Towards a PV-$\theta$ view of the general circulation

By BRIAN J. HOSKINS, Department of Meteorology, University of Reading, 2 Earley Gate, Whiteknights, PO Box 239, Reading RG6 2AU, UK

(Manuscript received 18 July 1990; in final form 28 February 1991)

ABSTRACT

In recent years, there has been a resurgence of interest in using isentropic coordinates and Rossby-Ertel potential vorticity (PV) for diagnosing the behaviour of middle latitude synoptic systems. Such a PV-$\theta$ analysis may also prove important in providing insight into the global circulation of the atmosphere. Apart from the isentropic diagnostic of D. Johnson and collaborators, some quasi-geostrophic studies and recent studies of stratospheric behaviour, there has been little work in this area and our present understanding is very limited. The object of the present paper is to stimulate such studies by presenting some initial results from continuing research. A three-fold division of the atmosphere is discussed. The "Overworld" is the region encompassed by isentropic surfaces that are everywhere above the tropopause. In the "Middleworld", the region with isentropes crossing the tropopause but not striking the Earth's surface, the isentropic zonal and time mean of PV exhibits interesting regions of enhanced and diminished gradients. The isentropic transient eddy advection of PV exhibits a dipolar distribution about the tropopause, suggestive of PV mixing. The marked PV signature of the Asian summer monsoon on one particular Middleworld isentrope is shown and the mean isentropic advection of PV shows interesting features. For the "Underworld", in which isentropic surfaces intercept the surface of the Earth, a PV-$\theta$ analysis yields a novel constraint linking low-level drag and diabatic heating. This constraint links "westerlies" and "cooling", and "easterlies" and "heating" in some average sense. The result is discussed in terms of the Southern Hemisphere strong surface westerlies and the circulation associated with the Asian summer and winter monsoons.

1. Introduction

This tribute to the wide-ranging scientific, political and personal achievements of Professor Bert Bolin may be usefully introduced by the schematic of the atmosphere below 30 mb shown in Fig. 1. This shows isentropic surfaces every 30 K from 270 K to 390 K, and also the extra-tropical and tropical tropopause. The picture suggests a subdivision of the atmosphere which is an extension of that originally proposed by Sir Napier Shaw (1930). The "Overworld" consists of those isentropic surfaces that everywhere lie above the tropopause. In the "Middleworld" the isentropic surfaces cross the tropopause. The lowest isentropic surface to do this also grazes the Earth's surface in the tropics. (Presumably this occurs by chance). Below this is the "Underworld". The Underworld is analogous to the oceanographer's ventilated thermocline in the sense that air parcels have a direct isentropic route to the surface. Similarly there is a direct isentropic route between the Middleworld troposphere and stratosphere. Haynes and McIntyre (1987), Holton (1990) and Haynes et al. (1990) discuss the indicated circulation of mass between the Middleworld and Underworld. The vertical branches of this circulation may be determined from the upper horizontal branch, which is itself related through the zonal momentum equation to eddy-induced forces. Johnson et al. (1985) have determined the Middleworld and Underworld mean meridional circulation in isentropic coordinates and discussed its maintenance. The most striking result is the extension of the Hadley Cell over each hemisphere.

The focus in this paper is on the Middleworld and Underworld. The framework for the discussion is Rossby-Ertel potential vorticity and potential temperature and their conservation; this is to be contrasted with the quasi-geostrophic potential
vorticity, \(q\). The latter has been used for a number of investigations of the general circulation. For example, Green (1970) used the relation:

\[
\rho [v^* q^*] = \frac{\partial}{\partial y} \left\{ -\rho [u^* v^*] \right\} \\
+ \frac{\partial}{\partial z} \left\{ \rho \frac{\partial q}{\partial \theta} [v^* \theta^*] \right\}, \tag{1}
\]

and mixing length theory for the quasi-conserved properties \(q\) and \(\theta\), to create a parametrisation for the eddy momentum flux convergence. Andrews and McIntyre (1977) introduced the notion of the residual meridional circulation, the Eliassen–Palm flux \(F\), whose \(y\) and \(z\) components are, to a first approximation, the quantities inside the curly brackets in (1), the interpretation of \(F\) as a flux of wave activity, and the effective eddy forcing of the zonal mean momentum by \(\rho [v^* q^*] = \nabla \cdot F\). These ideas were applied to atmospheric data by Edmon et al. (1980) and comparison made with idealised nonlinear baroclinic wave signatures. Holopainen et al. (1981) discussed implications of the 3-D distribution of the transient eddy flux of \(q\), and Lau and Holopanien (1984) took account also of the boundary potential temperature flux and determined the mean potential tendency associated with the transients.

However, there are difficulties in applying quasigeostrophic theory to planetary scale flow, and in particular to domains wide enough to include both the mid-latitudes and the tropics. Furthermore the theory is inaccurate near jet streams with relative vorticities comparable to \(f\), and sloping tropopause transitions in static stability. Recently there has been renewed interest in using full Rossby-Ertel potential vorticity (henceforth denoted PV) to illuminate the behaviour of middle latitude weather systems, including their interaction with low latitudes. Hoskins et al. (1985) gave a review of the associated concepts and their application, and Hoskins (1990) has used those concepts extensively in a review of extra-tropical cyclogenesis.

In this paper, some of the background material on PV is summarised in Section 2. Sections 3 and 4 then give some preliminary results and insights that are emerging in current research at the University of Reading on the general circulation using analyses based on the PV-\(\theta\) framework.

2. A brief review of PV

In a frictionless, adiabatic atmosphere a closed material contour on an isentropic surface remains on that surface and its absolute circulation

\[
C_a = \oint u_a \cdot dl = \int \xi_a \cdot n \, dS \tag{2}
\]

is conserved. For an elemental cylinder of cross-sectional area \(S\) between surfaces \(\theta\) and \(\theta + \delta\theta\),

\[
\frac{C_a}{m} \delta\theta = \frac{1}{\rho} \int \xi_a \cdot n \, \delta h
\]

is also conserved. In the limit this gives material conservation of the PV

\[
P = \frac{1}{\rho} \xi_a \cdot \nabla \theta. \tag{3}
\]

At the level of the hydrostatic approximation,
\( \zeta_n \cdot n = \zeta_{n0} \), Rossby’s “isentropic vorticity”, a quantity resembling the vertical component of absolute vorticity, but with horizontal derivatives evaluated on isentropic surfaces. Then (3) is similarly approximated by

\[ P = \frac{1}{\sigma} \zeta_{n0}, \quad (4) \]

where \( \sigma = -g^{-1} \frac{\partial p}{\partial \theta} \) plays the role of density in isentropic coordinates.

In a frictionless, adiabatic atmosphere PV is conserved by the 2-D motion on an isentropic surface and similarly \( \theta \) is conserved by the 2-D motion on an iso-PV surface. Further if the interior PV and boundary \( \theta \) distributions are known then the balanced motion may be determined by inversion of an elliptic operator.

Returning to the schematic in Fig. 1, the PV = 0 contour lies close to the equator. In the troposphere PV magnitudes generally increase polewards and upwards to values somewhat greater than 1 PVU (1 PVU = \( 10^{-6} \text{m}^2 \text{K s}^{-1} \text{kg}^{-1} \)). Above the extra-tropical tropopause the magnitude is about 4 PVU and increases rapidly polewards and upwards. Thus an iso-PV surface such as PV = 2 PVU marks a quasi-conservative tropopause poleward of about latitude 25°. Danielsen (1984) used the value 1.6 PVU to delineate the tropopause. However, given the large jump in PV at the tropopause the actual value chosen is not critical. The contours of \( \theta \) on such a surface provide a particularly illuminating view of motion in the Middleworld. Hoskins and Berrisford (1988) have used a sequence of such maps to display the upper-level processes involved in an explosive cyclogenesis.

A low-\( \theta \) (high-\( \theta \)) tropopause, as measured by \( \theta \) on PV, is associated, through inversion, with a cyclonic (anticyclonic) circulation and reduced (increased) static stability below. A low-\( \theta \) (high-\( \theta \)) lower boundary is associated with an anticyclonic (cyclonic) circulation and increased (decreased) static stability above. When the tropopause and boundary \( \theta \) anomalies are in the same sense then the associated circulations oppose one another. The extent of the cancellation depends on the height scale of these circulations compared with the depth of the tropopause \( H_T \). Using standard notation, for planetary horizontal length scales \( fL/NH_T \) is large in middle and high latitudes and the cancellation is very large. For example, the polar vortex in Fig. 1 has a low-\( \theta \) tropopause (cyclonic) and low-\( \theta \) lower boundary (anticyclonic). The net result is a much reduced cyclonic circulation than suggested by the tropopause \( \theta \) alone, having a maximum near the tropopause. Equivalent barotropic Rossby waves, having comparable meridional displacements at all levels, also have a large cancellation between interior PV and boundary \( \theta \). The essence of baroclinic instability, as summarised in Hoskins (1990) for example, is an interior PV and boundary \( \theta \) anomaly interaction involving significant cancellation, but also cooperation to produce amplification of the anomalies. Here \( fL/NH_T \sim 1 \). The shortwave cut-off in the Eady and 2-layer model occurs at scales below which such interaction is too weak for an unstable structure to be maintained. The significant cancellation between boundary and interior PV contributions to the large-scale circulation which has recently been highlighted by Holopainen (pers. comm.) means that the PV view of the general circulation may need careful determination.

For a mass of fluid between two isentropic surfaces, \( \theta_1 \) and \( \theta_2 \), (4) gives

\[ \int P \, dm = \int_{\theta_1}^{\theta_2} C_s \, d\theta, \quad (5) \]

where \( C_s = C_s(\theta) \) is the circulation on an isentropic circuit bounding the mass, and \( dm \) is the mass element \( \sigma \, dx \, dy \, d\theta \). (5) states that the mass-weighted potential vorticity integral is equal to the integral of the isentropic circulation around the region. It is therefore unaltered by any frictional or external forces, or by heating processes that act within the region. This is reflected by the PV equation in the flux form given by Haynes and McIntyre (1987):

\[ \frac{\partial}{\partial \theta} \Big|_\phi \sigma P = -\nabla \cdot J, \quad (6) \]

where

\[ J = (u, v, 0) \sigma P + \frac{\partial}{\partial \theta} (v, -u, 0) \]

\[ + (-F_y, F_x, 0), \]

Tellus 43AB (1991), 4
Fig. 2. Schematic of the result of a local diabatic heating in the northern hemisphere. The vertical coordinate is pressure. (a) The heating represented by the cross. The absolute circulation in the initial state is indicated. (b) The result of the heating with the signs of the consequent changes in PV indicated. (c) The convergence and divergence associated with adjustment towards balance. (d) The adjusted state with the absolute circulation decreased above and increased below.

and $F = (F_x, F_y, 0)$ is the horizontal component of the external force per unit mass.

Haynes and McIntyre (1987, 1990) stress the importance of the constraint implied by the appearance of a 2-D, "horizontal" divergence on the right hand side of (6), especially when the behaviour of PV is being compared with that of chemical tracers. Viewed in the isentropic framework, the mass-weighted PV can change only through the convergence of a flux along an isentropic surface. However it should be noted that (4) implies that (6) is a form of the vorticity equation in isentropic coordinates. The qualitative thermodynamic information which is directly available from $P$ has been lost. It could of course by recovered by inversion of $\sigma P$ and use of the appropriate balance condition. An alternative that keeps this direct thermodynamic information is to consider the Lagrangian form

$$\frac{DP}{Dt} = \frac{1}{\rho} \zeta \cdot \nabla \theta + \frac{1}{\rho} \nabla \times F \cdot \nabla \theta. \quad (7)$$

The hydrostatic form of this equation may be written in isentropic coordinates

$$\frac{\partial}{\partial \theta} \left| \frac{P + v \cdot \nabla \theta}{P} = -\frac{\partial P}{\partial \theta} + P \frac{\partial \theta}{\partial \theta} + \frac{1}{\sigma} \left\{ \frac{\partial v}{\partial \theta} \times \nabla \theta + \nabla \theta \times F \right\}. \quad (8)$$

In a small Rossby number, frictionless flow the final bracketed term is negligible and the source following the two-dimensional motion on an isentropic surface is made up of a term due to advect-
ECMWF data and based on monthly time-averages, $\bar{v}$, of (8) and deviations from this average, $\bar{\sigma}v'$, will be presented for the Middle-world.

The first example is taken from the research of M. Masutani. Shown in Fig. 3a is the zonal average of $\bar{P}$ on isentropic surfaces for January 1990. It should be recalled that in this isentropic framework, the surface of the Earth would slope, though less steeply than the tropopause, from about 300 K at the equator to 250 K at the North Pole. The spreading of PV contours in the tropical troposphere up to 350 K probably represents a partial approach towards the uniform angular momentum, zero PV limit of a 2-D frictionless Hadley circulation (Held and Hou, 1980; K. Emanuel, pers. comm.). This is capped by a lower stratospheric region in which slightly enhanced equatorial PV gradients indicate a possible barrier to inter-hemispheric transport (M. E. McIntyre, pers. comm.).

As an example of the terms in the time-averaged PV equation in the form (8), the zonally averaged contribution of transient eddies to the horizontal advection, $-\bar{v}' \cdot \nabla_{\theta} \bar{P}'$, for January 1990 is given in Fig. 3b. Note that this term is different from the convergence, $-\nabla_{\theta} \cdot \bar{v}' \bar{P}'$, and the flux convergence, $-\nabla_{\theta} \cdot (\bar{\sigma}v') \bar{P}'$. The Southern Hemisphere with its zonally oriented and extended storm track shows a particularly strong dipole signature of transient motions in the region of the tropopause, with advection of sub-tropical, tropospheric air along isentropic surfaces towards the pole and polar, stratospheric air towards the equator. The baroclinic wave life cycle pictures in Hoskins (1990) show these advections occurring in an

---

*Fig. 3.* Zonally averaged statistics in isentropic coordinates for January 1990. The domain is from 310 to 400 K with tick marks every 10 K and from North Pole (left) to South Pole. The approximate position of the tropopause is marked in both frames by a thick line. (a) The zonal and time mean potential vorticity. The contour interval is 1 PVU, and values between $-2$ PVU and $+2$ PVU are shaded. (b) The transient advection term $-\bar{v}' \cdot \nabla \bar{P}'$ with contour interval 0.5 PVU day$^{-1}$. Negative values are stippled.
individual system. A similar signature is apparent in the Northern Hemisphere, but the zonal averaging hides considerable longitudinal structure which is evident in maps on isentropic surfaces. This cross-tropopause mixing of PV, along with the lower tropospheric thermal mixing is probably the dominant effect of transient motions that should be represented in any parametrization of them.

The second example is from the research of D. Hotchkiss, (for other results see Hotchkiss, 1989). It highlights the mean circulation in the upper troposphere and lower stratosphere associated with the Asian summer monsoon. The mean horizontal velocity field on the 360 K surface for August 1988 (Fig. 4a) is dominated by the anticyclonic gyre centred on the northern Indian Ocean. There is a significant cross-equatorial component to the wind in this sector, and a circulation in the anticyclonic sense in the Southern Hemisphere also.

The mean PV distribution at this level (Fig. 4b) shows marked gradients in the middle latitude of both hemispheres. There is also a weak maximum in the poleward gradient near the equator, consistent with 360 K being at the lower level of the equatorial PV barrier in this season also. Comparison of Figs. 4a and b shows that the mean winds are generally closed aligned with the PV contours. The mid-Pacific trough is associated with high PV which appears to be advected around the eastern and south-eastern flanks of the

Fig. 4. Mean fields on the 360 K isentropic surface for August 1988. (a) \( \vec{v} \) with the arrow at the top right corresponding to 50 ms\(^{-1}\). (b) \( F \) with contour interval 1 PVU and values between \(-2\) PVU and \(+2\) PVU shaded.
Northern Hemisphere monsoon anticyclone. There is some negative PV advection on the northwestern flank of the anticyclone.

The PV values in the centre of this anticyclone are between \( \frac{1}{4} \) and \( \frac{1}{2} \) PVU. Elsewhere such values are found only within a few degrees of the equator. This can be viewed as a signature of the convective heating beneath. The explanation is then either in terms of dilution or in terms of a negative material source term, (i.e., the 2nd term on the right-hand side of (8)).

Corresponding to the twin region in the Southern Hemisphere is a band of relatively weak negative PV (anticyclonic). It is clear from Figs. 4a and b that, associated with the cross-equatorial wind component, there is strong positive advection of PV in this region which acts to support the anticyclone. Little convection occurs in the underlying atmosphere here so that the same explanation as in the Northern Hemisphere is not applicable. The positive PV advection description is essentially equivalent to that given by Sardeshmukh and Hoskins (1988). They showed that, from a vorticity perspective, the Southern Hemisphere anticyclone could be viewed as being associated with the divergent outflow from the Northern Hemisphere monsoon region. In particular the cross-equatorial advection of vorticity by the mean divergent wind, \(-\mathbf{v}\cdot\nabla\zeta\), was important.

Fig. 4 is suggestive that the Northern Hemisphere summer monsoon could play an important role in troposphere-stratosphere exchange in the neighbourhood of the 360 K surface. The consequences of this are currently unexplored.

4. An underworld budget result

From (5), the total mass-weighted PV in a layer between \( \theta \)-surfaces is equal to the \( \theta \)-integral of the circulation around the region of interest. For a complete spherical layer around the Earth in the Middleworld or the Overworld the total mass weighted PV must therefore be zero (Haynes and McIntyre, 1987, 1990). However, for an underworld layer between \( \theta_1 \) and \( \theta_2 \) it is clear that the total mass-weighted PV is equal to the integral from \( \theta_1 \) to \( \theta_2 \) of \( C \), the absolute circulation along a \( \theta \) contour on the Earth’s surface, and that this quantity can change.

For the present purposes, we will consider the incremental mass-weighted PV on an isentropic surface \( S \) which intersects the Earth along a contour \( \Gamma \). The surface integral of \( \sigma P \) is equal to the circulation, \( C \), around the isentropic contour \( \Gamma \). It will be shown elsewhere that

\[
\frac{d}{dt} \int_S \sigma P \, dS = \frac{d}{dt} C = \oint_{\Gamma} \left( F_i + \frac{\partial}{\partial \theta/j} \zeta \right) \, ds,
\]

where \( s \) is the distance along \( \Gamma \) in a cyclonic sense (direction \( i \)), \( F_i \) is the component of \( \mathbf{F} \) in this direction, \( n \) is the distance along the surface normal to \( \Gamma \) rotated anticlockwise from \( i \), and \( \zeta \) is a quantity resembling the vertical component of absolute vorticity but this time with horizontal derivatives evaluated on the surface of the Earth. The appearance of \( F_i \) is expected, but the interpretation of the heating term is worthy of discussion. When there is low-level cooling around a circuit enclosing colder air, surface material contours shrink relative to \( \Gamma \). Since the absolute circulation is conserved on the material contours in the absence of friction, it must increase on \( \Gamma \).

Averaging (9) over a long enough period gives the result

\[
\oint_{\Gamma} F_i \, ds = -\oint_{\Gamma} \frac{\partial}{\partial \theta/j} \zeta \, ds.
\]

This constraint again illustrates the power of PV-thinking. It provides an intriguing relationship between mechanical and thermal sources at the Earth’s surface. No eddy flux terms appear but it should be noted that the position of \( \Gamma \) will change through the period.

If we consider a situation in which there are low-level “westerlies”, i.e., winds in the positive sense along \( \Gamma \), then the integrated frictional force must be negative. The usual situation is that \( \partial \theta/j \) and \( \zeta \) would be negative and positive respectively, in the Northern Hemisphere and the opposite signs in the Southern Hemisphere. From (10) this implies that a weighted \( \theta \) is negative, i.e., low-level diabatic cooling. Similarly low-level “easterlies” and heating are associated.

To give an indication of the sense of the frictional integral in (10), the December–February
mean surface potential temperature contours and wind are shown in Fig. 5. The trade-wind easterlies are so prevalent when compared with the anticipated excursion of $\theta_*$ contours that the left-hand side of (10) is almost certainly positive on such contours. To counter the implied increase in circulation it is necessary for there to be diabatic heating. This is entirely in accord with the anticipated heating distribution.

In the westerly belt of the Southern Hemisphere in the summer and winter season, the winds are so strong and the $\theta_*$ contour excursions small enough that the left hand side of (10) is almost certainly negative for such contours. The implied low-level cooling is indeed found in diabatic heating determinations (Hoskins et al., 1989). The significance of such low-level cooling in the westerly belt is, perhaps, a new result.

In the Northern Hemisphere, it is not at all apparent that there are any "westerly" contours. The determination of both sides of (10) requires a proper determination of the integrals as the contours make relatively large excursions. There could be a significant correlation between anomalous equatorward displacement and diabatic heating.

On the other hand, it is probable that the characteristic surface westerly wind errors in many large-scale models is accompanied by anomalous surface cooling. Although cause and effect cannot be determined from such a diagnostic result, it raises new possibilities for the origin of such an error.

The Tibetan Plateau shows as a "warm" region in Fig. 5. In the June–August season it becomes the "hottest" spot on the Earth's surface. (10) shows that sensible heating over the Plateau in this season is responsible for maintaining near surface cyclonic circulation about it. The cyclonic nature of this circulation is associated with the south-westerly flow into India and S.E. Asia which is a crucial component of the Asian summer monsoon. The importance of the Tibetan Plateau sensible heating in the Monsoon is thus directly apparent.

5. Concluding remarks

There are numerous frameworks for viewing the general circulation. The necessity to understand the behaviour of both the atmosphere and complex numerical model simulations means that we must have available to us an increasing array of diagnostic armoury. The studies briefly reported on here, and to be related in full elsewhere, are a contribution to the PV-$\theta$ view of the general circulation, a view that will probably be an important part of that armoury.

6. Acknowledgements

Numerous people have contributed significantly to the present paper. The results in Section 3
are taken from the work of M. Masutani and D. Hotchkiss. P. Sardeshmukh produced Fig. 5. Valuable comments have been received from D. Andrews, M. Blackburn, P. Haynes, M. McIntyre, A. Thorpe and anonymous referees.

REFERENCES


Tellus 43AB (1991), 4