terms to the model's MLM mass budget. We compute this budget for a range of PV contours in the model's extratropics. The model contains only two processes that cause PV to be nonconserved. These are the mass source, which has the form of a Newtonian relaxation on the layer thickness; and the ∇^6 hyperdiffusion, which acts on the vorticity field in the model.

This analysis highlights some interesting features of the equilibrated vortex. In the surf zone, the hyperdiffusion balances a substantial degree of radiative disequilibrium. In the bulk of the vortex edge region, the PV contours remain undular and unfilamented, and the hyperdiffusion has almost no effect. This region is thus close to being (as is then required) in radiative equilib-

rium. Between the surf zone and the main vortex edge is a narrow ''entrainment zone'' where the PV contours still tightly enclose the vortex but are frequently deformed into filaments that are drawn out to lower latitudes.

In section 6, the RDF-LGR technique is used to estimate the rate at which small-scale features in the PV field are produced. This involves first smoothing the model PV and wind fields and then using these as input to the RDF particle advection technique to regenerate high-resolution PV fields. We then use an improved version of the local gradient reversal algorithm (SPW) to quantify the rate of production of small-scale features from the high-resolution fields.

As mentioned in the introduction, we assume that these small-scale features are destined to be dissipated quickly due to the scale selectivity of the model dissipation (hyperdiffusion), and that the effect of that dissipation on unfilamented regions is relatively negligible, so that the removal of filaments is the primary role of the dissipation. With these assumptions, the RDF-LGR calculation can be interpreted as an estimate of the contribution of the dissipation to the MLM mass budget that does not incorporate any explicit knowledge of the form of the dissipation, other than that it is weak and acts predominantly at small scales. This estimate is compared to the actual contribution of the hyperdiffusion to the MLM mass budget. The agreement is good in a number of respects, and in those respects the results are only weakly sensitive to variations in the free parameters in the local gradient reversal technique. The RDF-LGR results are also compared to the results obtained from the same model data by PWP using the CA-CG technique, and the two techniques are shown to agree quite closely. Section 7 discusses the results of the preceding sections and draws conclusions.

3. The model

a. Formulation

The shallow water model is the one used by PWP. Similar models have also been used in studies by Juckes (1989) and Salby et al. (1990).[1] The model dynamical equations are

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\mathbf{u}\zeta_a) + G$$

$$\frac{\partial \delta}{\partial t} = -\frac{1}{2}\nabla^2(\mathbf{u}\cdot\mathbf{u}) + \mathbf{k}\cdot\nabla \times (\mathbf{u}\zeta_a) - g\nabla^2(h + h_b)$$

$$\frac{\partial h}{\partial t} = -\nabla \cdot (h\mathbf{u}) + Q,$$

where ζ is the relative and ζ_a the absolute vorticity, δ the divergence, h the layer thickness, $\mathbf{u}$ the fluid velocity, h_b the fixed bottom topography, Q the mass source, and G represents any friction or other vorticity sources or sinks. As in PWP, Q takes the form of a relaxation back to a prescribed, zonally symmetric ''radiative equilibrium'' state, h_e,

$$Q = \tau_e^{-1}[h_e(y) - h],$$

and a sixth-order hyperdiffusion is applied to the vorticity field

$$G = \nu\nabla^6\zeta.$$

The model equations are solved on the sphere by a spectral method. The code (with the exception of the thermal relaxation and topography) is described by Browning et al. (1989).

b. Simulation particulars

Two model runs have been diagnosed. The first run is identical to that shown in Fig. 8 of PWP, except that it was run for 150 instead of 100 days. The mountain height is set to 1500 m, radiative relaxation time τ_e = 10 days, and radiative equilibrium profile $h_e(y)$ as given in PWP (constant from the South Pole to about 20°N, then decreasing sharply to the North Pole). The hyperdiffusivity was set to 5×10^{26} m^6 s^{-1}, which gives a 5-h damping time for total wavenumber 42, so that although the model was run at T85, the effective resolution is really approximately T42.[2] This run will be referred to as the ''T42'' or ''low-resolution'' run. The second run was identical except for having a hyperdiffusion coefficient of 1×10^{25} m^6 s^{-1}, a factor of 50 smaller than in the first experiment. This yields a true T85 resolution, with a damping time of about 5 h on total wavenumber 85. This one will be referred to as the ''T85'' or ''high-resolution'' run. As noted by PWP, changing the hyperdiffusion coefficient in this range makes no qualitative difference at large scales in the simulations.

All the diagnostics have been performed for a 50-day period, beginning on day 80 of the model runs, by which time the model is in a nearly statistically steady state. We chose a 50-day period because model wave-breaking events tend to occur somewhat irregularly, with quiet and disturbed periods of roughly 10–20 days duration. A period of 50 days was judged long enough to allow us to interpret averages over this period as reasonably representative of the model's climatology for this particular set of forcings.

The flow fields have been archived at time intervals of 1 day and at spatial resolution of 1° × 1°. Since T85 spectral resolution corresponds roughly to 1.4° gridpoint

[1] See also Juckes (1996) and Polvani et al. (1996).

[2] In PWP, this coefficient was incorrectly listed as 1×10^{26} m^6 s^{-1}. The value used here, which is five times greater, was also used in PWP for their runs at T85 resolution.

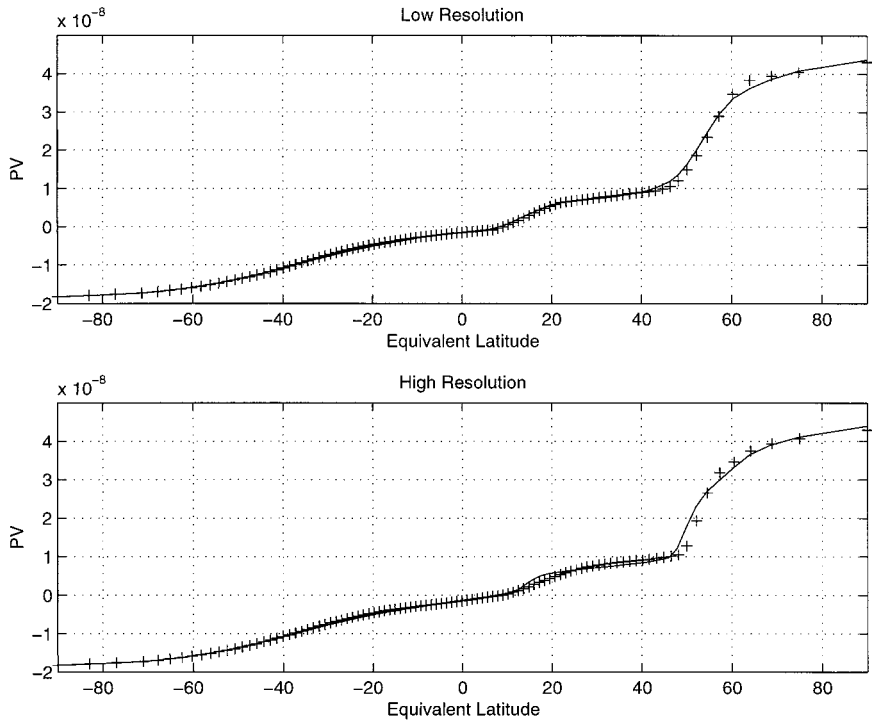

FIG. 1. The PV (m^{-1} s^{-1}) vs equivalent latitude, for both the low-resolution and high-resolution runs. Solid lines represent day 80, pluses day 130.

resolution, this means that we are slightly oversampling the fields, to guarantee that there will be no aliasing.

Figure 1 shows the profiles of PV versus equivalent latitude on the first and last days of our 50-day period, for both runs. Figure 2 shows the mass inside each contour as a function of equivalent latitude on these two days. As usual, equivalent latitude is defined as the latitude of the zonal circle that encloses the same area as the PV contour in question.

In both runs the model is clearly near a steady state, as defined by these diagnostics, during this period. Also, the structure is very nearly the same between the two runs, supporting the notion that the dynamics are at most weakly influenced by the choice of hyperdiffusion coefficient.

4. Modified Lagrangian mean diagnostics in the shallow water system

a. Basic MLM equation

The MLM framework is presented in three dimensions by Nakamura (1995). Here we present the shallow water version of the MLM equation. The necessary modifications are straightforward.

In the shallow water system, the mass (strictly, volume) enclosed by a tracer contour is

$$m(q, t) = \iint_{A(q,t)} h \, dA,$$

where $A(q, t)$ is the area enclosed at time t by the tracer contour whose value is q. We will use the Northern Hemisphere PV convention for the tracer, that is, assume that the area enclosed on the northward side by a tracer contour decreases as the contour's tracer value increases. The shallow water MLM mass continuity equation is then

$$\frac{\partial}{\partial t} m(q, t) = -\frac{\partial}{\partial q} \iint_{A(q,t)} h\dot{q} \, dA + \iint_{A(q,t)} Q \, dA, \quad (1)$$

where $\dot{q} = dq/dt$, the derivative following the flow.

The second term on the rhs above is simply the area integral of the mass source. The first term, which is identical in form to the analogous term in the 3D MLM equation, is a boundary term due to the motion of the contour relative to the fluid, which occurs due to any nonconservative processes acting on the tracer field on (as opposed to inside) the contour. By a well-known identity [e.g., see Eq. (2.1) of Nakamura (1995)] this term may be rewritten in terms of a contour integral around the boundary of A:

$$\frac{\partial}{\partial t} m(q, t) = \oint_{C(q,t)} \frac{h\dot{q}}{|\nabla q|} \, dl + \iint_{A(q,t)} Q \, dA, \quad (2)$$

where C is the tracer contour bounding the region A, and dl is an increment of scalar distance around C. Note that for the above to remain finite we must assume that $|\nabla q|$ does not vanish anywhere, except at isolated maxima and minima. In general, it is reasonable to assume

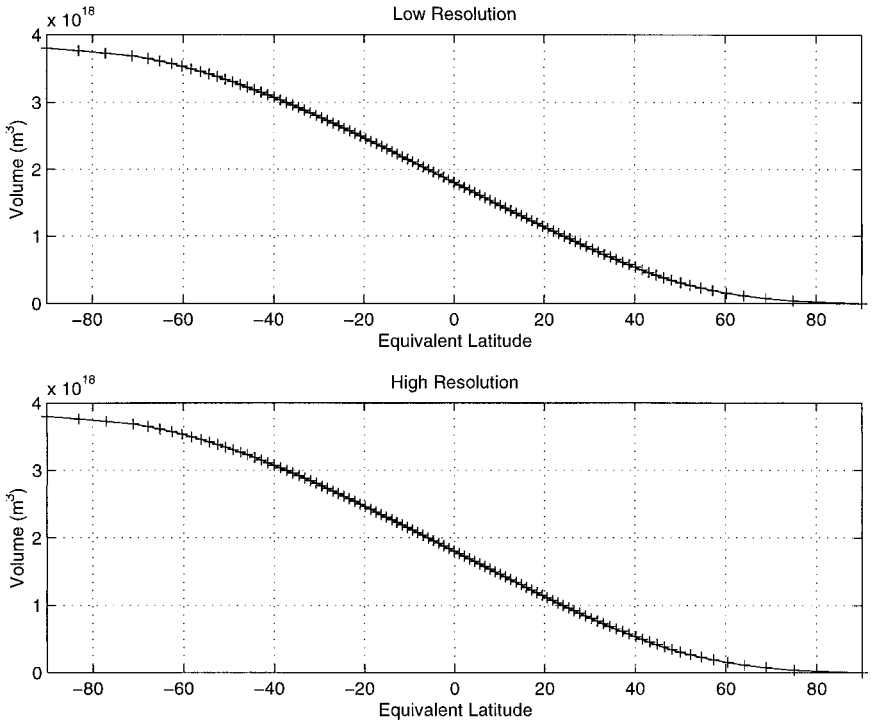

FIG. 2. Mass enclosed by PV contours (on the northern side) vs equivalent latitude, for both the low-resolution (T42) and high-resolution (T85) runs. Solid lines represent day 80, pluses day 130.

that the maxima and minima are quadratic. This means that in a Taylor expansion about the location of an extremum in q,

$$q(x_0 + \Delta x, y_0 + \Delta y) = q(x_0, y_0) + \left(\Delta x \left.\frac{\partial q}{\partial x}\right|_{x_0,y_0} + \Delta y \left.\frac{\partial q}{\partial y}\right|_{x_0,y_0}\right) + \frac{1}{2}\left[(\Delta x)^2 \left.\frac{\partial^2 q}{\partial x^2}\right|_{x_0,y_0} + (\Delta y)^2 \left.\frac{\partial^2 q}{\partial y^2}\right|_{x_0,y_0}\right] + \cdots, \quad (3)$$

where (x_0, y_0) is the location of the extremum and Δx and Δy are small, the quadratic term (which is the lowest-order term whose gradient is not automatically zero at an extremum) does not vanish. In this case, $|\nabla q|$ will go to zero linearly as the extremum is approached. The contour length will shrink to zero linearly as well, so that the contour integral in (2) will remain finite.

b. The inviscid steady state

Consider the MLM mass budget of a shallow water atmosphere that is inviscid ($G = 0$ everywhere) but subject to thermal forcing ($Q \neq 0$). Let the tracer that defines the MLM coordinate be the shallow water potential vorticity P:

$$P = \zeta_a/h.$$

Then since $G = 0$, the material derivative of P is (following the notation introduced in section 3a)

$$\frac{dP}{dt} = -\frac{PQ}{h}. \quad (4)$$

No restriction has been placed on the form of Q. Substituting (4) into (2), the result is

$$\left.\frac{\partial m}{\partial t}\right|_{P,t} = -P \oint_{C(P,t)} \frac{Q}{|\nabla P|}\, dl + \int_P^{P_{\max}} dP' \oint_{C(P',t)} \frac{Q}{|\nabla P'|}\, dl, \quad (5)$$

where the area integral in the second term has been rewritten in terms of contour integrals, and $P_{\max}$ is the the global maximum value of P. Now defining

$$F(P) = \oint_{C(P,t)} \frac{Q}{|\nabla P|}\, dl,$$

(5) becomes

$$\left.\frac{\partial m}{\partial t}\right|_{P,t} = -PF(P) + \int_P^{P_{\max}} F(P')\, dP'. \quad (6)$$

Now consider a steady state, so that the lhs is zero. Note that this does not imply an assumption of steady flow. Taking the derivative with respect to P, and noting that

for positive argument the integral on the rhs increases as the lower bound decreases, we have

$$2F(P) + \frac{dF(P)}{dP} = 0. \tag{7}$$

We assume that the steady state holds over the region between P and $P_{\max}$.[3] By the argument in the preceding section concerning local maxima and minima, $F(P)$ remains finite at $P_{\max}$. This, and (6) with the LHS set to zero, implies that

$$P_{\max}F(P_{\max}) = 0,$$

which implies that $F(P_{\max})$ vanishes, since by assumption $P_{\max}$ does not. This provides the necessary boundary condition on (7), whose general solution is

$$F(P) = CP^{-2},$$

with C a constant of integration. The boundary condition requires $C = 0$, so that

$$F(P) = 0$$

is the solution throughout the region. Integrating with respect to q implies that

$$\iint_A Q\, dA = 0$$

also. In other words, there can be no net radiative disequilibrium in an MLM steady state, nor any net mass flux into or out of the PV contours, unless $G \neq 0$. This result is a generalization of those of Schneider (1987) and McIntyre and Norton (1990), who reached similar conclusions under assumptions of steady flow. We emphasize that the present result only assumes that $m(P, t)$ is constant in time, so that the contours may wobble and distort while containing the same mass; this is a much weaker assumption. The relationships between these various results are explained in more detail in appendix B. Also, we note that the above result holds trivially for an exactly conserved tracer, since then the first term on the rhs of (1) or (2) vanishes. The present result is not trivial due to the nonconservation of P as per (4).

Note that at each contour, the definition of $F(P)$ means that the integral around the contour of $Q/|\nabla P|$, rather than of Q itself, must vanish. If the steady state is statistical rather than exact, so that $\partial[m(q, t)]/\partial T = 0$ where T is some slow timescale defined by a linear time-averaging operation, the same conclusion applies. That is, the area integral of Q must vanish in the same time average.

For a general vorticity forcing G, we have

$$\frac{dP}{dt} = \frac{G - Qq}{h}.$$

Then, in an MLM steady state, we have

$$-\oint_{C(P,t)} \frac{G}{|\nabla P|}\, dl = -P \oint_{C(P,t)} \frac{Q}{|\nabla P|}\, dl + \iint_{A(P,t)} Q\, dA. \tag{8}$$

The quantity on the left-hand side—explicitly representing the vorticity forcing (in our case, the hyperdiffusion)—may in fact be viewed as a measure of the *degree of radiative disequilibrium*, defined at a particular PV contour, in the shallow water system. We will use this term in what follows.

Strictly, this quantity cannot be considered to be a purely local property at a particular contour, since the right-hand side of (8) contains an area integral. However, the area integral term will turn out to be small relative to the other terms in the polar vortex region of the shallow water model, for reasons to be made clear below. It is presently unclear to what extent, and under what circumstances, this situation may generalize to the three-dimensional case. Because of this, the definition of radiative disequilibrium used here should presently be viewed as strictly meaningful only within the context of the shallow water system. This is related to the fact that Q is not strictly analogous to a heating rate in three dimensions, but only to the part of the heating rate that changes the amount of mass between two isentropic surfaces.[4] Nonetheless there are clear dynamical similarities between the shallow water and three-dimensional cases; the precise consequences of the differences between them is a subject of current investigation.

In sum, we see that in the MLM framework as applied to the shallow water system, the degree of radiative disequilibrium in steady state is directly and solely related to the degree of vorticity forcing, since the former must be zero if the latter is. Knowledge of this fact aids our interpretation of the numerical results below.

5. Numerical results: Balances maintaining quasi equilibrium

From the shallow water model output, we have computed the terms on the right-hand side of the MLM continuity equation [(1)] for every day of the 50-day period, once per day. It may be straightforwardly shown from the model equations that, in the numerical model, the shallow water PV obeys the equation

[3] It may be added that, strictly, the results to follow require that $G = 0$ between P and $P_{\max}$, but not outside this region.

[4] It may be shown that Q is analogous to $\partial(\sigma\dot{\theta})/\partial\theta$, with θ the potential temperature, $\dot{\theta}$ the heating rate, and $\sigma = -g\partial p/\partial\theta$ the isentropic density. Thus the heating rate may be nonzero while the quantity analogous to Q remains zero.

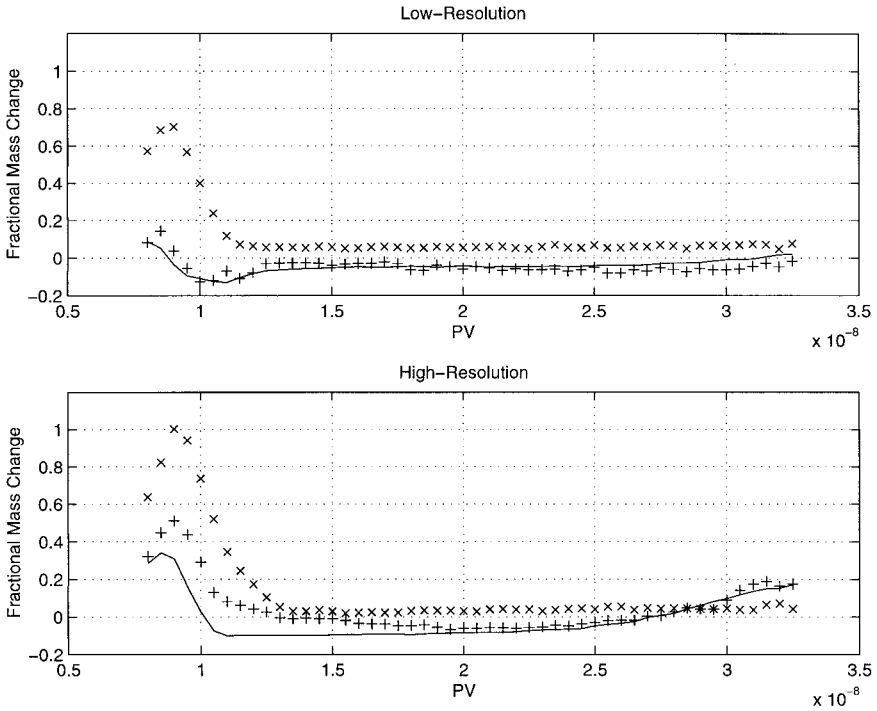

FIG. 3. Comparison of true change in mass between days 80 and 130, and that computed by MLM diagnostics, for both the low-resolution and high-resolution runs. Solid curve is the actual change, pluses the MLM result, and ×'s the MLM result without including the hyperdiffusion term. The mass change is shown as a fraction of each contour's average mass during the period.

$$\frac{dP}{dt} = P(\tau_e h)^{-1}(h - h_e) + \nu h^{-1}\nabla^6\zeta. \tag{9}$$

So, for a range of PV values spanning the model Northern Hemisphere, each term on the rhs of

$$\left.\frac{\partial m}{\partial t}\right|_{P,t} = -\frac{\partial}{\partial P}\iint_{A(P,t)} [\tau_e^{-1}P(h - h_e) + \nu\nabla^6\zeta]\, dA + \tau_e^{-1}\iint_{A(P,t)} (h - h_e)\, dA \tag{10}$$

has been computed. The hyperdiffusion term was dumped directly from the model's time-stepping routine, rather than being diagnosed from the gridded vorticity field. We computed the area integrals using simple "box counting," that is, counting an entire grid box as being enclosed inside a PV contour if the PV value at the center of the grid box was greater than the contour's value. The computations were done using a grid size of $1° \times 1°$. Sensitivity tests were performed by interpolating onto a grid with double the resolution and repeating the area integrations. These showed that the larger grid size was adequate for all terms but the hyperdiffusion term in the high-resolution run. The hyperdiffusion term in this case varied at the smallest spatial scale of both the model and the grid on which the data were output, and thus it proved difficult to perform very accurate area integrals of this term.

Having computed the daily values for the three contributions to the mass tendency, we computed a "predicted" change to the total mass inside each contour over the entire period by integrating in time, that is, multiplying each daily tendency by one day and summing them. This was also done without including the hyperdiffusion term, in order to allow this term to be estimated as a residual as an alternative to its direct computation. The predicted mass change can be compared to the actual change in mass over the period, the mass at the end minus that at the beginning, computed directly using only the fields of height and PV on the initial and final days. The result of this comparison is shown for both runs in Fig. 3. The *x* axis is PV, rather than equivalent latitude. This means that relative to physical space, the coordinate is stretched for PV values greater than about 1×10^{-8} m^{-1} s^{-1} and contracted for lower values. This is advantageous since we are particularly interested in the region at and just above 1×10^{-8} m^{-1} s^{-1} (see Figs. 4 and 5).

The agreement between predicted and true mass changes is very good at T42. At T85 it is not as good; as noted, the $1° \times 1°$ grid was inadequate for performing the area integrals of the hyperdiffusion term in this case. However, increasing the accuracy of the calculation by increasing the resolution did not significantly improve the agreement; we suspect that our once per day time sampling may be insufficient and responsible for the

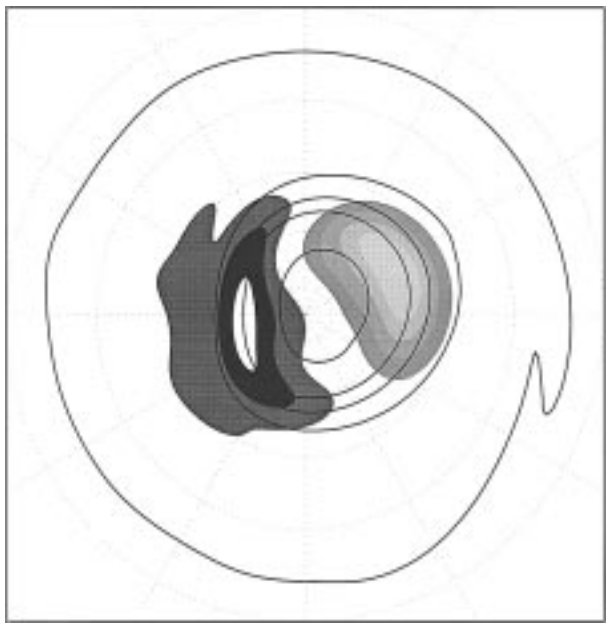

FIG. 4. Thermal forcing of PV (shaded) superimposed on PV (contours) on day 100 of the high-resolution run. Forcing contour interval is 2×10^{-15} m^{-1} s^{-2}; values less than -2×10^{-15} m^{-1} s^{-2} (the blob on the right) are shaded in lighter grays; values greater than 2×10^{-15} m^{-1} s^{-2} (the blob on the left) are shaded in darker grays, except that the highest values (greater than 6×10^{-15} m^{-1} s^{-2}) are left white, and absolute values less than 2×10^{-15} m^{-1} s^{-2} left white as well, so that the color white is double valued to improve contrast. PV contour interval is 1×10^{-8} m^{-1} s^{-1}; minimum contour value is zero.

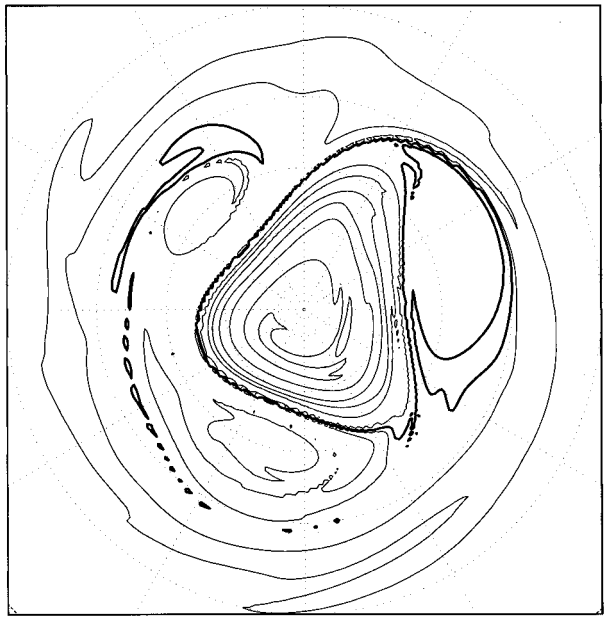

FIG. 5. The PV field resulting from a 5-day RDF run ending on day 90 of the simulation. Contour interval is 5×10^{-9} m^{-1} s^{-1}; heavy contour is 1×10^{-8} m^{-1} s^{-1}.

disagreement. The calculation was repeated for intervals of different duration (i.e., other than 50 days) and similar agreement was obtained. Of course this is no great achievement, since the predicted value is not really predicted but computed directly from known model quantities.

The hyperdiffusion term can be reliably estimated as the difference between the actual mass change and that computed from the MLM mass budget without including hyperdiffusion, circumventing the numerical difficulty mentioned above (note that the terms involving the mass source Q are presumed to have been reliably computed in both runs; the good agreement at T42 is viewed as a demonstration of this for that case, and the Q field varies equally smoothly in space and time in both runs, implying that the calculation should be just as accurate at T85). This calculation is displayed more clearly in Figs. 8 and 9 (to be discussed below) and shows that the hyperdiffusion term is small, though not zero, in the main vortex edge region, which we define here as 1.5×10^{-8} m^{-1} s^{-1} $< P <$ 3.5×10^{-8} m^{-1} s^{-1}. In the outermost vortex edge and surf zone, defined here as $P < 1.0 \times 10^{-8}$ m^{-1} s^{-1}, the hyperdiffusion term becomes much larger. In between there is a transition region, which we will call the entrainment zone (1.0×10^{-8} m^{-1} s^{-1} $< P <$ 1.5×10^{-8} m^{-1} s^{-1}). Of the other two terms [equivalent to the two on the rhs of (5)], the area integral of the height forcing is everywhere negligible compared to the boundary term, as mentioned in section 4b.

As discussed in section 4b, the hyperdiffusion term can be viewed as the degree of radiative disequilibrium at a PV contour. Therefore, Fig. 3 shows that the main vortex edge region is very close to net radiative equilibrium. This is due to almost exact cancellation between regions of positive and negative Q. A contour plot of the PV forcing due to Q on a typical model day is shown in Fig. 4, superimposed on the PV field for that day. In the zonal mean, this PV forcing is everywhere positive throughout the extratropics, but the vortex leans toward the region of negative forcing so that averaged over the main vortex the forcing is approximately zero. This highlights the inadequacy of zonal mean diagnostics for capturing essential aspects of polar vortex behavior.

In the real polar vortex, at least in midwinter, there are probably not large regions of negative thermal PV forcing. Recent radiative transfer calculations show cooling throughout the vortex in monthly averages (Rosenfield et al. 1994), and while there are occasions when regions of radiative warming exist (Rosenfield et al. 1990), these appear too small in spatial extent and magnitude to compete with the regions of cooling in an areal average over the polar vortex. The cooling at any given level is presumably driven by descent forced by the breaking of Rossby or gravity waves at higher levels. Depending partly on the vertical structure of the cooling, this may be expected to increase the PV forcing in the positive sense throughout the vortex in the middle and

lower stratosphere. No analogy to this effect is present in the shallow water model.

Because in the shallow water model the area integral of Q over the main vortex nearly vanishes, the last term in the MLM mass budget [(10)] is, as noted above, negligible relative to the first two, for contours in the outermost vortex edge. The balance at these contours may be viewed as one between two different processes that can change m by changing PV at a contour, rather than by changing the depth of the fluid well within the contour. Intermittently, the contour is deformed into filaments, which are eventually dissipated by the hyperdiffusion, resulting in irreversible entrainment of fluid into the surf zone. Elsewhere on the contour, the thermal forcing tends in the net to increase PV, so that the contour expands to enclose new fluid.

The fact that the hyperdiffusive contribution to the PV budget is quite similar in both runs, throughout most of the domain and despite a factor of 50 difference in the hyperdiffusion coefficients, suggests that the magnitude of this term is determined not by the coefficient *but by the large-scale dynamics*, which is effectively inviscid. In other words, the results suggest that the amount of fluid irreversibly entrained into the surf zone is independent of the hyperdiffusion coefficient. This is reminiscent of the famous result in statistically steady 3D turbulence that the energy dissipation is independent of the viscosity. One might reasonably suppose that changing the form of the operator (say, to ∇^2 instead of ∇^6) should not change the statistically steady MLM mass budget, as long as the operator maintains some scale selectivity and the coefficient is sufficiently small, so that the large-scale dynamics remains inviscid. The exception to this is in the main, unfilamented vortex edge region. One presumes that the small-scale dissipation controls the width of this region directly, since the vortex-stripping process, by itself, would tend to drive the width to zero, as shown quite effectively by truly inviscid contour dynamics models (e.g., Legras and Dritschel 1993). However, as these models have (at least until recently) tended to lack thermal forcing as well, it is not yet entirely clear that thermal relaxation toward a smooth basic state could not arrest the stripping process and yield an edge of finite width in the absence of small-scale dissipation.

6. Transport calculations using RDF and LGR

We applied an alternate approach to diagnosing the rate at which material is stripped from the vortex edge in filamentation events. Our method is similar to the CA-CG method used by Waugh et al. (1994), Plumb et al. (1994), PWP, and Waugh (1996). CA-CG involves first performing a CA calculation (Waugh and Plumb 1994; Norton 1994) on a tracer field, to generate filaments along a contour (which presumably did not have them initially). Then, the contour surgery algorithm (Dritschel 1989a) is applied repeatedly, with successively larger values of a scale cutoff parameter, to remove the filaments; this technique was called coarse graining by Waugh (1992). The material contained in the filaments is then counted as transported across the contour. Again, this procedure is consistent with the MLM approach, given two assumptions. The first is that the filament, once stretched to sufficiently small scales, can have no other fate than eventual dissipation (see SPW for a relevant discussion). The second is that the dissipation affects the MLM mass budget primarily by acting to remove filaments.

Here, we apply the RDF-LGR technique. This differs in its particulars from CA-CG but embodies the same overall approach. In our technique, the initial advection calculation is performed using the RDF method (Sutton et al. 1994; see also Schoeberl and Newman 1995; Yang 1995; SPW) rather than CA. RDF is reviewed in appendix A. Then, the LGR algorithm (described below) is used to perform the ''surgery,'' that is, to remove the filaments. The basic LGR algorithm was introduced by SPW, but in that paper the filaments were not actually removed. Here, the method has been refined so that it is more analogous to contour surgery.

a. Local gradient reversal

1) ALGORITHM

LGR is a simple pattern recognition algorithm that identifies narrow filaments or small blobs in a tracer field that is otherwise smooth in at least one coordinate direction. In SPW, the authors introduced the LGR criteria that determine whether a grid point is part of a filament or not and computed transport by simply summing the areas of these points. Here, a smoothing is applied locally in a region around the filaments, as defined by the LGR criteria. This homogenizes the region around and including each filament, which is equivalent to removing the filament. The total mass contained in the filaments, for a given tracer contour, is estimated by subtracting the mass enclosed by the contour in the processed field from the same quantity in the original field, as in CA-CG. As in CA-CG, this may be considered as representing transport across the contour.

The LGR criteria are reviewed here. A more detailed discussion is given in SPW. Consider a Cartesian domain with coordinates (x, y) and grid points evenly spaced and equal in both dimensions, $\Delta x = \Delta y$. 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 $x = i\Delta x$, $y = j\Delta y$. Those points that satisfy either

$$q(i, j) > q(i - n_{\text{long}}, j) + \delta q_{\text{long}} \quad \text{and}$$
$$q(i, j) > q(i + n_{\text{long}}, j) + \delta q_{\text{long}}$$

or

$$q(i, j) > q(i, j - n_{\text{lat}}) + \delta q_{\text{lat}} \quad \text{and}$$
$$q(i, j) > q(i, j + n_{\text{lat}}) + \delta q_{\text{lat}},$$

with δq_{lat} and δq_{long} free parameters, are considered small-scale PV anomalies in the positive sense, or as being contained in ''outward'' breaking filaments, again using the Northern Hemisphere convention. The criteria for negative anomalies, or ''inward'' breaking filaments, are defined analogously.

In SPW, an equal area grid was used for the LGR calculations. Because of this, in that study n_{lat} and n_{long} were identical. Here, a standard latitude–longitude grid is used. Therefore, we choose

$$n_{\text{long}} = \text{int}[n_{\text{lat}}/\cos(\phi)],$$

where n_{lat} is held fixed, ϕ is the latitude, and int( ) represents rounding off downward to the nearest integer. Though it makes little difference in practice, we also adjust the threshold δq in the longitudinal direction so that

$$\delta q_{\text{long}} = \delta q_{\text{lat}} n_{\text{long}} [n_{\text{lat}}/\cos(\phi)]^{-1}.$$

This means that the critical *gradient* is the same in the two directions, while the critical scale over which this gradient must exist is nearly, but not exactly, the same. Hereafter, δq_{lat} will be referred to simply as δq without the subscript.

In the implementation used in SPW, transport was computed simply by counting the points that satisfied the criteria. Here we have made the transport computation more sophisticated, as mentioned above. We locate the points satisfying the LGR criteria as before; in the discussion that follows these will be referred to as the ''center points.'' Then, however, we generate a smoothed tracer field by applying a smoothing procedure to the field in a region extending for n_{lat} grid points in both latitude and longitude from each center point. The smoothing procedure averages the PV values of the center point with those of its neighbors to make the region locally homogeneous or nearly so.

In this study we have used only $n_{\text{lat}} = 2$ and $n_{\text{lat}} = 4$, for fields on a $1° \times 1°$ latitude–longitude grid. This amounts to assuming a minimum cutoff scale of roughly 200–450 km. The LGR smoothing procedure will be referred to as the ''two-point'' or ''four-point'' smoother, depending on which value of n_{lat} is used. Further, the truncation and smoothing have been applied only to outward breaking filaments, that is, features with PV greater than their environments. This is because we are interested only in the region near the vortex edge, and a careful survey of the model output from the entire period revealed that no inward wave-breaking events occurred there. An event would be considered inward breaking if a thin region of low-PV material were drawn into the vortex so that it was surrounded by large regions of higher PV.

Various versions of the smoothing procedure have been tested on a number of instantaneous model PV fields. These tests have shown that the only small-scale negative PV anomalies (as defined by the LGR criteria) in the vortex edge region occur when small regions of surf zone air are pinched between the main vortex edge and the base of a narrow vortex filament that is being drawn around an anticyclone. It seems inconsistent to call this inward transport at the same time that the formation of the (outward breaking) filament itself is considered outward transport. Lacking justification for making a more complex scheme to deal with this ambiguity, we simply make the consistent choice that a filament formed in an outward breaking event is considered an agent of pure outward transport; that is, we smooth only positive PV anomalies. The algorithm needs further modification to handle a situation in which inward and outward breaking both occur to a significant extent,[5] or in which filaments wrap tightly around the vortex.

In the smoothing procedure, each point involved is smoothed according to

$$\begin{aligned} q(i, j, m+1) &= 0.5q(i, j, m) \\ &\quad + 0.125[q(i+1, j, m) + q(i-1, j, m) \\ &\qquad + q(i, j+1, m) + q(i, j-1, m)], \end{aligned}$$

where i, j are now the indices of each point being affected by the smoothing. Here m is an index indicating how many times the smoothing has been applied; in all calculations below, we iterate the smoothing procedure three times. After three iterations, further application of the smoothing would make little difference since the region has been well homogenized by this point.

The smoothing is applied to each point less than n_{lat} points away from the center point in latitude, longitude, and along the diagonals. Note that the smoothing is applied over the same number of *points* (n_{lat}) in latitude and longitude, although the LGR criteria themselves are applied over equal *distances* in latitude and longitude. This simplification is considered acceptable because first, most of the points satisfying the criteria do so with respect to latitude rather than longitude anyway (that is, the filaments tend to be oriented quasi-zonally), and second, the filamentation events we will study occur within a restricted range of latitudes. Because of this, the spatial inhomogeneity in the smoothing is slight in practice.

2) EXAMPLE

Figure 5 shows the output from a 5-day RDF run ending on day 90, while Fig. 6 shows the same PV field after processing by the two-point smoother with $\delta q = 0.125 \times 10^{-8}$ m^{-1} s^{-1}. Again only positive anomalies, representing outward breaking filaments, have been

[5] In the case of inward breaking only, the signs of the inequalities in the local gradient reversal criteria (section 6a) may simply be reversed.

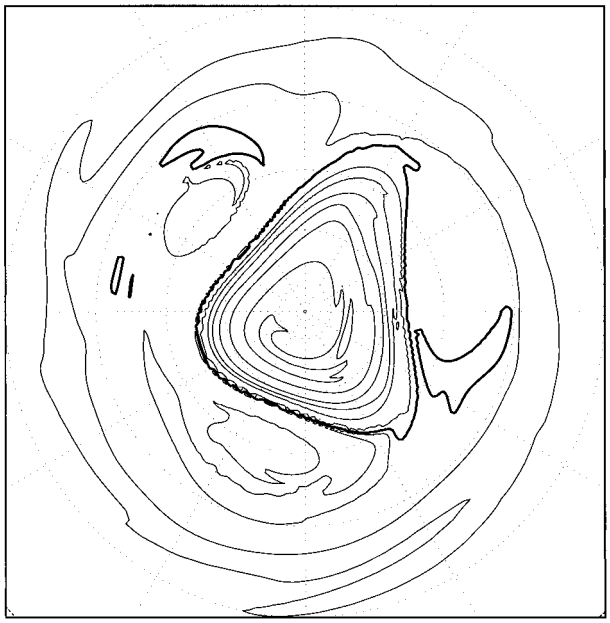

FIG. 6. The PV field resulting from a 5-day RDF run ending on day 90 of the simulation, after application of the two-point LGR filter (see text). Contour interval is 5×10^{-9} m^{-1} s^{-1}; heavy contour is 1×10^{-8} m^{-1} s^{-1}.

FIG. 7. Difference between the fields shown in Figs. 5 and 6. Contour interval is 2×10^{-9} m^{-1} s^{-1}; minimum contour value is 2×10^{-9} m^{-1} s^{-1}.

smoothed. Figure 7 shows the difference between the fields in Figs. 5 and 6.

3) INTERPRETATION

The difference in mass enclosed by a given PV contour in the smoothed and unsmoothed fields may be treated as an estimate of the amount of mass contained in small-scale structures into which that contour has been deformed over the period. We will call this the ''RDF-LGR estimated transport.'' As previously discussed, we assume that the small-scale features identified by the LGR technique are due to be sheared out and eventually dissipated by the hyperdiffusion, and that in a time average, the integral of the hyperdiffusion around filamented PV contours is dominated by the hyperdiffusive dissipation of filaments, so that the effect of the hyperdiffusion along the rest of the contour is negligible. If these assumptions hold, the time integral of the RDF-LGR estimated transport over a long period should be similar to time integral of the hyperdiffusion term in the MLM mass budget over the same period.

b. RDF-LGR procedure

For both model runs, we applied a spectral filter to the PV and wind fields, producing fields truncated at T15 for the 50-day period of interest. The purpose of this is to remove any filaments that may be present in the raw model fields, so that they can be regenerated by RDF calculations. This has the effect of making the subsequent procedure analogous to that which one could perform using observational analyses, the only difference being that the latter would contain some noise. We then performed 10 5-day RDF runs, spanning the period, using the filtered fields as input. Lastly, we applied the two-point and four-point LGR smoothing to the results.

c. Results

Figure 8 compares the contribution of the hyperdiffusion to the MLM mass budget, computed as the residual of the total mass change (the solid curve in Fig. 3) with the terms other than hyperdiffusion (pluses), with the estimate obtained as described above using the two-point LGR filter (solid curves). Figure 9 shows the same quantities, but using the four-point filter. In both plots, three values of δq (where here $q = P$), spanning a factor of 4 in this parameter, are shown. In all cases a smaller δq corresponds to a larger absolute value (deeper minimum) of the fractional mass change.

In the main vortex edge region (1.5×10^{-8} m^{-1} s^{-1} $< P < 3.5 \times 10^{-8}$ m^{-1} s^{-1}), where the hyperdiffusive contribution to the MLM mass budget is small, the LGR technique also estimates a transport that is very small (or zero), regardless of the choice of δq or n_{lat}. This is simply because to a first approximation no filaments are formed from the contours in this region. There does tend to be a small, consistent difference between the hyperdiffusive contribution to the MLM mass budget and the RDF-LGR estimated transport in the main vortex edge region. This is due to the small *direct* effect of hyperdiffusion on the smooth, unfilamented contours

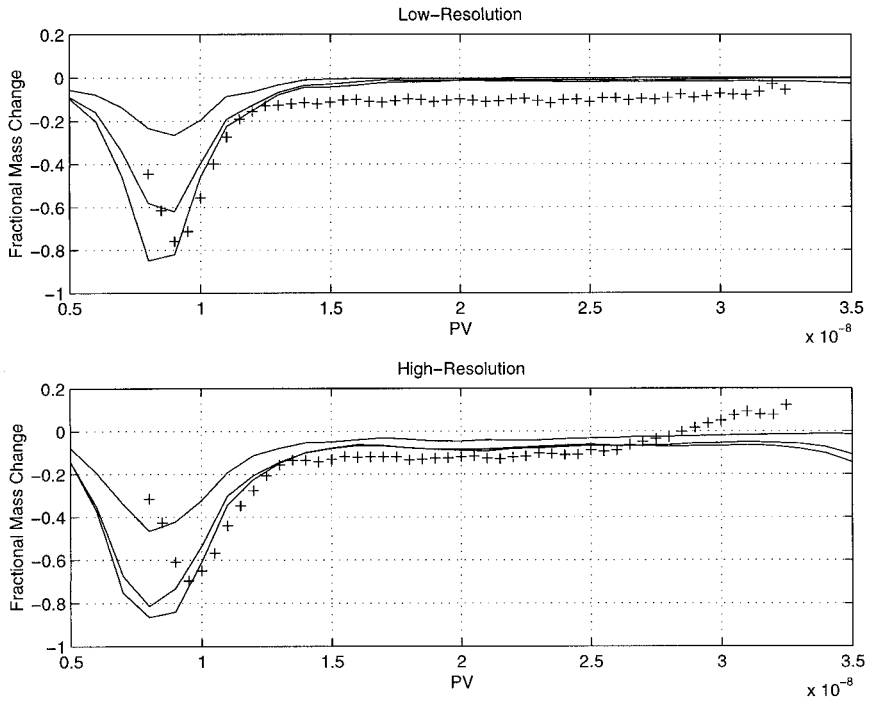

FIG. 8. Mass transported across PV contours in filamentation events between days 80 and 130, estimated by the two-point LGR filter (solid curves) and the hyperdiffusion term in the MLM mass budget (pluses). The three solid curves represent three values of the parameter δq, those being 0.063, 0.125, and 0.250 $\times$ 10^{-8} m^{-1} s^{-1}.

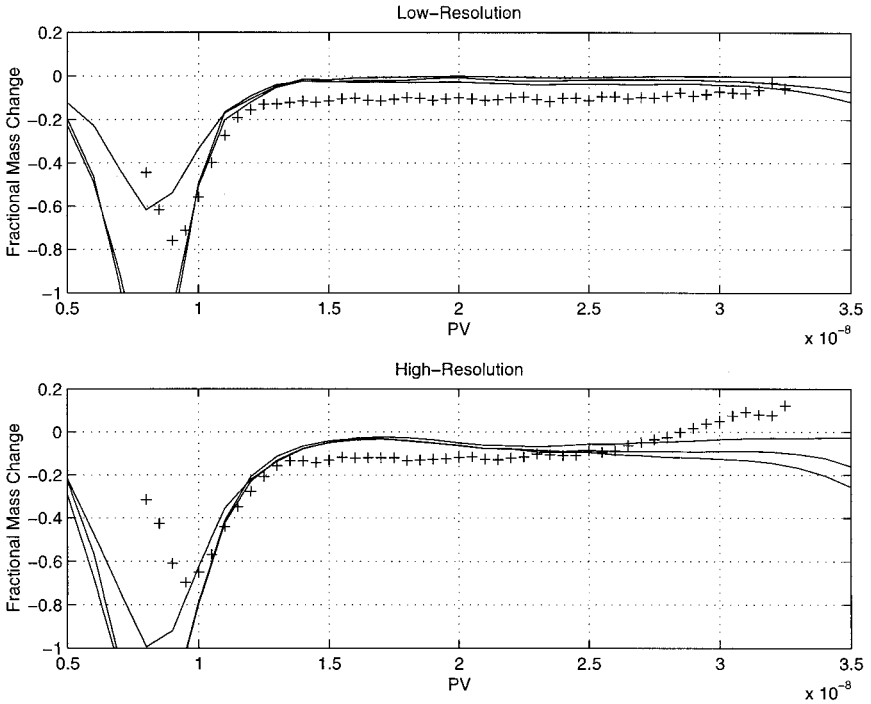

FIG. 9. As in Fig. 8 but with the four-point LGR filter.

in this region, which the LGR technique cannot capture. The details of this transport cannot be taken as realistic, since the hyperdiffusion is a model artifact. What is significant is that both the RDF-LGR technique and the MLM mass budget indicate the presence of a strong transport barrier in the same place, with a net transport of no more than a few percent of the vortex mass per month.

In the outermost vortex edge and surf zone ($P < 1.0 \times 10^{-8}$ m^{-1} s^{-1}), the RDF-LGR and MLM results sometimes agree and sometimes disagree. In any case the results are sensitive to δq and n_{lat} for the parameter ranges chosen; hence, they do not appear to be entirely robust. This region is essentially a ''filament graveyard,'' where previously created filaments and small vortices undergo various processes such as decay and merger. Since this region consists primarily of leftover small-scale structure, it is not surprising that the RDF-LGR technique, which estimates the rate at which small-scale structure is *produced* from large-scale structure, may not give robust results here. This is discussed further in the following section. The lack of a clean scale separation between vortex and filaments, such as exists at higher PV, means that diagnosis of small-scale structures depends strongly on the details of the definition.

In the entrainment zone (1.0×10^{-8} m^{-1} s^{-1} $< P <$ 1.5×10^{-8} m^{-1} s^{-1}) the RDF-LGR and MLM results agree quite well in general. This means that the RDF-LGR technique, which is an ad hoc but plausible way to estimate the rate at which material is irreversibly entrained from the vortex edge into the surf zone, produces estimates that correspond closely to those produced by a fully rigorous method. As shown in section 4b, the entrainment rate thus estimated is directly related to the radiative disequilibrium at the outer edge PV contours, which in steady state equals the convergence of mass into those contours. These are well-defined, macroscopic properties. Remember that the hyperdiffusion is the only process in the model that can balance any noncancellation of the terms involving the mass source, which is the nearest shallow water analog to radiative heating.

d. Sensitivity to duration of calculation

Though it is true that, as we have discussed, the only explicit parameters in our LGR filtering scheme are n_{lat} and δq, there is really one more parameter in our RDF-LGR procedure whose role requires some scrutiny. This is the duration of the RDF calculation. The RDF technique, like CA, is absolutely nondiffusive, but unlike CA it is not absolutely conservative. With either RDF or CA, as the duration of a calculation is extended, finer and finer scale features develop. In the case of RDF, however, features smaller than the grid spacing in any dimension cannot be resolved. As the run's duration is increased for a fixed initial time, the small-scale features eventually simply disappear. For a given grid spacing, then, the optimal run duration would be long enough to allow some finescale structure to develop, but short enough that that structure does not disappear.

To test the LGR procedure's sensitivity to the run duration, we repeated the calculations for both model runs and both values of n_{lat} using 5 runs of 10 days each spanning the period, instead of 10 runs of 5 days each. The results, shown in Figs. 10 and 11, are somewhat smaller transport estimates than in the 5-day run case and are smaller than the MLM results by up to a factor of 2. Note that, although the slope of the curve in the entainment zone is still correct, the curve is displaced to slightly lower PV. Note also that the depth of the minimum appears to be fairly robustly and accurately predicted in this case, though it was not so robustly predicted in the 5-day calculations mentioned in the previous section. It is not clear whether the more consistent agreement in the 10-day calculations is an accident, or whether the technique can, if applied subject to some a priori constraints on the parameters (which remain undetermined here), reliably predict this minimum in general.

To investigate the issue of duration depencence further, we took a single 10-day period, divided it up into periods of periods of 1, 2, 3, 5, and 10 days' duration, and repeated the RDF-LGR calculations in each case (for the 3-day periods, the calculations span 9 days rather than 10; the results have been multiplied by 10/9 to compensate). The results of this are shown in Fig. 12. The six curves represent the three values of δq used in the earlier calculations, with larger transport corresponding to smaller δq, and the two values of n_{lat} (see figure caption). A 10-day period is not long enough that we can reasonably expect the RDF-LGR calculation to agree with the hyperdiffusive contribution to the MLM mass budget over this period; remember that we expect the two calculations to agree only in a long-time average. Rather than having a verifiably correct value to which to compare, then one ought simply to look for something like convergence, at least in an asymptotic sense. For durations of 1 or 2 days, the result is strongly dependent on δq. At 10 days, this dependence is largely gone. However, between 5 and 10 days one begins to see a clear downward trend in the computed transport. This may be expected to continue at longer durations, as filaments have more time to vanish during the course of a single RDF run. The optimal parameter choices are those for which the dependence on parameters is the weakest, since then the technique is most robust. At a run duration of 10 days, the dependences on n_{lat} and δq are smallest, but the dependence on duration itself appears to have a stationary point somewhere between 5 and 10 days, particularly if the results for the highest value of δq (0.25×10^{-8} m^{-1} s^{-1}) are discarded. The latter are outliers in most of the calculations done, in any case; the dependence on δq appears minimized (indeed, appears nearly to vanish in the entrainment zone) for the higher two values of δq (0.125 and 0.063×10^{-8}

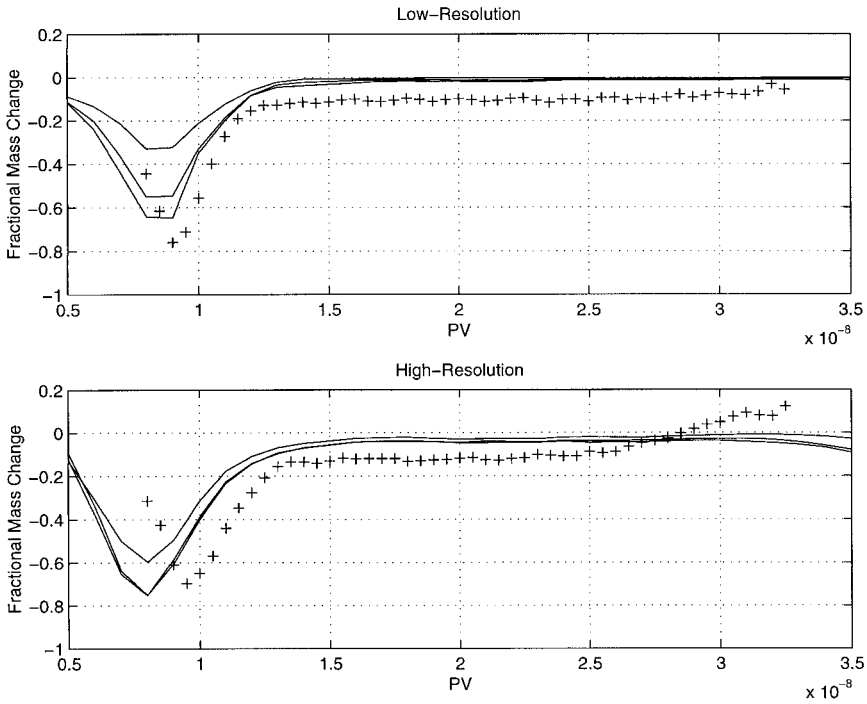

FIG. 10. As in Fig. 8 but with RDF run durations of 10 rather than 5 days.

m^{-1} s^{-1}), which still differ from one another by a factor of 2. Note, also, that the dependence on n_{lat} and δq appears to be smaller in an average over a larger number of 5-day RDF runs (see Figs. 8 and 9) than in this case where the total period length is only 10 days.

One can conclude from these results that there is an optimal range of RDF run duration, centered somewhere between 5 and 10 days, within which the results are relatively insensitive to duration and to the other two parameters. This conclusion has been obtained inde-

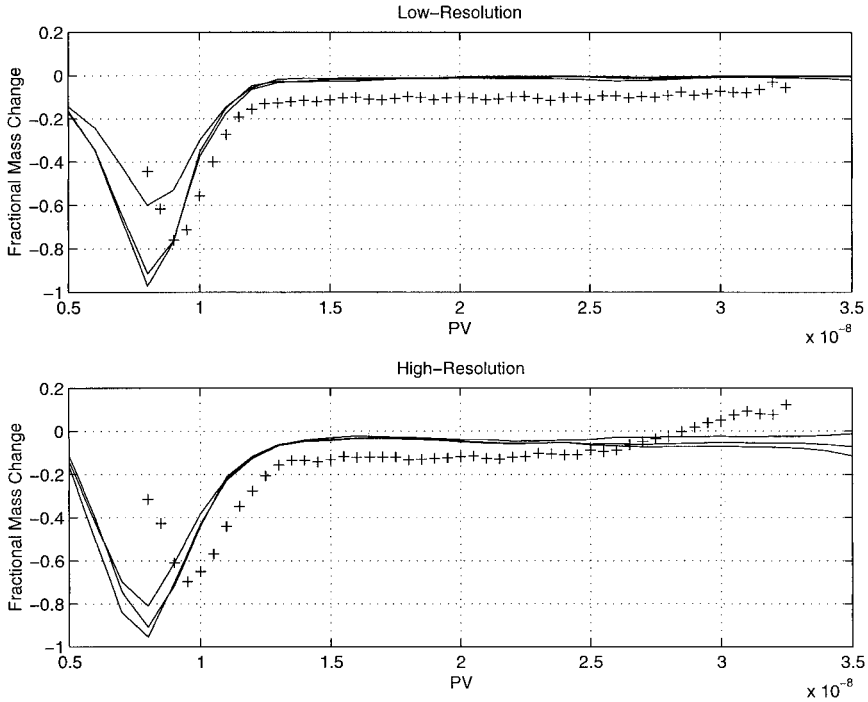

FIG. 11. As in Fig. 9 but with RDF run durations of 10 rather than 5 days.

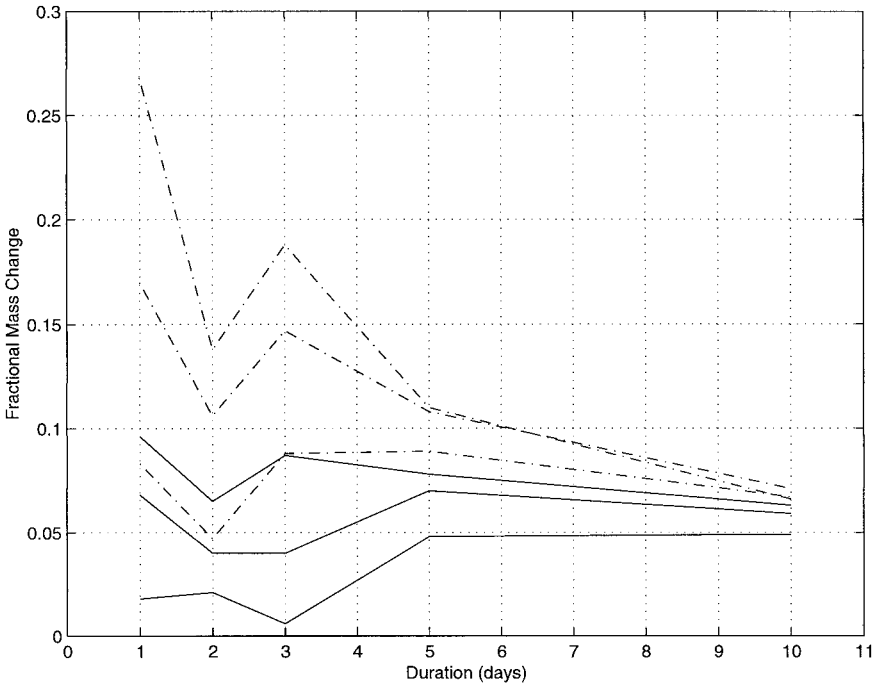

FIG. 12. Transport across the 1.1×10^{-8} m^{-1} s^{-1} contour over a single 10-day period, computed by the LGR technique with RDF run durations varying from 1 to 10 days. The six curves represent the three different values of δq used in the preceding figures, and two values of n_{lat}: $n_{\text{lat}} = 2$ (solid) and $n_{\text{lat}} = 4$ (dot–dash). For each n_{lat}, larger δq corresponds to smaller fractional mass change at a duration of 1 day.

pendently of the comparisons to the MLM results, by considering only the dependence of the RDF-LGR results on parameter variations. It can then be viewed as showing that the good agreement between the MLM and RDF-LGR calculations is robust, rather than being a consequence of an arbitrary parameter choice.

e. Comparison to CA-CG

PWP used the CA-CG technique to estimate transport across the vortex edge, which they defined as the 1.125×10^{-8} m^{-1} s^{-1} PV contour. They computed transport every 10 days of their 100-day run, which is identical to the first 100 days of our T42 run. We have computed transport above for days 80–130. For comparison to PWP's results, we also used LGR to compute transport across the 1.125×10^{-8} m^{-1} s^{-1} contour between days 60 and 100. For this comparison we did calculations of 10 days' duration, and unlike in the calculations presented above, we did *not* perform any spectral truncation or other processing on the wind or PV fields prior to the RDF runs, since PWP did not do so. Again following their procedure, we performed our transport calculations in terms of area rather than mass. This tends to decrease the transport by about 20% when expressed as a fraction of the main vortex, since the fluid is thicker in the outer vortex edge (where the filaments come from) than in the interior.

TABLE 1. Comparison of transport calculations using the RDF-LGR technique vs those using the CA-CG technique, as done by PWP. The numbers in parentheses correspond to δq (in units of 10^{-8} m^{-1} s^{-1} for the RDF-LGR results) and to *dm* (the dimensionless maximum surgery scale for the CA results. Transport is expressed as a fraction of the area enclosed by the contour on day 70.

	LGR vs CA comparison			
Day	LGR (0.063)	LGR (0.125)	CA (0.10)	CA (0.05)
60–70	1.1	1.0	1.0	0.9
70–80	4.0	3.6	4.5	3.9
80–90	3.4	2.4	3.2	2.6
90–100	0.3	0.1	0.2	0.2
Total	8.8	7.1	8.9	7.6

Table 1 shows the results of this comparison. Outward transport across the contour for each 10-day period is shown, as well as the total over the entire 40-day period. Results are shown for the two lower values of δq as used above ($\delta q = 0.25 \times 10^{-8}$ m^{-1} s^{-1} is not shown, as the analysis above showed it to be a less satisfactory value), as well as two different values of the maximum surgery cutoff scale (the relevant parameter in the coarse-graining technique used by PWP) for the CA-CG results.[6] For the RDF-LGR results the two-point

[6] The CA-CG results, some of which were not included in PWP, were kindly provided by D. Waugh (1997, personal communication).

LGR smoother was used; results using the four-point smoother (not shown) are not significantly different.

Excepting the RDF-LGR results with $\delta q = 0.25 \times 10^{-8}$ m^{-1} s^{-1} (which value also yielded the least satisfactory agreement with the MLM results), the agreement between the LGR and CA-CG results is remarkably good. This is consistent with the idea that because of the clear scale separation between the vortex and the filaments, where the distinction between the two is drawn is relatively insensitive to the precise criteria used to draw it. It also means that the comparison to the MLM calculations can be viewed as providing concrete support for CA-CG as much as for RDF-LGR.

7. Discussion

The balance maintaining the model's polar vortex–surf zone structure has been quantitatively diagnosed in detail. The processes that enter the balance directly, in the MLM framework, are the large-scale radiative mass source and the small-scale hyperdiffusion. The quasi-inviscid large-scale dynamics links these two nonconservative processes through the stretching-induced scale cascade.

We proved that if an inviscid shallow water atmosphere is in an MLM steady state, then the mass source enclosed by PV contours must vanish, which is the shallow water version of radiative equilibrium. This result extends earlier results that made the stronger assumption of steady flow fields, and clarifies the relationship between the different processes in the shallow water model. Any net mass convergence into PV contours, or in steady state, any departure from radiative equilibrium, is directly related to the viscous (or hyperviscous) dissipation of PV. This feature of the MLM formulation stands in contrast to the Eulerian or transformed Eulerian mean formulation, in which mean heating is a necessary consequence of eddy thermal dissipation even in the absence of friction, in a steady state.

We diagnosed the contributions to the strict MLM mass budget from both nonconservative model processes for PV contours in and near the vortex edge region. The main vortex edge region was shown to be near MLM radiative equilibrium, meaning that the integral of the mass source over the area enclosed by the PV contours nearly vanishes. This is not due to the source's being negligibly small everywhere within these contours, but rather due to the fact that large regions of mass source and sink exist within the contours, whose effects nearly cancel in the area (or contour) integral. In contrast, the outermost vortex edge–surf zone region is far from radiative equilibrium, implying that this region behaves, in a sense, as though it were much more viscous. This is clearly due to the prevalence in this region of small-scale structures, which are more readily subject to dissipation.

It was shown that the net hyperdiffusive contribution to the mass budget, even in regions where that contribution is large, is at most weakly dependent on the value of the hyperdiffusivity. This supports the hypothesis that, in a time-averaged sense, the rate at which vortex edge fluid is entrained into the surf zone (which, in steady state, equals the rate at which it is replaced by thermal forcing) is not sensitive to the manner in which vortex filaments are dissipated, as long as the dissipation is sufficiently weak and scale selective that it does not directly affect the dynamically active large scales. To be sure, a more thorough test of this hypothesis would involve varying the form as well as the coefficient of the dissipative operator. Ideally, one would like to compare the results of a ∇^6 operator to the more physically plausible ∇^2. However, much higher resolution than T85 is required in order to run the model with a ∇^2 diffusion and still obtain an inertial range, that is, vortex filaments. The limited computational resources available for this study precluded this step.

The RDF-LGR technique was used to estimate the transport across the contours due to filamentation events. This technique incorporates the assumption that filaments are due eventually to be dissipated once they reach a minimum scale, regardless of the details of the dissipative process. By also assuming that the small-scale dissipation contributes to the MLM mass budget primarily by dissipating these filaments, the time-averaged net transport across the contours due to the small-scale dissipation can be estimated simply by estimating the mass contained in filaments produced during a given period of time. The LGR filter makes use of a particular definition of a filament. The similarity of the RDF-LGR and CA-CG results shows that at least in a certain regime, the results will not be sensitive to the precise definition.

The transport computed by the RDF-LGR technique was shown to agree with the the hyperdiffusion term in the MLM budget, except in the outermost vortex edge and surf zone. In the main vortex edge, the RDF-LGR technique showed essentially zero transport. This is in near agreement with the MLM results that show this region to be near radiative equilibrium. The MLM results show a small hyperdiffusive transport in this region, which could not be reliably estimated by LGR due to the transport's being generally unrelated to filament production.[7] In the entrainment zone, the RDF-LGR-estimated transport agreed remarkably well with the MLM hyperdiffusion term. The RDF-LGR results in the main vortex edge and entrainment zone are fairly insensitive to variations in the scale cutoff parameter n_{lat}, the amplitude threshold δq, and the RDF run duration.

The integral constraint on a hypothetical inviscid po-

[7] There were a few RDF calculations that did show weak filamentation in the main vortex edge, which is why the RDF-LGR-estimated transport is not always identically zero there. This filamentation was never visible in the raw model output, presumably having been suppressed by the hyperdiffusion.

lar vortex shows that the shallow water polar vortex studied here can, insofar as it is statistically steady, depart from MLM radiative equilibrium only because of the presence of the hyperdiffusion. The magnitude of the hyperdiffusion term in the MLM mass budget may thus be viewed as a quantitative measure of the degree of radiative disequilibrium, or equivalently the mass convergence into the contours; this is then a well-defined, macroscopic, physically meaningful quantity. The numerical results discussed above show that the RDF-LGR technique can provide a reliable estimate of this quantity over much of the region of interest, despite incorporating no explicit information about the nature of the nonconservative processes. This shows that this quantity, despite being ostensibly a direct measure of nonconservative effects, *is in fact controlled, or at least strongly mediated, by quasi-inviscid, large-scale dynamics*. It also provides direct support for both the RDF-LGR and (indirectly) the CA-CG method, and guidance in interpreting results from these methods.

In the surf zone, the RDF-LGR results sometimes disagree substantially with the MLM results and are very sensitive to the parameters. This is not surprising, as the technique estimates the rate at which small-scale structure is *produced*, while the surf zone consists primarily of small-scale structure that is left over from previous wave-breaking events. The clean scale separation between dynamically active and passive structures is lacking in this region.

The RDF-LGR technique was applied to low-resolution versions of the model wind and PV fields and requires no knowledge of the small-scale dissipation. Hence, the technique may be applied as is to observations. The same can be said of CA-CG, which already has been applied to observations (Waugh et al. 1994; Plumb et al. 1994; Waugh 1996); the present results may be viewed as providing further support for the results of these studies. The only important difference between observationally derived fields and the smoothed model fields used here is the presence of noise in the former. The effect of noise has not been directly addressed here, but the results presented by SPW (using a precursor of the technique developed here) suggest that RDF-LGR and similar techniques (i.e., CA-CG) will be less sensitive to modest amounts of noise in the data than more direct approaches such as ''contour crossing.''

Acknowledgments. We are much indebted to Lorenzo Polvani for providing and assisting us with the shallow water model, and to Darryn Waugh for providing the CA-CG results, assisting us in interpreting them, and providing helpful comments on the first draft of this paper; we also thank Drs. Polvani and Waugh for many stimulating discussions on this research and related subjects. A suggestion from Glenn Flierl set us on a line of investigation that led to the main result of section 4b, and Dr. Flierl subsequently helped us to fill a logical gap in the proof of this result. Peter Haynes and Martin Juckes made us aware of some useful references and helped us to understand the relationship between the various results discussed in appendix B. Noboru Nakamura and Keith Ngan provided insightful and constructive comments on the first draft of this paper; we particularly thank Dr. Nakamura for catching a couple of errors. We also thank David Dritschel, Kerry Emanuel, Isaac Held, Paul Kushner, and Richard Lindzen for discussions. This research was supported by NASA Grant NAGW-1727, and by a NASA Graduate Student Fellowship in Earth System Science for AHS. AHS was supported during the final stages of manuscript preparation and revision by the NOAA Postdoctoral Program in Climate and Global Change, administered by the University Corporation for Atmospheric Research. Computer time used for the shallow water model integrations was provided by a grant from the Pittsburgh Supercomputing Center.

APPENDIX A

Reverse Domain Filling

The reverse domain filling (RDF) technique (Sutton et al. 1994; Schoeberl and Newman 1995) may be viewed as a nondiffusive offline advection scheme that can be applied for a predetermined, finite time period. Given a set of wind fields covering some particular time period, and a tracer field at the beginning of that period, RDF allows the generation a high-resolution tracer field at the end of the period that, in general, includes structure at smaller scales than were present in the initial field.

RDF works as follows. One initializes a large number of Lagrangian trajectory calculations with a set of hypothetical point parcels located at the nodes of a regular grid. One then performs the trajectory calculations backward in time, starting at the time at which the high-resolution field is desired (the final time) and ending at some earlier time (the initial time). At the initial time, the parcels are given tracer values by interpolating the given tracer field to the parcels' positions at that time. Finally, the parcels are placed back on the nodes where their trajectories were initialized, but labeled by the tracer values they have been given as a result of the calculation. One is then left with a numerical representation of a tracer field on a regular grid, representing the result of advecting the initial tracer field by the given flow from the initial until the final time. The RDF calculations discussed here were performed on a $1^\circ \times 1^\circ$ latitude–longitude grid.

Comparisons of RDF and CA have been performed by Schoeberl and Newman (1995), Yang (1995), and SPW.

APPENDIX B

Relationship between Different Integral Constraints

This appendix explains the relationship between the results of Schneider (1987), McIntyre and Norton (1990), Juckes (1987), and the main result of section 4b.

First, write the inviscid, shallow water vorticity equation in flux form:

$$\frac{\partial \zeta_a}{\partial t} + \nabla \cdot (\mathbf{u}\zeta_a) = 0.$$

Then perform an area integral of the entire equation over the area enclosed by some contour, for the moment unspecified, and note that ζ_a is simply the layer thickness times the PV. The result is

$$\int\int \frac{\partial (hq)}{\partial t}\, dA + \int\int \nabla \cdot (\mathbf{u}hq)\, dA = 0. \quad \text{(B1)}$$

Both Schneider (1987) and McIntyre and Norton (1990) proceeded by assuming a steady state ($\partial/\partial t = 0$), so that the first term in (B1) vanishes. The difference between the two results is that Schneider let the contour be one of constant absolute vorticity, so that he obtained

$$\zeta_a \int\int \nabla \cdot \mathbf{u}\, dA = 0,$$

implying that if the contour has a value other than zero, the integrated flow divergence ($\nabla \cdot \mathbf{u}$) must vanish. McIntyre and Norton let the contour be one of constant PV, so that they obtained

$$q \int\int \nabla \cdot (\mathbf{u}h)\, dA = 0,$$

so that for a nonzero PV contour the integrated mass flux divergence [$\nabla \cdot (\mathbf{u}h)$] must vanish. Nominally, they derived their result in three dimensions, so that the thickness h was replaced by the isentropic density σ. However, they made an additional assumption that has the same effect as assuming no vertical gradient of absolute vorticity, which makes their system formally identical to the shallow water system.

One could also let the contour be one of constant thickness, in which case the result would be

$$h \int\int \nabla \cdot (\mathbf{u}q)\, dA = 0,$$

or no net PV flux into or out of the contour. Another possibility that yields an interesting result is to choose a streamfunction contour, as discussed by Sardeshmukh and Hoskins (1988).

The difference between all of these results and the one presented in section 4b is that the latter assumes only that the mass enclosed by a PV contour is steady; *the flow fields need not be*. In other words, the condition $\partial m/\partial t = 0$ does *not* imply that $\partial(u, v, h, \ldots)/\partial t = 0$, where the quantities in parentheses are the shallow water flow variables. In the results of Schneider (1987) and McIntyre and Norton (1990), one must assume $\partial/\partial t = 0$ everywhere, or else the first term in (B1) does not vanish. For a contour that is moving, taking the time derivative outside the integral in that term results in the introduction of a boundary term. The contribution of the present work is to show that this does not change the result.

Juckes (1987) proved a related approximate result. He proved that under quasigeostrophic scaling, in three dimensions, the net mass flux across a PV contour is smaller by a factor of order Rossby number than the zonal mean mass flux. Juckes did *not* assume a pure steady state, but rather only a steady state of the sort assumed in section 4b, that is, with the time derivative outside the integral (before it is taken to be zero).

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Browning, G. L., J. J. Hack, and P. N. Swarztrauber, 1989: A comparison of three numerical methods for solving the shallow water equations on the sphere. *Mon. Wea. Rev.*, **117,** 1058–1075.

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. *Comput. 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.

Frisch, U., 1995: *Turbulence.* Cambridge University Press, 296 pp.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. P. Rood, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33,** 403–439.

Juckes, M. N., 1987: Studies of stratospheric dynamics. Ph.D. thesis, University of Cambridge, 162 pp. [Available from DAMTP, Silver Street, Cambridge CB3 9EW, United Kingdom.]

——, 1989: A shallow water model of the winter stratosphere. *J. Atmos. Sci.*, **46,** 2934–2955.

——, 1996: Comments on ''On the subtropical edge of the stratospheric surf zone.'' *J. Atmos. Sci.*, **53,** 3770–3771.

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

McIntyre, M. E., 1980: Towards a Lagrangian-mean description of stratospheric circulations and chemical transports. *Philos. Trans. Roy. Soc. London*, **296A,** 129–148.

——, and W. A. Norton, 1990: Dissipative wave-mean interactions and the transport of vorticity or potential vorticity. *J. Fluid Mech.*, **212,** 403–435.

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.

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.

Plumb, R. A., and Coauthors, 1994: Intrusions into the lower strato-

spheric arctic vortex during the winter of 1991/92. *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.

——, ——, and ——, 1996: Reply. *J. Atmos. Sci.,* **53,** 3772–3775.

Rosenfield, J. E., M. R. Schoeberl, L. R. Lait, P. A. Newman, M. H. Proffitt, and K. K. Kelly, 1990: Radiative heating rates during the Airborne Arctic Stratospheric Experiment. *Geophys. Res. Lett.,* **17,** 345–348.

——, 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.

Salby, M. L., R. R. Garcia, D. O'Sullivan, and P. Callaghan, 1990: The interaction of horizontal eddy transport and thermal drive in the stratosphere. *J. Atmos. Sci.,* **47,** 1647–1665.

Sardeshmukh, P. D., and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.,* **45,** 1228–1251.

Schneider, E. K., 1987: A simplified model of the modified Hadley circulation. *J. Atmos. Sci.,* **44,** 3311–3328.

Schoeberl, M. R., and P. A. Newman, 1995: A multiple level trajectory analysis of vortex filaments. *J. Geophys. Res.,* **100,** 25 801–25 816.

Sobel, A. H., R. A. Plumb, and D. W. Waugh, 1997: Methods of calculating transport across the polar vortex edge. *J. Atmos. Sci.,* **54,** 2241–2260.

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., 1992: The efficiency of symmetric vortex merger. *Phys. Fluids,* **4A,** 1745–1758.

——, 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.

Yang, H., 1995: Three-dimensional transport of the Ertel potential vorticity and N_2O in the GFDL SKYHI model. *J. Atmos. Sci.,* **52,** 1513–1528.