A study on assimilating potential vorticity data

By YONG LI*,†,‡, RICHARD MÉNARD‡, LARS PETER RIISHØJGAARD‡, STEPHEN E. COHN
and RICHARD B. ROOD, *Data Assimilation Office, NASA/Goddard Space Flight Center, Code 910.3,
Greenbelt, Maryland 20771, USA*

ABSTRACT

The correlation that exists between the potential vorticity (PV) field and the distribution of chemical tracers such as ozone suggests the possibility of using tracer observations as proxy PV data in atmospheric data assimilation systems. Especially in the stratosphere, there are plentiful tracer observations but a general lack of reliable wind observations, and the correlation is most pronounced. The issue investigated in this study is how model dynamics would respond to the assimilation of PV data. First, numerical experiments of identical-twin type were conducted with a simple univariate nuding algorithm and a global shallow water model based on PV and divergence (PV-D model). All model fields are successfully reconstructed through the insertion of complete PV data alone if an appropriate value for the nudging coefficient is used. A simple linear analysis suggests that slow modes are recovered rapidly, at a rate nearly independent of spatial scale. In a more realistic experiment, appropriately scaled total ozone data from the NIMBUS-7 TOMS instrument were assimilated as proxy PV data into the PV-D model over a 10-day period. The resulting model PV field matches the observed total ozone field relatively well on large spatial scales, and the PV, geopotential and divergence fields are dynamically consistent. These results indicate the potential usefulness that tracer observations, as proxy PV data, may offer in a data assimilation system.

1. Introduction

In the stratosphere, there is a general lack of reliable wind observations. Aside from radiosonde reports in the lower stratosphere, the only source of stratospheric wind measurements at present originates from the High Resolution Doppler Image (HRDI) on board the Upper Atmosphere Research Satellite (UARS). However, these data

have been little used because they are daylight-only retrievals, have large observational biases, and have shown only modest impact in stratospheric assimilations (Richard Swinbank, personal communication). Traditionally, it has been attempted to obtain wind observations indirectly, ever since Charney et al. (1969) showed that the wind field could be recovered through a model with temperature being replaced periodically by observations. Preliminary success was documented in their idealized experiments, but problems such as unrealistic gravity waves due to data insertion were encountered. Today satellite temperatures or radiances are being assimilated routinely and spurious gravity waves are ameliorated by multivariate covariance modeling and various types of initialization. Nevertheless, wind information remains indirect with wind analysis incre-

---

* Corresponding author: Yong Li, JCET/UMBC and Data Assimilation Office, NASA/Goddard Space Flight Center, Code 910.3, Greenbelt, Maryland 20771, USA.

† Current affiliation: Danish Centre for Earth System Science, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen O, Denmark.

‡ Also affiliated with: Joint Center for Earth Systems Technology, The University of Maryland Baltimore County, Baltimore, Maryland 21228, USA.

ments being essentially geostrophic. Space-based lidars have been proposed for future stratospheric and tropospheric wind measurements, but they present technological difficulties and consequent expense. Alternative technologies such as the Stratospheric Wind Interferometer For Transport Studies (SWIFT) are under development and will require some time to become operational (Rowlands et al. 1996; Tarasick et al. 1996).

Another source of wind information comes from measurements of chemical constituents. Many constituents (e.g., O_3, CH_4, and N_2O) can be considered as tracers since they have long lifespans. Recently, efforts to infer wind fields from tracer data have been made by using tracer transport models. These approaches involve the coupling of the transport model and the dynamical model to retrieve the wind information through an advanced dynamical assimilation method as illustrated by Daley (1995; 1996) with a Kalman Filter and Riishøjgaard (1996) with a 4-dimensional variational (4D-Var) method.

One purpose of this study is to examine whether dynamical information can also be extracted from chemical tracer observations by using them as proxy data for potential vorticity (PV). It has been noticed that the potential vorticity (PV) field and the ozone field are often highly correlated, particularly in the lower stratospheric and upper tropospheric extra-tropics. Early observational evidence was presented by Danielsen (1968) and was further confirmed and examined by Danielsen et al. (1987). Recent studies which suggest, examine or make use of such correlations include Douglass et al. (1990), Lait et al. (1990), Allaart et al. (1993), Lary et al. (1995), Riishøjgaard and Källén (1997) and Riishøjgaard et al. (1997). These studies suggest the possibility of using tracer information as proxy PV data.

It has been known for some time that PV is a dynamically significant quantity (see e.g., Hoskins et al. 1985 for a review). The notion of PV as a preferred field to which other variables are slaved has been discussed in many published articles (Allen 1993; Warn et al. 1995). One may even derive both the wind field and the mass field from the global distribution of PV if certain balance conditions are assumed. For example, Vallis (1996) proposed a variety of models based either on the shallow water equations or on the stratified primitive equations. Different balance conditions were proposed in an attempt to describe different types of flow. It was demonstrated in the numerical experiments that other dynamical quantities can be recovered with some success through these PV inversion approaches, although some limitations are imposed by the balance conditions. It was indeed noticed that the reconstruction of the divergence field is, in most situations, unsatisfactory.

We can expect improvements if balance conditions are avoided, such as in the case of primitive equation models. The PV field, however, contains information concerning both the wind field and the mass field. There are nonunique combinations of these two fields which would yield the same PV if no balance condition is assumed. How, then, would a primitive equation model respond to externally imposed PV data? Can an appropriate dynamical adjustment between the wind field and the mass field be accomplished inherently by a primitive equation model? We attempt to address these questions through the present study.

We use a global shallow water model based on the advection of PV (referred to as the PV-D model hereafter) which was developed by Bates et al. (1995) and extended to include orography by Li and Bates (1996). A simple nudging method is adopted here for assimilating proxy PV data. The method is intentionally univariate in PV, with no externally-imposed balance constraint. Though this method serves as a simple and useful test of the idea of using tracer observations as proxy PV data, it cannot fully exploit the tracer property of the data, as can be accomplished by including a tracer transport model and involving a fully 4-dimensional assimilation method such as the Kalman Filter or 4D-Var (Daley 1995, 1996; Riishøjgaard 1996). On the other hand, because of this, the method can be more easily extended to be applied to a static operational data assimilation system such as the 3D-Var or PSAS (Physical Space Statistical Analysis System).

The paper is planned as follows. In Section 2, we briefly describe the PV-D model and the data assimilation algorithm. In Section 3, we describe the design of numerical experiments in which synthetic PV data are assimilated and we discuss the results. In Section 4, a linear analysis of these

results is presented. In Section 5, we present results of assimilating Total Ozone Mapping Spectrometer (TOMS) total ozone observations as proxy PV data. Finally in Section 6, we offer concluding remarks.

2. The shallow water PV-D model and the data assimilation algorithm

2.1. The PV-D model

We present a brief description of the model here. A detailed description can be found in Bates et al. (1995) and in Li and Bates (1996) with orography included.

The model equations consist of the PV, divergence and continuity equations. The PV equation is

$$\frac{d}{dt}\left(\frac{\nabla^2\psi + f}{\Phi}\right) = 0, \tag{1}$$

where ψ is the streamfunction, f the Coriolis parameter, and Φ the geopotential. After semi-Lagrangian semi-implicit (SLSI) discretization, this equations becomes

$$\left(\frac{\nabla^2\psi + f}{\Phi}\right)^{n+1}$$
$$+ \frac{\beta\tau_1'}{\Phi_*^n}\left[\frac{1}{a\cos\phi}\frac{\partial\psi}{\partial\lambda} + \frac{1}{a}\frac{\partial\chi}{\partial\phi}\right]^{n+1} = r_1, \tag{2}$$

where χ is the velocity potential, β the meridional derivative of f, r_1 the forcing term (Li and Bates 1996), $\tau_1' = \frac{1}{2}\Delta t$ with Δt denoting the time step, and $(\ )_*$ denotes a departure-point value. The term associated with β arises from a centered implicit discretization of the β-term after a Taylor expansion of f_*. Such a treatment alleviates the numerical instability of Rossby waves (Bates et al. 1995).

The SLSI-discretized divergence equation is obtained through first discretizing the horizontal momentum equation in its vector form and then applying the divergence operator, which gives

$$\left(\nabla^2\chi + \tau_1\left[\nabla^2\Phi - f\nabla^2\psi\right.\right.$$
$$\left.\left. + \beta\left(\frac{1}{a\cos\phi}\frac{\partial\chi}{\partial\lambda} - \frac{1}{a}\frac{\partial\psi}{\partial\phi}\right)\right]\right)^{n+1}$$
$$= r_2 - \tau_1\nabla^2\Phi_s, \tag{3}$$

where $\tau_1 = \frac{1}{2}(1 + \varepsilon)\Delta t$, with ε being the first-order uncentering parameter (chosen to be zero in this study), Φ_s is the geopotential height of surface orography, and r_2 the forcing term defined as in Li and Bates (1996).

For the continuity equation,

$$\frac{d\Phi}{dt} = -\Phi\nabla^2\chi, \tag{4}$$

the SLSI discretization is straightforward:

$$\Phi^{n+1}(1 + \tau_1\nabla^2\chi)^{n+1} = r_3. \tag{5}$$

This model, with an independent equation for PV, conveniently facilitates our study on PV data assimilation.

2.2. The data assimilation algorithm

The data assimilation algorithm is a simple nudging technique which has been studied extensively (Davies and Turner 1977; Kuo et al. 1993) since its early use (Anthes 1974). This technique forces model fields toward the corresponding available data fields through linear terms proportional to the model-data misfit. Because of its simplicity, the algorithm is a useful tool for the proof-of-concept type of study presented here.

When we insert PV data, the only equation that we force directly is the PV equation, which then becomes

$$\frac{d}{dt}\left(\frac{\nabla^2\psi + f}{\Phi}\right) = \alpha_{pv}(PV_{data} - PV), \tag{6}$$

where the nudging coefficient is $\alpha_{pv} = 1/T_{pv}$ with T_{pv} denoting the nudging time-scale. Thus the algorithm is univariate: the model geopotential and divergence fields must adjust internally to the imposed PV data, without explicit external forcing. The temporal discretization of eq. (6) takes the following form:

$$PV^{n+1} + \frac{\beta\tau_1'}{\Phi_*^n}\left[\frac{1}{a\cos\phi}\frac{\partial\psi}{\partial\lambda} + \frac{1}{a}\frac{\partial\chi}{\partial\phi}\right]^{n+1}$$
$$= r_1 + \frac{1}{2}(\alpha_{pv}\Delta t)(PV_{data}^{n+1} + PV_{data}^n - \hat{PV}^{n+1} - PV^n), \tag{7}$$

where $\hat{PV}^{n+1}$ is chosen to be PV^{n+1} if an implicit scheme is used and $2.0 \times PV^n - PV^{n-1}$ if an explicit scheme is used. For the implicit scheme, the multigrid solver of the PV-D model (Ruge et al.

1996) needs to be modified. The explicit scheme, adopted here, can be shown to be linearly stable as long as $\alpha_{pv}\Delta t < 1$. We have not found a noticeable difference in the numerical results between these two approaches.

A divergence damping term has been included in the formulation as an option after observing spurious growth of the longest waves, especially in the divergence field, when PV data are inserted with a large nudging coefficient. The SLSI-discretized divergence equation in this case is

$$\left(\nabla^2\chi + \tau_1\left[\nabla^2\Phi - f\nabla^2\psi\right.\right.$$
$$\left.\left.+ \beta\left(\frac{1}{a\cos\phi}\frac{\partial\chi}{\partial\lambda} - \frac{1}{a}\frac{\partial\psi}{\partial\phi}\right)\right]\right)^{n+1}$$
$$= r_2 - \tau_1\nabla^2\Phi_s - \mu_D\Delta t(\nabla^2\chi)^n, \qquad (8)$$

where $\mu_D = 1/T_{div}$ is the divergence damping coefficient with a time-scale T_{div}. Note that the divergence damping used in this study differs from the usual one (Talagrand 1972; Haltiner and Williams 1980; Bates et al. 1993), which is applied on the second-order horizontal derivatives of divergence with the intention of damping out the undesired small-scale gravity waves. The damping
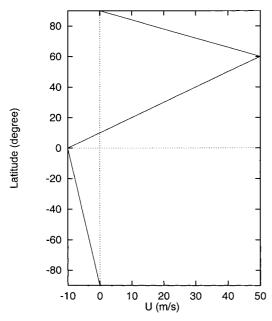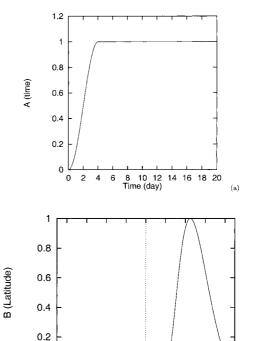

Fig. 2. Functions used to specify the orographic forcing. (a) $A(t)$; (b) $B(\phi)$.

term used here is proportional to the divergence field itself because the spurious modes that arise when α_{pv} is large, as we will see, are on large spatial scales.

3. Numerical experiments

3.1. Experimental design

Experiments of identical-twin type are carried out, in which one model integration, the control, provides the "data". These data are then inserted into model integrations which start from different initial conditions. We refer to these integrations as "assimilation runs".

A polar vortex erosion case, similar to that used by Bates and Li (1996) for a comparison study of numerical simulations, is used here. Despite its simplicity, such a simulation captures basic features (e.g., Rossby wave breaking) of a major flow
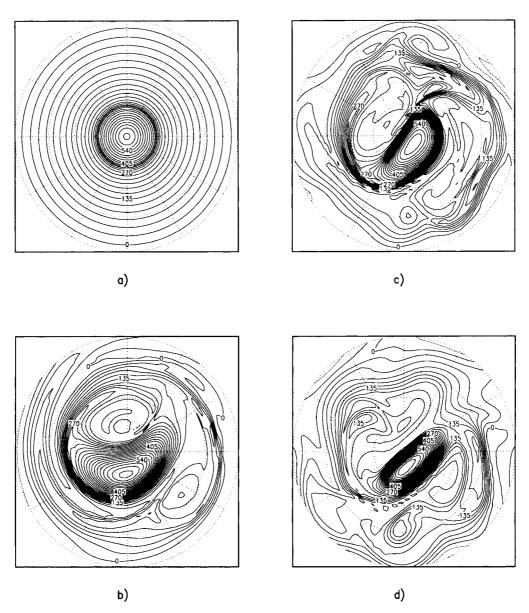

Fig. 1. Initial zonal wind (m/s).

a)   c)

b)   d)

Fig. 3. PV evolution of the control, i.e., the "data", plotted in a polar stereographic projection of the Northern Hemisphere. (a) Day 0; (b) Day 20; (c) Day 30; (d) Day 50. Contour interval $= 22.5$ ($\times 10^{-11}$ m^{-2} s).

pattern for the winter polar stratosphere. Since the work of Juckes and McIntyre (1987), many others have conducted model simulations using various derivatives of this case for dynamical investigations (e.g., Mariotti et al. 1994; Norton 1994; Polvani et al. 1995).

For the control, the model is initialized with a zonal flow displayed in Fig. 1. The initial geopotential is derived from the gradient wind balance. The model is then integrated for 50 days with the following orographic forcing:

$$h(\lambda, \phi, t) = -H_s A(t) B(\phi) \sin \lambda, \qquad (9)$$

with h denoting the height of the orography. We

choose $H_s = 0.72$ km and assign $A(t)$ *and* $B(\phi)$ the form shown in Figs. 2(a) and (b). The evolution of the PV field is displayed in Figs. 3(a)–(d).

For the assimilation runs, the model starts from Day 20 but is initialized with the model fields of Day 10 of the control. Thus the model initial condition bears an error of a magnitude characterized by two states with a 10-day phase shift. PV data are inserted at every gridpoint and every time step. In all the experiments a horizontal resolution of 128×65 points on a regular latitude/longitude grid is used with a timestep of 30 minutes.

3.2. PV data insertion

Fig. 4a displays the global RMS error of the PV field for various nudging time-scales. The curve labeled "Model", for which no data are inserted, is included for reference. All the assimilation runs show an initial decrease of the RMS PV error. The shorter the nudging time-scale, the faster the initial rate. With $T_{pv} = 24$, 48 or 96 h, the RMS errors decrease nearly monotonically, but for $T_{pv} = 6$ h and $T_{pv} = 12$ h, they start to increasse at later times. In fact, for the divergence field (Fig. 4b), we see that the RMS errors increase immediately and rapidly for the two cases $T_{pv} = 6$ h and $T_{pv} = 12$ h. The RMS error of geopotential height (Fig. 4c) starts to increase at a time later than the divergence but earlier than PV for these two cases. For $T_{pv} = 48$ h and $T_{pv} = 96$ h, the RMS errors for all three fields are quite small at Day 50: complete PV data alone suffice to drive all fields to the evolving control state.

We now examine more closely two assimilation runs, a good one ($T_{pv} = 48$ h) and a bad one ($T_{pv} = 6$ h). Figs. 5a–c display the divergence fields at Day 30 for the control, the assimilation run with $T_{pv} = 48$ h and that with $T_{pv} = 6$ h, respectively. The assimilation run with $T_{pv} = 48$ h (Fig. 5b) represents a success — the divergence field arising from PV data assimilation is almost identical to the control. Fig. 6 displays the RMS errors of PV, geopotential, divergence and vorticity for this run, normalized by their respective initial values. The errors of PV, geopotential and vorticity decrease at the same rate, and that of the divergence follows the trend closely. For $T_{pv} = 6$ h (Fig. 5c), the divergence field deviates dramatically from the control.

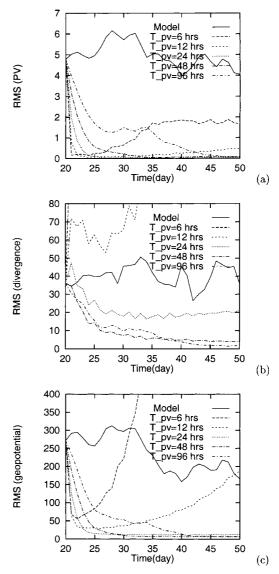*Fig. 4.* RMS error for the PV data insertion experiments with various nudging time-scales. (a) PV ($\times 10^{-10}$ m^{-2} s); (b) Divergence ($\times 10^{-8}$ s^{-1}); (c) Geopotential (m).

There is evidently strong spurious growth of the longest waves.

3.3. PV data insertion with divergence damping

As we have seen, the divergence field takes the lead in the spurious growth associated with the
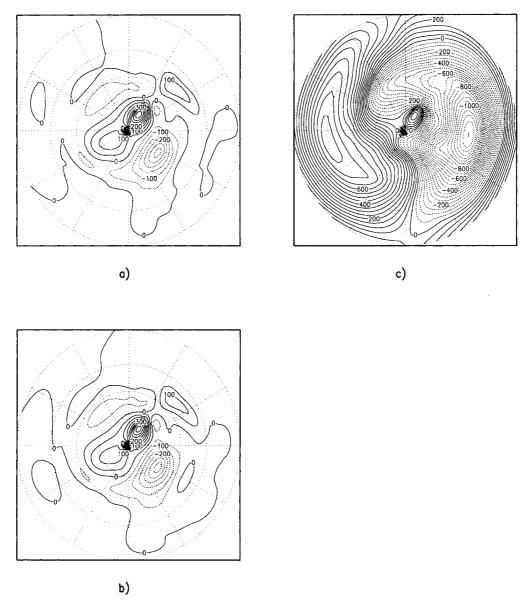
a)

c)

b)

Fig. 5. Divergence at Day 30 for (a) the control run; (b) the assimilation run with $T_{pv} = 48$ h; (c) the assimilation run with $T_{pv} = 6$ h. Contour interval $= 50.0$ ($\times 10^{-8}$ s^{-1}).

longest waves when $T_{pv} = 6$ or 12 h. Since the initial decrease in the RMS PV and geopotential errors (Figs. 4a, c) is nonetheless most rapid for these two cases, a desirable feature, we examine the possibility of including a divergence damping term. For $T_{pv} = 6$ h and $T_{div} = 6$, 12 and 24 h, respectively, the resulting divergence fields at

Day 30 are shown in Figs. 7a–c. Indeed, the spurious long-wave instability is now suppressed, and in all three cases the divergence field is nearly identical to the control (Fig. 5a). Further numerical experiments (results not presented) show that the long-wave instability reappears when T_{div} increases to 48 h, and that the critical

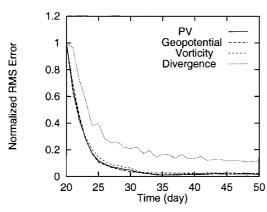*Fig. 6.* Normalized RMS error for the case of PV data insertion with $T_{pv} = 48$ h.

T_{div} is proportional to the T_{pv} being used.

4. Linear analysis

The numerical results we have seen so far point our attention to the potential power of assimilating PV data. We now conduct a simple linear analysis in order to understand better these results. The shallow water equations will be linearized about a state of rest, with constant geopotential height $\bar{\Phi}$.

Assuming no variation in the y-direction, the potential velocity perturbation PV' and divergence perturbation δ' are, respectively,

$$PV' = \frac{1}{\bar{\Phi}} \frac{\partial v'}{\partial x} - \frac{f}{\bar{\Phi}^2} \Phi', \qquad (10)$$

$$\delta' = \frac{\partial u'}{\partial x}, \qquad (11)$$

and the linearized shallow water equations can be written as:

$$\frac{\partial PV'}{\partial t} + \frac{\beta}{\bar{\Phi}} v' = 0, \qquad (12)$$

$$\frac{\partial \delta'}{\partial t} + \left[\frac{\partial^2 \Phi'}{\partial x^2} - \frac{f^2}{\bar{\Phi}} \Phi' - (\bar{\Phi}f)PV' + \beta u' \right] = 0, \qquad (13)$$

$$\frac{\partial \Phi'}{\partial t} + \bar{\Phi}\delta' = 0, \qquad (14)$$

where Φ' is the geopotential perturbation. Now

letting

$$\begin{pmatrix} PV' \\ \delta' \\ \Phi' \end{pmatrix} = \begin{pmatrix} \hat{PV} \\ \hat{\delta} \\ \hat{\Phi} \end{pmatrix} e^{ikx}, \qquad (15)$$

we then have

$$\frac{du}{dt} + Au = 0, \qquad (16)$$

where

$$u \equiv \begin{pmatrix} \hat{PV} \\ \hat{\delta} \\ \hat{\Phi} \end{pmatrix} \qquad (17)$$

and

$$A \equiv \begin{bmatrix} -i\omega_R & 0 & -i\dfrac{\omega_R \omega_I}{\bar{\Phi}^2} \\ -\omega_I \bar{\Phi} & -i\omega_R & -\dfrac{1}{\bar{\Phi}}(\omega_G^2 + \omega_I^2) \\ 0 & \bar{\Phi} & 0 \end{bmatrix}. \qquad (18)$$

Here $\omega_G = (k^2 \bar{\Phi})^{1/2}$, $\omega_I = f$ and $\omega_R = \beta/k$ are, respectively, the frequencies of pure gravity waves, inertia waves and Rossby waves. We define f and β at $45°$ N, and take $\bar{\Phi} = 10$ km.

With initial condition $u(t_0) = u_0$, the solution of the free system (16) is

$$u(t) = e^{-A(t-t_0)}u_0. \qquad (19)$$

If A has eigendecomposition

$$A = P\Lambda P^{-1}, \qquad (20)$$
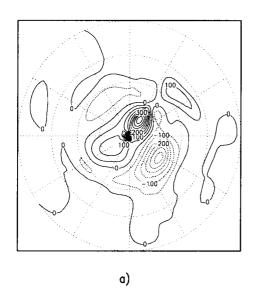
then the solution can be written as
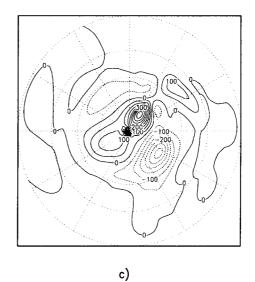
$$u(t) = P\, e^{-\Lambda(t-t_0)}P^{-1}u_0. \qquad (21)$$

Denoting the eigenvalues λ of A by $\lambda = -i\omega$, the dispersion relation $\det(A - \lambda I) = 0$ reads:

$$\omega^3 - 2\omega_R\omega^2 - (\omega_G^2 + \omega_I^2 - \omega_R^2)\omega + \omega_G^2\omega_R = 0. \qquad (22)$$

The solution (21) is purely oscillatory, as expected, since the three roots (eigenfrequencies) ω of (22) are all real. These eigenfrequencies are displayed in Fig. 8. They correspond to the westward-propagating slow mode, the westward-propagaging fast mode and the eastward-propagating fast mode.

A forced problem corresponding to data assimilation by nudging can be written in the following
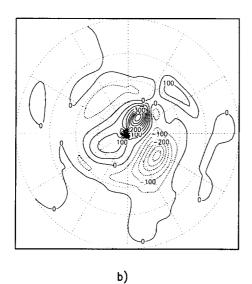
a)

c)

b)

Fig. 7. Divergence at Day 30 for the case of PV data insertion ($T_{pv} = 6$ h) with divergence damping; (a) $T_{div} = 6$ h; (b) $T_{div} = 12$ h; (c) $T_{div} = 24$ h. Contour interval $= 50.0$ ($\times 10^{-8}$ s^{-1}).

matrix form:

$$\frac{dv}{dt} + Av = B(u - v), \quad v(t_0) = v_0, \tag{23}$$

with B independent of t, where u is the solution of the free (unforced) problem above, toward which it is attempted to drive the state v. The vector Bu corresponds to the data to be assimilated. The solution of the forced problem is

$$v(t) = e^{-(A+B)(t-t_0)}(v_0 - u_0) + u(t). \tag{24}$$

Thus the "error field" $e(t) \equiv v(t) - u(t)$ is given by

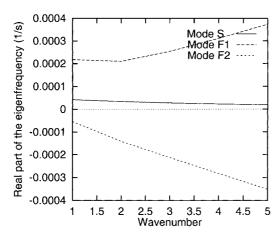$$e(t) = e^{-(A+B)(t-t_0)}e_0, \tag{25}$$

Fig. 8. The eigenfrequencies $\omega = \omega(k)$ of the three free modes as a function of zonal wavenumber. Mode S: westward-propagating slow mode, Mode F1: westward-propagating fast mode, Mode F2: eastward-propagating fast mode.

where $e_0 \equiv e(t_0) = v_0 - u_0$.

Remark 1. If $B = \alpha I$, meaning that complete data are available for all three fields (PV, δ, Φ) and are inserted with identical nudging parameter α, then

$$e(t) = e^{-(A + \alpha I)(t - t_0)}e_0. \tag{26}$$

Since

$$A + \alpha I = P(\Lambda + \alpha I)P^{-1}, \tag{27}$$

we have

$$e(t) = e^{-\alpha(t - t_0)}Pe^{-A(t - t_0)}P^{-1}e_0, \tag{28}$$

so the error field does indeed decay to zero as $e^{-\alpha(t - t_0)}$, which is a scalar.

Remark 2. In general, A and B will not commute, so that

$$e^{-(A + B)(t - t_0)} \neq e^{-A(t - t_0)} e^{-B(t - t_0)}. \tag{29}$$

Therefore even if the eigenvalues of B are all positive, it is not necessarily true that the error field will decay.

Remark 3. Suppose that e_0 is a slow eigenmode for A. This would be the case for instance if u_0 is a slow eigenmode for A and $v_0 = 0$, since then $e_0 = -u_0$. But e_0 is not necessarily a slow eigenmode for $A + B$, hence the error field evolution in (25) may not be slow.

Now let's consider specifically the forced prob-

lem with PV data insertion, for which

$$B = \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{30}$$

where $\alpha = \alpha_{pv} = 1/T_{pv}$. Again, denote the eigenvalues of $A + B$ by $\lambda = -i\omega$. Then from $\det(A + B - \lambda I) = 0$, we obtain the following cubic equation for ω:

$$\omega^3 - (2\omega_R + i\alpha)\omega^2 - (\omega_G^2 + \omega_I^2 - \omega_R^2 - i\alpha\omega_R)\omega$$
$$+ \omega_G^2\omega_R + i\alpha(\omega_G^2 + \omega_I^2) = 0, \tag{31}$$

whose roots will be complex when $\alpha \neq 0$.

Denoting the diagonal matrix of eigenvalues of $A + B$ by $\Lambda = -i\Omega$ and the corresponding matrix of eigenvectors by P, it follows from (25) that the error field is

$$e(t) = P\,e^{i\Omega(t - t_0)}P^{-1}e_0. \tag{32}$$

The real parts of the eigenfrequencies ω give the purely oscillatory parts of the error field. The error field decays to zero as $t \to \infty$ for every e_0 if and only if

$$\text{Im } \omega > 0 \tag{33}$$

for every root of (31) and every wavenumber k.

The real parts of the eigenfrequencies turn out to be essentially unaffected by the PV data insertion: plots (not shown) of the real parts for T_{pv} as small as 6 h are nearly indistinguishable from Fig. 8. Thus the error system has well-defined slow and fast modes. However, the imaginary parts of the eigenfrequencies are always nonzero and the modes are not neutrally stable, unlike those of the free system. In Fig. 9, we display the imaginary parts of the eigenfrequencies for the PV data insertion cases with $T_{pv} = 6$ h (panel a) and $T_{pv} = 48$ h (panel b). These represent the exponential growth or decay rates of the corresponding modes.

From examination of Fig. 9, we reach the following conclusions.

(a) The component of the error field corresponding to the slow modes always decays. The decay rate, which is proportional to the nudging coefficient, is nearly uniform over the wave spectrum, meaning that all spatial scales of the error field are eliminated equally effectively for the slow modes.

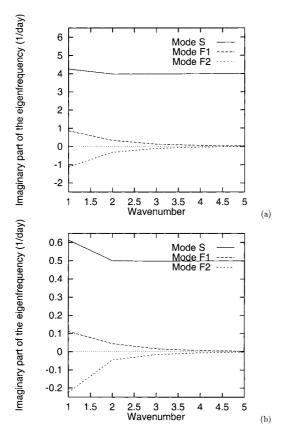(b) The component of the error field corresponding to the westward-propagating fast modes

Fig. 9. As in Fig. 8, but for the imaginary part of the eigenfrequencies of the system with PV data insertion: (a) $T_{pv} = 6$ h; (b) $T_{pv} = 48$ h. Note that the vertical axes of the two plots are scaled differently.

always decays, though much more slowly than for the slow modes. The rate of decay decreases with increasing wavenumber.

(c) The component of the error field corresponding to the eastward-propagating fast modes grows. The growth rate decreases with increasing wavenumber, and is most pronounced for wavenumber 1 when the nudging time-scale is small.

In order to check whether the long-wave instability observed in the numerical experiments of Section 3 with small nudging time-scale does in fact correspond to eastward-propagating fast modes, the hourly divergence field for the case with $T_{pv} = 6$ h was examined. It was found that the propagation is indeed eastward. The linear analysis, although suggestive, tends to over-estimate the magnitude of the error growth as

compared with the numerical experiments, probably due to the model dynamics being nonlinear and the numerics being not completely without intrinsic time-smoothing. In fact, the numerical experiments with $T_{pv} = 24$ h, $T_{pv} = 48$ h and $T_{pv} = 96$ h displayed no instability.

The linear analysis confirms that large-scale instability can be expected for simple univariate nudging toward PV data, especially when the nudging coefficient is large. On the other hand, the linear analysis suggests that any small-scale instability should be modest, and perhaps controllable by intrinsic spatial smoothing of a numerical model. Small-scale noise was not observed in any of our numerical experiments. We conclude that PV data assimilation by univariate nudging does not tend to generate small-scale noise as often encountered in other data assimilation problems.

5. Assimilating TOMS total ozone data

In this section, we describe as assimilation experiment in which total ozone data from the Nimbus-7 TOMS instrument were used as proxy for PV observations. PV is known to be well-correlated with ozone mixing ratio at individual levels near the tropopause in the extratropics. In a study of pattern correlation coefficients, Allaart et al. (1993, henceforth A93) showed furthermore that total ozone, C_{TOVS}, as calculated from the TOVS ozone channel and assimilated with a barotropic transport model driven by ECMF W 6-h forecast winds, can be reasonably well reproduced by a PV-based estimate $\tilde{C}$ defined as

$$\tilde{C} = C_0 + \lambda \int_{p_t}^{p_s} PV(p)\, dp \qquad (34)$$

in the northern hemisphere. In eq. (34), the total column ozone is determined by a rescaled partial PV column integral plus a background term C_0 to within an error $\varepsilon \equiv C_{TOVS} - \tilde{C}$ that is of the same magnitude as the TOVS total ozone retrieval error. C_0 and λ are slowly-varying functions of the seasons, while the error ε is a random variable with mean zero. The fact that ε is comparable to the instrument retrieval error is taken as evidence that the ozone mixing ratio and the PV are proportional in the height region that contributes to the variability in total ozone. After subtraction of the quasi-constant offset C_0, the total ozone

must therefore be proportional to the pressure-weighted average PV of that region. A93 empirically found the highest correlation between $\tilde{C}$ and C_{TOVS} when p_t was set to 50 hPa, and p_s to 400 hPa.

Riishøjgaard et al. (1997) examined the validity of eq. (34) in further detail using total ozone from TOMS and GEOS-DAS (DAO 1996) analysis PV. They found pattern correlation coefficients very similar to those of A93, however, based on a time series analysis they showed that the agreement in the tropics is largely fortuitous, and that only in the middle and sub-polar latitudes does eq. (34) reflect a truly dynamical relationship. They also found marked differences between the northern and southern hemispheres, with the northern hemisphere winter and spring showing by far the best correlation.

With these caveats, we take the combined results of A93 and Riishøjgaard et al. (op. cit.) as justification for assimilating actual total ozone measurements from TOMS globally as proxy observations of vertically averaged PV. From the point of view of the assimilation system the interesting aspect of this experiment is not only the degree of meteorological realism that can be achieved. This will in any case be limited by the fact that the shallow water equations are used to constrain the system. The key point is the question of whether a PV assimilation system can actually digest such highly imperfect observations and generate internally consistent geopotential and PV fields from them.
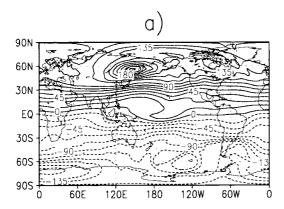
The Nimbus-7 platform was flying in a polar, nearly sun-synchronous orbit during the assimilation period, and observations were taken near local noon at a resolution of roughly 50 by 50 km at the nadir point. The TOMS instrument operates in a cross-track scanning mode, and each point on the globe except for the polar night region is observed within a 24-h period. As in the numerical experiments discussed in Section 3, the proxy-observational data are inserted at each model time step. For the TOMS data, however, only a limited part of the globe is covered within a time step, while the rest is unobserved. All data obtained within a ± 15 min window surrounding a given model time are taken as valid at that time, and for the cell surrounding each model grid point a simple mean of all valid measurements within the cell is defined to be the observation. The model is run without orography, from an initial state of rest.
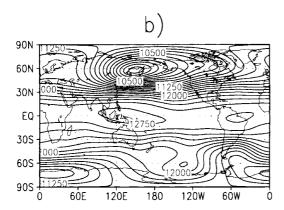
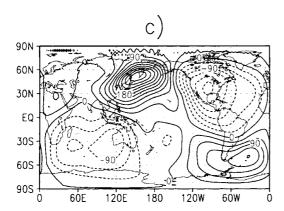In order to transform the total ozone observa-

tions into proxy PV observations, the background contribution C_0 in eq. (34) was first subtracted from them. The difference $C_{\text{TOMS}} - C_0$ was then multiplied by an empirical scaling factor so that the range of the observations corresponds to the expected range of model PV. Note that although the background contribution C_0 was extracted from the statistical study of A93, this approach could not be used for the scaling factor (λ in eq. (34)), since the actual values of shallow water PV are different from those of vertically-averaged atmospheric PV. Also note that the correlation between PV and ozone bears some seasonal variation (Allaart et al. 1993) and geographical dependence (Riishøjgaard et al. 1997). These factors have not been taken into account in this experiment, which takes the parameters λ and C_0 to be constant.

In order to separate the temporal range of influence of an observation from the choice of timestep used in the model integration, we insert any given observation gradually during a 6-h window. At a given model time t_m, observations outside the interval $(t_m - 3 \text{ h}; t_m + 3 \text{ h})$ are not used. Inside that interval the observations being inserted are weighted with a cosine curve (the positive half-period) that has its maximum at t_m. The nudging time-scale is 24 h.

We now show results from an assimilation run beginning on 5 February 1992, 00Z. After an initial spin-up period, the flow develops into a mature pattern and subsequently keeps evolving with the inserted total proxy PV data. PV, geopotential, and divergence after 10 days of assimilation (15 February 1992, 00Z) are presented in Figs. 10a–c. For comparison, aggregate TOMS total ozone data resampled to the model resolution for the 24 h from 15 February 00Z through 16 February 00Z are shown in Fig. 11. The model PV matches the observed total ozone field relatively well on the largest scales, especially in the northern (winter) hemisphere. The high total ozone region extending along the 60° N parallel from eastern Siberia across the Pacific is well-captured. Also the ridge/trough pattern in the sub-tropical central Pacific region seen in the TOMS data is found in the PV field. The geopotential field (Fig. 10b) seems to be consistent with the PV field, and it could conceivably be a low-order approximation to an actual geopotential height field at the mean tropopause level. We have not attempted

to validate the fields of Fig. 10 against meteorological analysis fields. As expected from the results of Riishøjgaard et al. (1997), the link between the dynamical fields of Fig. 10 and the ozone field of Fig. 11 is strongest in the northern hemisphere.

It is worth emphasizing that even though we are inserting small patches of proxy PV data that are very unevenly distributed in time and space, the flow fields that are reconstructed from these data are actually mainly large-scale, including the divergence field (Fig. 10c). Although no divergence damping was used for this experiment, the absolute values of the divergence remain at moderate levels, with no tendency towards growth as experienced in some of the experiments discussed in Section 3. Contrary to what might have been expected, the highly inhomogeneous data distribution does not give rise to small-scale noise in the fields.

The correlation coefficient between the model PV and the proxy PV data (scaled total ozone) is shown in Fig. 12 as a function of time. The value increases from zero to about 0.6 in approximately one day and then fluctuates throughout the 10-day period. The correlation coefficient is larger by about 8% if it is only for the northern hemisphere extratropics (not shown here). Fig. 12 demonstrates that convergence toward the proxy PV field is rapid but incomplete. Planetary-scale features in Fig. 10a and Fig. 11 match well, for instance, but regional-scale features do not.

The short-term variability in total ozone is known to be dynamically controlled also on the regional scales. These scales, however, are conspicuously absent from the plots in Fig. 10, that are totally dominated by planetary-scale features. One possibility is that the shallow water dynamics is different from the real dynamics of the lower stratosphere that dominates the total ozone signal. Another possibility would be that the simple nudging scheme cannot exploit the tracer property of the data, which carries information about the advective velocities on the regional scales. As we shall argue in the following, a simple explanation can be that the absence of a regional-scale signal

Fig. 10. Model fields at Day 10 (15 February 1992, 00Z) from the assimilation of TOMS total ozone data. (a) PV (contour interval $= 15.0 \times 10^{-11}$ m^{-2} s); (b) geopotential height (contour interval $= 150.0$ m); (c) divergence (contour interval $= 30.0 \times 10^{-8}$ s^{-1}).
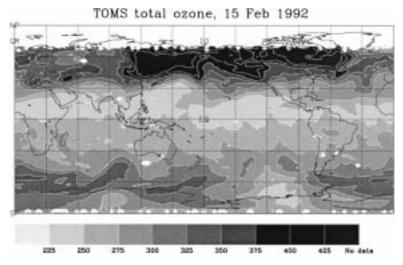
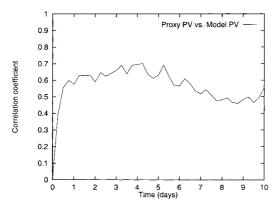*Fig. 11.* TOMS total ozone on 15 February 1992, in Dobson units.

Fig. 12. Correlation coefficient between proxy PV, i.e., scaled total ozone data, and model PV. The proxy PV data are assembled over 24-h intervals centered at the times of correlation calculation. This is to ensure a quasi-global coverage of proxy PV data for each correlation calculation.

is due to the univariate way in which the PV insertion is done in the simple nudging scheme that we use.

Shallow water PV is the ratio between absolute vorticity and geopotential. Thus, the PV can increase either through an increase in absolute vorticity, through a decrease in geopotential, or through a combination of changes to the two. Correspondingly, total ozone can change because of vertical motion leading to a change in tropopause height, through horizontal advection of air

from a latitude with a climatologically different ozone content, or, as is generally the case, through a combination of these two effects (see e.g. Riishøjgaard and Källén, 1997). A decrease in tropopause height will be accompanied by convergence in the stratosphere and divergence in the troposphere, and vice versa. The shallow water analog of this is a change in geopotential due to divergence. Horizontal intrusions of airmasses from different latitudes, on the other hand, tend to be of larger scale and are often barotropic in structure. A barotropic degenerate of the shallow water system is vorticity advection. In this way we can directly link each of the two components of the PV to a specific process affecting total ozone.

By applying the PV forcing to eq. (1), but not to eq. (4), an asymmetry is introduced between the ways in which vorticity and geopotential are updated. The vorticity is left free to respond to the data insertion, whereas the geopotential remains partly locked by the lack of a forcing term on the right-hand side of eq. (4)*. The data insertion will therefore tend to feed the data into

---

* Applying forcing also to the continuity equation would require an assumption about the relationship between PV and Φ, as well as solving the actual equation for that relation. However, by forcing only the PV equation and not the continuity equation we are implicitly assuming that the change in PV is caused mainly by a change in the ageostrophic vorticity component. This assumption is questionable if the changes are large.

the vorticity component rather than into the geopotential component of the PV. The resulting model fields may therefore capture changes associated with horizontal advection better than those caused mainly by vertical motion. Because of the characteristic scales of these two kinds of changes in the total ozone fields, this would bias the model fields towards the larger scales.

6. Conclusions

The correlation between a tracer field such as ozone and the PV field indicates the potential availability of proxy PV data. It has been known for some time that PV is a dynamically significant quantity through which many other quantities can be derived if certain balance conditions are assumed. In other words, PV can be regarded as a "master" quantity for atmospheric dynamics. From the data assimilation point of view, PV contains information concerning both the mass field and the wind field, and thus has much to offer. But this may also be a problem. If PV data are inserted univariately, with no externally-imposed balance constraint, the model is required to adjust quickly enough in order to partition the inserted PV information between the wind field and the mass field in a dynamically consistent manner. How would a model behave in such a situation?

We have made an initial attempt to address this issue. The shallow water PV-D model and a simple univariate nudging algorithm were used. A model-simulated polar vortex was used as a testbed for identical-twin experiments. A more realistic experiment was also conducted, in which actual TOMS total ozone data were treated as proxy PV data. A linear analysis, which agrees reasonably well with the numerical experiments, gave us additional insight into the experimental results.

We may summarize our results as follows: (1) The shallow water model recovers the full dynamical fields from complete PV data alone if an appropriate nudging coefficient is chosen; (2) If the nudging time-scale is too small, there is spurious growth of large-scale features, which corresponds, as revealed by the linear analysis and verified by the numerical experiments, to the longest eastward-propagating fast modes of the system. This instability can be controlled by including a divergence damping term in the divergence equation; (3) PV data insertion does not tend to

generate small-scale noise; (4) The linear analysis shows that there is little scale-dependence of convergence for the slow mode. This suggests that PV data insertion leads to recovery of the small scales as efficiently as for the large scales; (5) The model absorbs the TOMS total ozone observations as proxy PV data very well. The resulting geopotential and divergence fields are dynamically consistent and mainly of large scale. The absence of regional-scale features in the TOMS data assimilation experiment may be due to incompatibility between total ozone observations and shallow water dynamics, or it may be an artifact of the nudging scheme.

In this study, we have concentrated purposely on a particular data assimilation strategy, namely on the use of chemical tracer observations as proxy PV data rather than exploiting the tracer property itself of the data. The nudging methodology we have used is also primitive in at least two respects: it accounts neither for the correlation between errors in different fields, nor for the spatial correlation of the error in a single field. Thus, while this study suggests that tracer observations, as proxy PV data, may offer considerable information in a data assimilation system, we expect that use of a modern multivariate statistical data assimilation scheme would more fully realize the potential of this strategy by providing better dynamical consistency. Our study is also limited by the use of a shallow water model for the dynamics. Much more effort will be needed to extend this work to three dimensions.

REFERENCES

Allaart, M. A. F., Kelder, H. and Heijboer, L. C. 1993. On the relation between ozone and potential vorticity. *Geophys. Res. Lett.* **20**, 811–814.

Allen, J. S. 1993. Iterated geostrophic intermediate models. *J. Phys. Oceanogr.* **23**, 2447–2461.

Anthes, R. A. 1974. Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.* **31**, 702–719.

Bates, J. R., Moorthi, S. and Higgins, R. W. 1993. A global multilevel atmospheric model using a vector semi-Lagrangian finite-difference scheme. Part 1: Adiabatic formulation. *Mon. Wea. Rev.* **121**, 244–263.

Bates, J. R., Li, Y., Brandt, A., McCormick, S. F. and Ruge, J. 1995. A global shallow-water numerical model based on the semi-Lagrangian advection of potential vorticity. *Q. J. R. Meteorol. Soc.* **121**, 1981–2005.

Bates, J. R. and Li, Y. 1996. Simulation of stratospheric vortex erosion using three different global shallow water numerical models. In: *Numerical methods in atmospheric and oceanic modelling*. The André Robert Memorial volume (Companion Volume to Atmosphere-Ocean), (ed. C. Lin, R. Laprise, H. Ritchie), NRC Research Press, 624pp.

Charney, J., Halem, M. and Jastrow, R. 1969. Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.* **26**, 1160–1163.

Daley, R. 1995. Estimating the wind field from chemical constituent observations: Experiments with a one-dimensional extended Kalman filter. *Mon. Wea. Rev.* **123**, 181–198.

Daley, R. 1996. Recovery of the one and two dimensional windfields from chemical constituent observations using the constituent transport equation and an extended Kalman filter. *Meteorol. Atmos. Phys.* **60**, 119–136.

Danielsen, E. F. 1968. Stratospheric-tropospheric exchange based on radioactivity, ozone and potential vorticity. *J. Atmos. Sci.* **25**, 502–518.

Danielsen, E. F., Hipskind, R. S., Gaines, S. E., Sachse, G. W., Gregory, G. L. and Hill, G. F. 1987. Three-dimensional analysis of potential vorticity associated with tropopause folds and observed variations of ozone and carbon monoxide. *J. Geophys. Res.* **92**, 2103–2111.

DAO. 1996. *Algorithm theoretical basis document, version 1.01*. Data Assimilation Office, NASA/Goddard Space Flight Center (also available at: http://dao.gsfc.nasa.gov/subpages/atbd.html).

Davies, H. C. and Turner, R. E. 1977. Updating prediction models by dynamical relaxation: an examination of the technique. *Q. J. R. Meteorol. Soc.* **103**, 225–245.

Douglass, A. R., Rood, R. B., Stolarski, R. S., Schoeberl, M. R., Proffitt, M. H., Margitan, J. J., Loewenstein, M., Podolske, J. R. and Strahan, S. E. 1990. Global three-dimensional constituent fields derived from profile data. *Geophys. Res. Lett.* **17**, 525–528.

Haltiner, G. J. and Williams, R. T. 1980. *Numerical prediction and dynamic meteorology*. John Wiley and Sons, 477p.

Hoskins, B. J., McIntyre, M. E. and Robertson, A. W. 1985. On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.* **111**, 877–946.

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

Kuo, Y.-H., Guo, Y.-R. and Westwater, E. R. 1993. Assimilation of precipitable water measurements into a mesoscale numerical model. *Mon. Wea. Rev.* **121**, 1215–1238.

Lait, L. R., Schoeberl, M. R., Newman, P. A., Proffitt, M. H., Loewenstein, M., Podolske, J. R., Strahan, S. E., Chan, K. R., Gary, B., Margitan, J. J., Browell, E., McCormick, M. P. and Torres, A. 1990. Global three-dimensional constituent fields derived from profile data. *Geophys. Res. Lett.* **17**, 525–529.

Lary, D. J., Chipperfield, M. P., Pyle, J. A., Norton, W. A. and Riishøjgaard, L. P. 1995. Three-dimensional tracer initialization and general diagnostics using equivalent PV latitude-potential-temperature coordinates. *Q. J. R. Meteorol. Soc.* **121**, 187–210.

Li, Y. and Bates, J. R. 1996. A study of the behaviour of semi-Lagrangian models in the presence of orography. *Q. J. R. Meteorol. Soc.* **122**, 1675–1700.

Mariotti, A., Legras, B. and Dritschel, D. G. 1994. Vortex stripping and the erosion of coherent structures in two-dimensional flows. *Phys. Fluids* **6**, 3954–3962.

Norton, W. 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.

Polvani, L. M., Waugh, D. W. and Plumb, R. A. 1995. On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.* **52**, 1288–1309.

Riishøjgaard, L. P. 1996. On four-dimensional variational assimilation of ozone data in weather prediction models. *Q. J. R. Meteorol. Soc.* **122**, 1545–1571.

Riishøjgaard, L. P. and Källén, E. 1997. On the correlation between ozone and potential vorticity for large-scale Rossby waves. *J. Geoph. Res.* **107**, 8793–8804.

Riishøjgaard, L. P., Lou, G. P., Menard, R. and Cohn, S. E. 1997. A study of the relationship between GEOS-DAS PV and TOMS total ozone from February 1992. *Proceedings of the XVIII Quadrennial Ozone Symposium*, in press.

Rowlands, N., Buttner, G., Raab, A., Shepherd, G., Gault, W., Cann, M., Dobbie, S., Sargoytchev, S., Ward, W., Mani, R., Rochon, Y. and Tarasick, D. 1996. A satellite instrument to measure stratospheric winds. In *Optical spectroscopic techniques and instrumentation for atmospheric and space research II*. SPIE No. 2830.

Ruge, J., Li, Y., McCormick, S. F., Brandt, A. and Bates,

J. R. 1996. A nonlinear multigrid solver for a semi-Lagrangian potential vorticity based shallow water model on the sphere. *SIAM J. Sci. Comput.*, in press.

Talagrand, O. 1972. On the damping of high-frequency motions in four-dimensional assimilation of meteorological data. *J. Atmos. Sci.* **29**, 1571–1574.

Tarasick, D. W., Gault, W. A., Shepherd, G. G., Ward, W. E. and Rowlands, N. 1996. Imaging Michelson interferometers for studies of ozone transport. *Proc. Quad. Ozone Symp.*, l'Aquila, Italy.

Vallis, G. K. 1996. Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Q. J. R. Meteorol. Soc.* **122**, 291–322.

Warn, T., Bokhove, O., Shepherd, T. G. and Vallis, G. K. 1995. Rossby number expansions, slaving principles, and balance dynamics. *Q. J. R. Meteorol. Soc.* **121**, 723–739.