Potential Vorticity Modeling of the ITCZ and the Hadley Circulation

WAYNE H. SCHUBERT, PAUL E. CIESIELSKI, DUANE E. STEVENS AND HUNG-CHI KUO

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 28 November 1990, in final form 27 February 1991)

ABSTRACT

A simple zonally symmetric balanced model of the Hadley circulation is presented. The model is based on potential vorticity arguments and consists of a predictive equation for the potential pseudodensity and an invertibility principle to diagnose the associated balanced wind and mass fields. When the theory is formulated in the potential latitude coordinate, the meridional advection is implicit in the coordinate transformation, which makes the prediction equation for potential pseudodensity analytically solvable. For convective heating patterns that simulate the ITCZ, the model produces upper and lower tropospheric potential vorticity anomalies of opposite sign. The associated winds are easterly at low levels and westerly aloft, except between the equator and the ITCZ, where there are low-level westerlies and upper-level easterlies. Since the potential vorticity anomalies develop within a background state that has potential vorticity increasing to the north, reversed poleward gradients of potential vorticity are produced, just as has been observed in the west African region. For typical convective heating rates, significant potential vorticity gradient reversals occur quickly—on the order of a couple of days. According to the Charney–Stern theorem, such zonal flows are expected to be unstable. In this sense the ITCZ is self-destructive and should not be viewed as a strictly steady state feature of the tropical circulation. In addition, according to this scenario, the potential vorticity dynamics of the west African region are not unique, and we should expect a similar ITCZ formation–breakdown cycle to occur in other tropical regions such as the tropical east Pacific and northern Australia.

1. Introduction

In recent years there has developed an awareness that many aspects of midlatitude balanced dynamics can be most easily understood in terms of the advection of Rossby–Ertel potential vorticity on isentropic surfaces and the determination of the associated balanced wind and mass fields from this potential vorticity field through the invertibility principle (Hoskins et al. 1985). The present paper extends this general potential vorticity approach from midlatitude dynamics to the Hadley circulation problem of tropical dynamics, which has been studied with a variety of other techniques by Schneider and Lindzen (1977), Schneider (1977), Held and Hou (1980), Stevens (1983), Lindzen and Hou (1988), Hack et al. (1989), Hack and Schubert (1990).

To motivate the use of potential vorticity arguments in the study of the Hadley circulation and the ITCZ, consider what happens when an east–west line of deep convection forms near the equator and begins to change the potential vorticity field due to latent heat release. In the Northern Hemisphere such convection induces a positive potential vorticity anomaly at low levels and a negative anomaly aloft. In the Southern Hemisphere low-level negative and upper-level positive anomalies are induced. Since the convectively induced potential vorticity anomalies develop from an initial state which has potential vorticity increasing toward the north, reversed poleward gradients of potential vorticity tend to be produced. The regions of potential vorticity gradient reversal are expected to be found on the poleward side of the ITCZ at low levels and on the equatorward side of the ITCZ at upper levels, just as can be seen in the observed cross section presented by Burpee (1972) for the west African region. This sets the stage for combined barotropic–baroclinic instability, the formation of tropical waves, and the breakdown of the ITCZ. The use of potential vorticity arguments seems the most direct way of understanding this process.

The outline of the paper is as follows. After introducing the zonally symmetric balanced model (section 2), we demonstrate the advantage of the simultaneous use of potential latitude and potential temperature coordinates in the potential pseudodensity equation (section 3). Under certain assumptions on heating, this equation can be solved analytically. Once the potential pseudodensity is known analytically for all time, the zonal wind and mass fields can be found by numerically solving the nonlinear invertibility principle (section 4). In section 5 we use this procedure to illustrate the evolution of the potential pseudodensity, potential vorticity, and associated wind and mass fields for ITCZs cen-

Corresponding author address: Dr. Wayne Schubert, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523.

tered on different latitudes and with different vertical structures of apparent heat source. A basic feature of all the results is the evolution of two regions of reversed poleward gradient of potential vorticity—one poleward of the ITCZ at low levels and the other equatorward of the ITCZ at upper levels.

2. Theory of balanced zonal flow

Using the potential temperature θ as the vertical coordinate, the equations for thermally forced, inviscid, zonally symmetric, balanced flow can be written

$$\frac{Du}{Dt} - \left(2\Omega \sin\phi + \frac{u \tan\phi}{a}\right)v = 0, \tag{2.1}$$

$$\left(2\Omega \sin\phi + \frac{u \tan\phi}{a}\right)u + \frac{\partial M}{a\partial\phi} = 0, \tag{2.2}$$

$$\frac{\partial M}{\partial\theta} = \Pi, \tag{2.3}$$

$$\frac{D\sigma}{Dt} + \sigma\left(\frac{\partial(v\cos\phi)}{a\cos\phi\,\partial\phi} + \frac{\partial\dot{\theta}}{\partial\theta}\right) = 0, \tag{2.4}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v\frac{\partial}{a\partial\phi} + \dot{\theta}\frac{\partial}{\partial\theta} \tag{2.5}$$

is the total derivative, u and v the zonal and meridional components of the wind, $\Pi = c_p(p/p_0)^\kappa$ the Exner function, $M = \theta\Pi + gz$ the Montgomery potential, $\sigma = -\partial p/\partial\theta$ the pseudodensity in θ-space, and $\dot{\theta} = D\theta/Dt$ the effect of heating. Using the definitions of Π and σ, the set (2.1)–(2.4) can be considered closed in the unknowns u, v, p and M. However, it is not a convenient set for prediction since (2.1) and (2.4) cannot be used as independent predictors. In particular, the prediction of u by (2.1) and the prediction of σ by (2.4) must be consistent with a continuous state of hydrostatic and zonal wind balance, as required by (2.2) and (2.3). This implies that (2.1)–(2.4) can be combined into a diagnostic equation which can then replace (2.1) or (2.4). To obtain this diagnostic equation, (2.4) is first written in the form

$$\frac{\partial(\sigma v\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial(\sigma\dot{\theta} - \partial p/\partial t)}{\partial\theta} = 0, \tag{2.6}$$

which implies that $\sigma v = -\partial\psi/\partial\theta$ and $\sigma\dot{\theta} - \partial p/\partial t = \partial(\psi\cos\phi)/(a\cos\phi\partial\phi)$, where ψ is a streamfunction in the meridional plane. If the first of these is inserted into the zonal momentum equation and the second is multiplied by $\Gamma = d\Pi/dp = \kappa\Pi/p$, we obtain

$$\frac{\partial u}{\partial t} + \dot{\theta}\frac{\partial u}{\partial\theta} + P\frac{\partial\psi}{\partial\theta} = 0, \tag{2.7}$$

$$\frac{\partial\Pi}{\partial t} + \dot{\theta}\frac{\partial\Pi}{\partial\theta} + \Gamma\frac{\partial(\psi\cos\phi)}{a\cos\phi\partial\phi} = 0, \tag{2.8}$$

where $P = \zeta/\sigma$ is the potential vorticity and $\zeta = 2\Omega \times \sin\phi - \partial(u\cos\phi)/(a\cos\phi\partial\phi)$ is the absolute *isentropic* vorticity. Equations (2.2) and (2.3) can be combined into the thermal wind equation, the time derivative of which yields

$$\frac{\partial}{\partial\theta}\left(\hat{f}\frac{\partial u}{\partial t}\right) + \frac{\partial}{a\partial\phi}\left(\frac{\partial\Pi}{\partial t}\right) = 0, \tag{2.9}$$

where $\hat{f} = 2\Omega\sin\phi + (2u\tan\phi)/a$. Substituting from (2.7) for $\partial u/\partial t$ and from (2.8) for $\partial\Pi/\partial t$, (2.9) becomes

$$\frac{\partial}{a\partial\phi}\left(\Gamma\frac{\partial(\psi\cos\phi)}{a\cos\phi\partial\phi}\right) + \frac{\partial}{\partial\theta}\left(\hat{f}P\frac{\partial\psi}{\partial\theta}\right) = \frac{\partial(\Pi,\dot{\theta})}{a\partial(\phi,\theta)}. \tag{2.10}$$

If the upper and lower boundaries are considered isobaric and isentropic surfaces along which $\dot{\theta} = 0$, then $\psi = 0$ is the appropriate boundary condition for (2.10). If the lower boundary is an isentropic surface along which $z = 0$ and $\dot{\theta} = 0$, then the lower boundary condition on ψ is more complicated. In any case, we can regard (2.7) as the single predictive equation and (2.10) as the diagnostic equation which must be solved every time step in order to predict u by (2.7). Note that (2.10) is elliptic in regions where $\hat{f}P > 0$, which is generally true except in small regions near the equator where air has moved from one hemisphere to the other. The variable coefficients in (2.10) introduce a basic anisotropy into the dynamical response of the meridional circulation to latent heat release in the ITCZ. Because $\hat{f}P$ is smaller on the equatorial side of the ITCZ, the cross-equatorial branch of the Hadley circulation is much stronger than the poleward branch, a point which is explored in detail by Lindzen and Hou (1988) and Hack et al. (1989).

With the view of balanced dynamics presented in this section [i.e., with (2.7) as the predictive equation and (2.10) as the diagnostic equation], further analytical progress becomes difficult. In the next section we shall take an alternative approach, one in which the predictive equation comes from the potential pseudodensity principle and the diagnostic equation from the invertibility principle. With this alternative approach, further analytical progress is possible.

3. Coordinate transformation and the potential pseudodensity equation

The potential vorticity principle associated with the balanced set (2.1)–(2.4) can be derived by first noting that the equation for the absolute isentropic vorticity takes the form

$$\frac{D\zeta}{Dt} + \zeta\frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi} - \frac{\partial u}{\partial\theta}\frac{\partial\dot{\theta}}{a\partial\phi} = 0. \tag{3.1}$$

Eliminating the isentropic divergence between (2.4) and (3.1) we obtain

$$\sigma \frac{DP}{Dt} = \frac{\partial u}{\partial \theta} \frac{\partial \dot{\theta}}{a \partial \phi} + \zeta \frac{\partial \dot{\theta}}{\partial \theta}. \quad (3.2)$$

This form of the Rossby–Ertel potential vorticity equation will prove useful after we make a coordinate transformation, which will simplify the total derivative operator.

The zonal momentum equation (2.1) can also be written in the absolute angular momentum form $D(\Omega a \times \cos^2\phi + u \cos\phi)/Dt = 0$. Since absolute angular momentum is conserved, one might expect certain advantages in using it as a coordinate in place of ϕ (see Shutts 1980). We follow this general approach but, in particular, use as a new coordinate the potential latitude Φ (Hack et al. 1989) defined as

$$\sin\Phi = \pm\left(1 - \frac{u \cos\phi}{\Omega a \sin^2\phi}\right)^{1/2} \sin\phi, \quad (3.3a)$$

or inversely,

$$\sin\phi = \pm\left(1 + \frac{u^* \cos\Phi}{\Omega a \sin^2\Phi}\right)^{1/2} \sin\Phi, \quad (3.3b)$$

where $u^* \cos\Phi = u \cos\phi$. Squaring both sides of these relations and rearranging, we obtain

$$\Omega a \cos^2\Phi = \Omega a \cos^2\phi + u \cos\phi, \quad (3.3c)$$

which allows us to interpret the potential latitude as the latitude to which an air parcel must be moved (conserving absolute angular momentum) in order for its zonal wind component to vanish. In general there are two such latitudes, as indicated by the plus and minus signs in (3.3a,b). Equations (3.3a–c) relate (Φ, ϕ, u), or, alternatively, (Φ, ϕ, u^*). If ϕ is the independent variable and $\Phi(\phi)$ is known, or if Φ is the independent variable and $\phi(\Phi)$ is known, (3.3c) can be used to find u or u^*. In section 4 we shall formulate a Φ-space invertibility principle in which $\phi(\Phi)$ is part of the solution, so that we shall use (3.3c) to find the zonal wind.

The problem of finding $\Phi(\phi)$ given $u(\phi)$, or of finding $\phi(\Phi)$ given $u^*(\Phi)$, is somewhat more difficult since it involves a choice of roots. To solve this problem we first note that, since the quantities in parentheses in (3.3a,b) must be nonnegative, we must limit our attention to flows for which $-\Omega a \sin^2\Phi \leq u^* \cos\Phi = u \cos\phi \leq \Omega a \sin^2\phi$. Although westerly flows at the equator are excluded by the preceding inequality,[1] this does not appear restrictive for our purposes, since frictionless flows which develop by thermal forcing from an initial state of rest are never westerly at the equator. In order to understand which root should be chosen, consider what happens when a narrow region of ITCZ convective heating is centered near 10°N. Let us assume this heating forces a meridional Hadley circulation that in turn induces a zonal flow that develops from a state of rest (in which case Φ can be identified with the initial latitude). At low levels the air north of the ITCZ is displaced southward and the air south of the ITCZ is displaced northward, as shown in Fig. 1a. The meridional shifts on the south side of the ITCZ are larger because the inertial stability $\hat{I}P$ is smaller there. Figure 1b depicts the associated zonal flow $u \cos\phi$ as a function of ϕ, while Fig. 1c depicts $u^* \cos\Phi$ as a function of Φ. Four parcels of interest are denoted in Fig. 1. Parcel A originated at −2.5°, but is presently on the equator with an easterly zonal flow. It is the only parcel for which the equality $u^* \cos\Phi = -\Omega a \sin^2\Phi$ holds. For all other parcels the inequality $u^* \cos\Phi > -\Omega a \sin^2\Phi$ holds. Parcel B's northward displacement is twice as large as its present distance from the equator, i.e., its initial position was in the Southern Hemisphere at a latitude equal to its present latitude in the Northern Hemisphere. Consequently, it has no zonal velocity. Since the sloping dashed line in Fig. 1a is the line along which $\Phi = 0$, the point C started at the equator. It presently has a westerly zonal velocity and is the only parcel for which the equality $u \cos\phi = \Omega a \sin^2\phi$ holds. For all other parcels the inequality $u \cos\phi < \Omega a \sin^2\phi$ holds. Finally, parcel D is near the middle of the ITCZ and has not experienced any meridional displacement or any zonal acceleration. From this discussion it is clear that all parcels between A and C are presently in the Northern Hemisphere (i.e., $\phi > 0$) but originated in the southern hemisphere (i.e., $\Phi < 0$), so that the negative root in (3.3) should be chosen. For all other parcels the positive root should be chosen. Note that the region between A and C consists of parcels which originated in the southern hemisphere with $P < 0$. Thus, in the region between A and C, $\hat{I}P < 0$ and (2.10) is nonelliptic.

As the thermally forced flow continues to evolve, the westerlies south of the ITCZ become stronger but narrower in meridional extent. In Fig. 1b, point A shifts to the left, point B shifts to larger values of ϕ and point C moves to the right along the upper branch of the $\Omega a \sin^2\phi$ parabola. Correspondingly, in Fig. 1c point A moves to the left along the lower branch of the $-\Omega a \sin^2\Phi$ parabola, point B shifts to more negative values of Φ, and point C shifts to the right.

In the upper troposphere the meridional displacements and associated zonal flows are reversed from those in the lower troposphere. If a $u \cos\phi$ curve for the upper troposphere were drawn in Fig. 1b, it would touch the $\Omega a \sin^2\phi$ parabola on its lower branch. Similarly, if a $u^* \cos\Phi$ curve for the upper troposphere were drawn in Fig. 1c, it would touch the $-\Omega a \sin^2\Phi$ parabola on its upper branch. Thus, in the upper troposphere, the nonelliptic region would lie just south of the physical equator ($\phi < 0$) but just north of the potential equator ($\Phi > 0$).

[1] An anonymous reviewer has pointed out that the restriction $u \leq 0$ at the equator can be circumvented by choosing a coordinate system rotating fast enough for this to be true at all times.

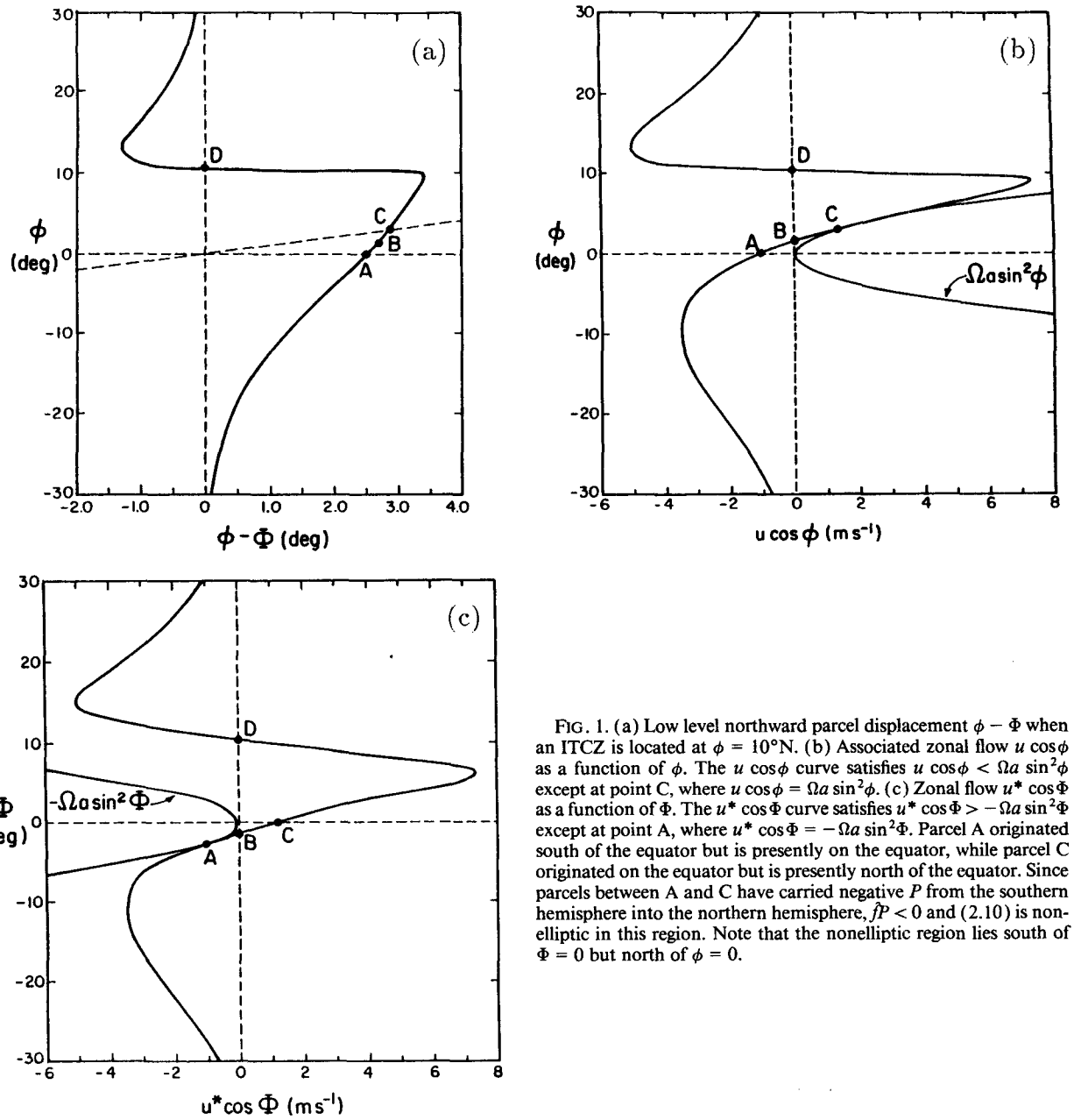


FIG. 1. (a) Low level northward parcel displacement $\phi - \Phi$ when an ITCZ is located at $\phi = 10°$N. (b) Associated zonal flow $u\cos\phi$ as a function of ϕ. The $u\cos\phi$ curve satisfies $u\cos\phi < \Omega a \sin^2\phi$ except at point C, where $u\cos\phi = \Omega a \sin^2\phi$. (c) Zonal flow $u^*\cos\Phi$ as a function of Φ. The $u^*\cos\Phi$ curve satisfies $u^*\cos\Phi > -\Omega a \sin^2\Phi$ except at point A, where $u^*\cos\Phi = -\Omega a \sin^2\Phi$. Parcel A originated south of the equator but is presently on the equator, while parcel C originated on the equator but is presently north of the equator. Since parcels between A and C have carried negative P from the southern hemisphere into the northern hemisphere, $\hat{f}P < 0$ and (2.10) is nonelliptic in this region. Note that the nonelliptic region lies south of $\Phi = 0$ but north of $\phi = 0$.

Let us now consider (Φ, Θ, T) space, where $\Theta = \theta$ and $T = t$. The symbols Θ and T are introduced to distinguish partial derivatives at fixed ϕ ($\partial/\partial\theta$ and $\partial/\partial t$) from partial derivatives at fixed Φ ($\partial/\partial\Theta$ and $\partial/\partial T$). Derivatives in (ϕ, θ, t) space are then related to derivatives in (Φ, Θ, T) space by

$$\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\theta}, \frac{\partial}{\partial t}\right) = \left(\frac{\partial\Phi}{\partial\phi}\frac{\partial}{\partial\Phi}, \frac{\partial\Phi}{\partial\theta}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial\Theta}, \frac{\partial\Phi}{\partial t}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial T}\right) \quad (3.4)$$

the first of which can also be written

$$\frac{\partial}{\cos\phi\,\partial\phi} = \left(\frac{\zeta}{2\Omega\sin\Phi}\right)\frac{\partial}{\cos\Phi\partial\Phi}. \quad (3.5)$$

We limit our attention to flows in which Φ is a monotonically increasing function of ϕ so that $\zeta/(2\Omega\sin\Phi) > 0$. In regions where $\zeta > 2\Omega\sin\Phi$ the Φ coordinate provides a natural stretching which is analogous to the stretching provided around fronts by the geostrophic coordinate in semigeostrophic theory (Hoskins and Bretherton 1972). It is of interest to note that application of (3.5) to $u\cos\phi$ leads to

$$\frac{2\Omega \sin\phi - \partial(u\cos\phi)/(a\cos\phi\partial\phi)}{2\Omega\sin\phi} = \frac{2\Omega\sin\Phi}{2\Omega\sin\Phi + \partial(u^*\cos\Phi)/(a\cos\Phi\partial\Phi)},$$

which shows that, as $-\partial(u^*\cos\Phi)/(a\cos\Phi\partial\Phi)$ approaches $2\Omega\sin\Phi$, the absolute vorticity becomes much larger than the local Coriolis parameter. From (3.4) we can easily show that (2.5) can also be written as

$$\frac{D}{Dt} = \frac{\partial}{\partial T} + \dot\theta\frac{\partial}{\partial\Theta}, \tag{3.6}$$

whose advantage over (2.5) is the elimination of the divergent wind component v.

Let us now introduce as new dependent variables the potential pseudodensity $\sigma^* = \sigma(2\Omega\sin\Phi)/\zeta$ and the Bernoulli function $\mathcal{M} = M + \frac{1}{2}u^2$. The potential pseudodensity σ^* is related to the potential vorticity P by $P\sigma^* = 2\Omega\sin\Phi$ and is simply the pseudodensity a parcel would acquire if ζ were changed to $2\Omega\sin\Phi$ under conservation of P. With the new variables u^* and $\mathcal{M}$, the balance equation (2.2) and the hydrostatic equation (2.3) transform to

$$(u^*, \Pi) = \left(-\left(2\Omega\sin\Phi - \frac{1}{\Omega a\cos\Phi}\frac{\partial\mathcal{M}}{a\partial\Phi}\right)^{-1}\frac{\partial\mathcal{M}}{a\partial\Phi}, \frac{\partial\mathcal{M}}{\partial\Theta}\right). \tag{3.7a,b}$$

Formally, (3.7b) is identical to (2.3) while (3.7a) is simpler than (2.2) in that (3.7a) allows only one u^* for a given $\partial\mathcal{M}/\partial\Phi$.

Our next task is to obtain the potential pseudodensity equation. Since σ^* is proportional to P^{-1}, the potential pseudodensity equation can be easily obtained from the potential vorticity equation (3.2). To begin we note that the first and second parts of (3.4) can be combined to yield $(\partial u/\partial\theta)(\partial/a\partial\phi) + \zeta(\partial/\partial\theta) = \zeta(\partial/\partial\Theta)$, which allows us to rewrite the right hand side of (3.2) and obtain the potential pseudodensity equation

$$\frac{D\sigma^*}{Dt} + \sigma^*\frac{\partial\dot\theta}{\partial\Theta} = 0. \tag{3.8}$$

In the absence of heating σ^* is conserved. However, the upward branch of the Hadley cell is highly nonconservative. As we shall see, a midtropospheric maximum in $\dot\theta$ leads to a decrease of σ^* in the lower troposphere and an increase of σ^* in the upper troposphere. The flux form of (3.8) is particularly convenient. With D/Dt given by (3.6), the flux form of (3.8) becomes

$$\frac{\partial\sigma^*}{\partial T} + \frac{\partial(\sigma^*\dot\theta)}{\partial\Theta} = 0. \tag{3.9}$$

The advantage of (3.9) is that, if the source term $\dot\theta$ is a known function of (Φ, Θ, T), then the problem of solving for the time evolution of σ^* has separated from the rest of dynamics. If $\dot\theta$ is simple enough, (3.9) can even be solved analytically, as was discussed by Schubert and Alworth (1987) and Schubert et al. (1989). Such analytic solutions of (3.9) will be further discussed in section 5.

4. Invertibility principle

If σ^* can be determined by solving (3.9), our final theoretical task is to determine both the wind and mass fields from the σ^* field. This task can be accomplished because the definition of σ^*, along with the gradient and hydrostatic constraints, lead to a coupled pair of equations which relate the known σ^* field to the unknown $\mathcal{M}$ and $\sin\phi$ fields. We shall refer to this pair of equations and its associated boundary conditions as the invertibility principle. To derive the invertibility principle we use the transformation relations (3.4) in the definition of σ^* to obtain the Jacobian form $\partial(s, p)/\partial(S, \Theta) + \sigma^* = 0$, where $s = \sin\phi$ and $S = \sin\Phi$. Using (3.7b) we can write this Jacobian form and the zonal balance condition as

$$\frac{\partial s}{\partial S}\frac{\partial^2\mathcal{M}}{\partial\Theta^2} - \frac{\partial s}{\partial\Theta}\frac{\partial^2\mathcal{M}}{\partial S\partial\Theta} + \Gamma\sigma^* = 0, \tag{4.1a}$$

$$2\Omega^2a^2S\left(\frac{s^2 - S^2}{1 - s^2}\right) + \frac{\partial\mathcal{M}}{\partial S} = 0. \tag{4.1b}$$

Equations (4.1a–b) constitute the desired relation between $\mathcal{M}$, s and σ^*. For boundary conditions we choose

$$\frac{\partial\mathcal{M}}{\partial\Theta} = \Pi_T \quad \text{at} \quad \Theta = \Theta_T, \tag{4.1c}$$

$$\Theta\frac{\partial\mathcal{M}}{\partial\Theta} - \mathcal{M} + \frac{\Omega^2a^2(s^2 - S^2)^2}{2(1 - s^2)} = 0 \quad \text{at} \quad \Theta = \Theta_B, \tag{4.1d}$$

$$s = 1 \quad \text{at} \quad S = 1, \tag{4.1e}$$

$$s = -1 \quad \text{at} \quad S = -1. \tag{4.1f}$$

Equation (4.1c) results from assuming that the upper isentropic surface $\Theta = \Theta_T$ is also an isobaric surface with Exner function Π_T. The lower boundary condition results from assuming the geopotential vanishes on the lower isentropic surface $\Theta = \Theta_B$, so that $M = \Theta\Pi$ there. Then, expressing M in terms of $\mathcal{M}$ and s, we can write the lower boundary condition as (4.1d). For the boundary conditions at the poles, symmetry requires the conditions (4.1e, f).

The diagnostic problem (4.1) involves nonlinearity in both the interior equations and the lower boundary condition. Because of this nonlinearity, (4.1) must be solved using an iterative technique. Note that in (4.1a) the coefficient Γ is in fact a function of the solution $\mathcal{M}$, and should therefore be considered part of the nonlinearity that must be included in the iterative procedure. In section 5 we present some solutions of (4.1),

which were obtained using the procedure described in the Appendix.

We can now summarize the results of our analysis as follows. If the time evolution of the σ^* field can be determined from (3.9), we can then solve the diagnostic problem (4.1) for $\mathcal{M}$ and s, after which u^*, u, Π and p are easily determined. This is all accomplished in (S, Θ) space. The transformation to other representations, e.g., $u(\phi, \theta)$ or $u(\phi, p)$, is straightforward.

5. Solutions of the potential pseudodensity equation and the invertibility principle for ITCZ forcing

We now turn to the problem of solving (3.9). For simplicity let us consider the case in which $\dot{\theta}$ is independent of time and is given by $\dot{\theta} = Q(S)\sin^2(\pi Z)$, where $Z = (\Theta - \Theta_B)/(\Theta_T - \Theta_B)$ and $Q(S)$ is the latitudinal distribution of the specified heating. We postpone the specification of $Q(S)$ since only the vertical dependence of $\dot{\theta}$ is required for our analytic result. Multiplying (3.9) by $\dot{\theta}$ we obtain

$$\frac{\partial}{\partial \tau}(\dot{\theta}\sigma^*) + \sin^2(\pi Z)\frac{\partial}{\partial Z}(\dot{\theta}\sigma^*) = 0, \quad (5.1)$$

where $\tau(S) = Q(S)T/(\Theta_T - \Theta_B)$ is the dimensionless "convective clock" time. According to (5.1) the quantity $\dot{\theta}\sigma^*$ is constant along each characteristic curve determined from $dZ/\sin^2(\pi Z) = d\tau$. By integration of this equation we can show that the characteristic through the point (Z, τ) intersects the $\tau = 0$ axis at a level $Z_0(Z, \tau)$ determined by $\pi Z_0(Z, \tau) = \cot^{-1}[\pi\tau + \cot(\pi Z)]$. Since $\dot{\theta}\sigma^*$ is constant along each characteristic, its value at (Z, τ) must equal its value at $(Z_0(Z, \tau), 0)$, which results in

$$\sigma^*(Z, \tau) = \sigma^*(Z_0(Z, \tau), 0) \times \frac{\sin^2\{\cot^{-1}[\pi\tau + \cot(\pi Z)]\}}{\sin^2(\pi Z)}. \quad (5.2a)$$

Although (5.2a) is indeterminant at the boundaries $Z = 0$ and $Z = 1$, use of l'Hopital's rule twice yields

$$\sigma^*(Z, \tau) = \sigma^*(Z_0(Z, \tau), 0) \quad \text{at} \quad Z = 0, 1. \quad (5.2b)$$

Since $Z_0(Z, \tau) \to 0$ as $Z \to 0$ and $Z_0(Z, \tau) \to 1$ as $Z \to 1$, the solution (5.2) constitutes an *internal* potential pseudodensity anomaly, i.e., a σ^* field which is unmodified at the upper and lower boundaries. The reason for this can be seen by referring back to (3.8) or (3.9) and noting that, for our specified $\dot{\theta}$ field, both $\dot{\theta}$ and $\partial\dot{\theta}/\partial\Theta$ vanish at the boundaries. For most of the results presented here we have specified the initial σ^* to be a constant, i.e., $\sigma^*(Z, 0) = \sigma_0$, which implies an initial state with no zonal flow.

Equations (5.2) constitute the analytic solution of the potential pseudodensity equation when the diabatic source has the $\sin^2(\pi Z)$ form. The complete solution $\sigma^*(S, \Theta, T)$ can be plotted once $Q(S)$, and hence $\tau(S)$, is specified. Since the τ clock runs faster where $Q(S)$ is large, the largest anomalies in the σ^* field will occur in the ITCZ.

For the latitudinal distribution of the heating we now choose the particular form

$$Q(S) = Q_0 4\alpha\pi^{-1/2}\{\text{erf}[\alpha(1 + S_c)] + \text{erf}[\alpha(1 - S_c)]\}^{-1}\exp[-\alpha^2(S - S_c)^2]. \quad (5.3)$$

By varying the parameters S_c and α we can consider simulated ITCZs centered at different latitudes and with different widths. By integration of (5.3) from the south pole to the north pole we can show that $\frac{1}{2}\int Q(S)dS = Q_0$, so that different values of S_c and α all result in the same area averaged heating Q_0. For the results shown here we have chosen $\alpha = 15$ and either $S_c = \sin(10°) \approx 0.174$ or $S_c = \sin(15°) \approx 0.259$. These can be interpreted as rather narrow ITCZs with approximately 85% of their rainfall occurring between 6°N and 14°N for the $S_c = \sin(10°)$ case or between 11°N and 19°N for the $S_c = \sin(15°)$ case. Because of the way the product $Q(S)T$ appears in the definition of $\tau(S)$, it is not really necessary to choose Q_0; rather, the solution can simply be obtained for different values of Q_0T. However, for purposes of physical interpretation let us choose $Q_0 = 0.30$ K day^{-1}, along with $\Theta_T = 360$ K and $\Theta_B = 300$ K. Then, the peak heating is $Q(S_c) \approx 5.1$ K day^{-1} and $T = 3$, 6 days correspond to $Q_0T = 0.9$, 1.8 K, or $\tau(S_c) \approx 0.26$, 0.51.

For the case of an ITCZ at 10°N the fields of $\sigma^*(\phi, \theta)$, $P(\phi, \theta)$, $u(\phi, \theta)$ and $p(\phi, \theta)$ at 3 and 6 days are shown in Figs. 2 and 3. The σ^* field has been normalized by $\sigma_0 = 1.458$ kPa K^{-1} and the P field by $2\Omega/\sigma_0$. In the ITCZ a region of small potential pseudodensity develops at lower levels and a region of large potential pseudodensity at upper levels. Due to vertical advection in the ITCZ, the lower tropospheric minimum in σ^* begins to form an indentation on the upper tropospheric maximum in σ^*. This same process occurs in a more extreme form in the development of a tropical cyclone (Schubert and Alworth 1987). The solution of the invertibility principle results in low level zonal flows, which are easterly except in a band that runs between a latitude just north of the equator and a latitude near the center of the ITCZ. At upper levels the zonal flow is westerly except in a band that runs between a latitude just south of the equator and a latitude near the center of the ITCZ. As the σ^* and P anomalies become larger, the associated zonal flows also become larger. The isolines of pressure in the bottom panels of Figs. 2 and 3 reveal only small adjustments in the mass field, with a slight stabilization at lower levels in the ITCZ and a slight destabilization aloft.

Perhaps the most striking result seen in Figs. 2 and 3 is that a narrow potential pseudodensity or potential vorticity anomaly produced in just a few days by convection in the ITCZ can result in significant zonal winds throughout the entire tropical and subtropical region.

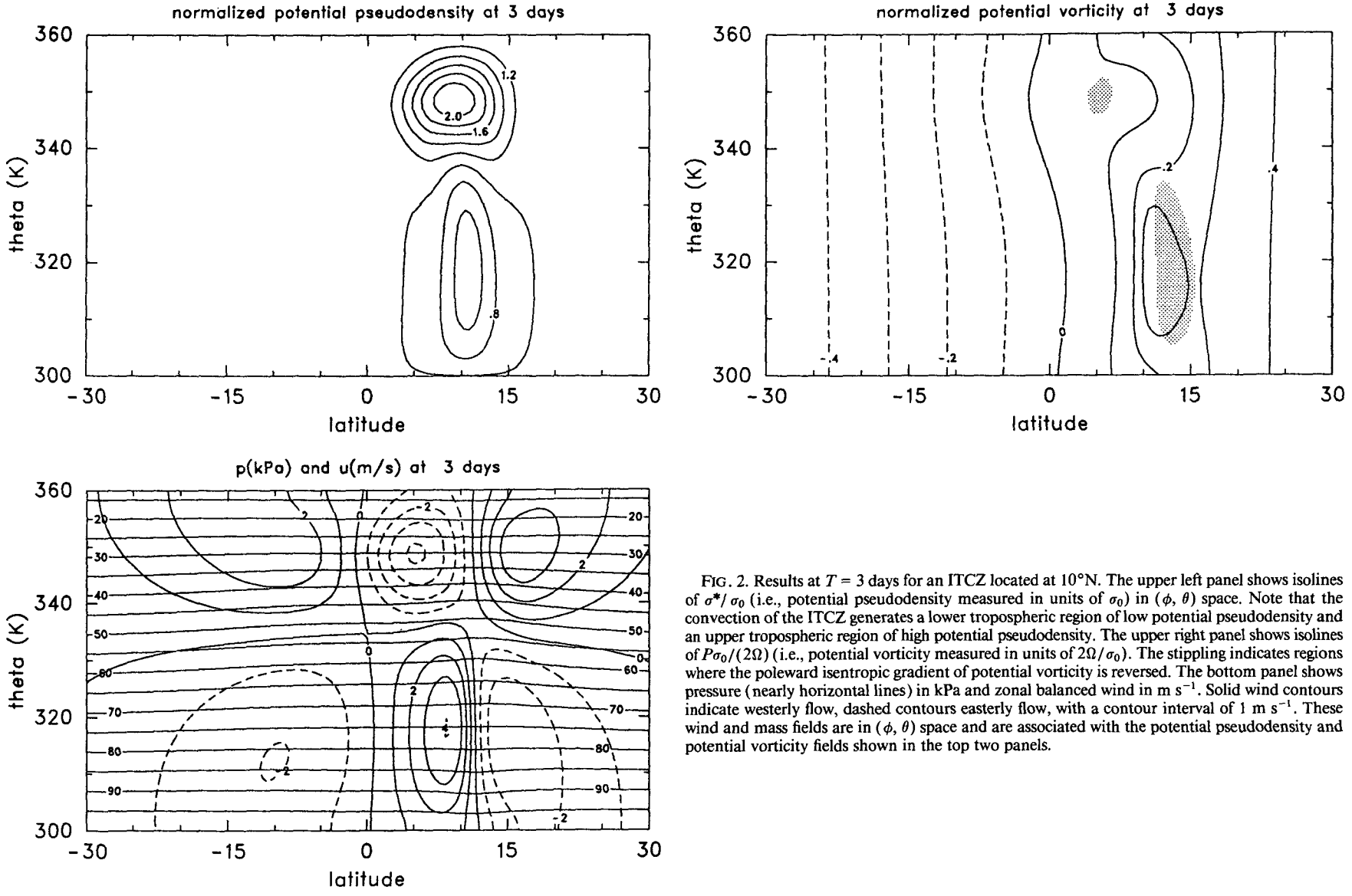
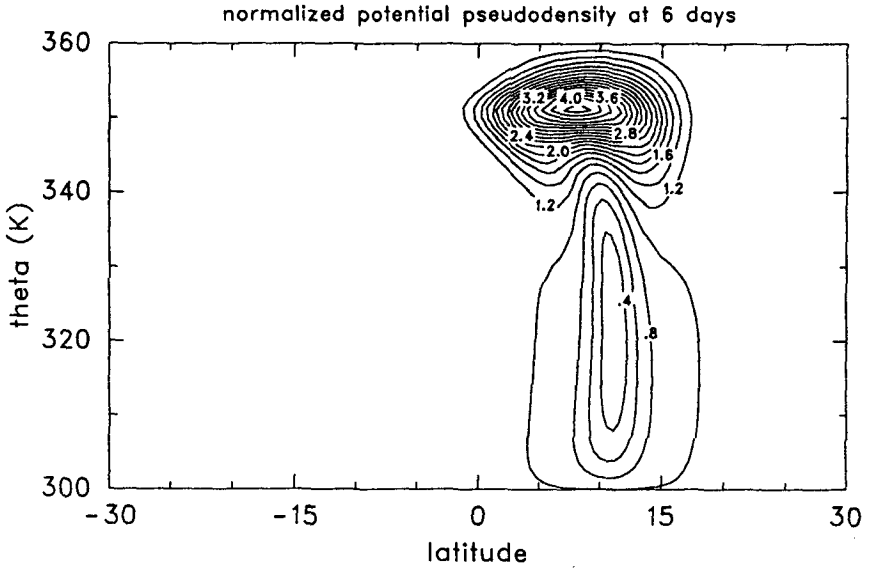
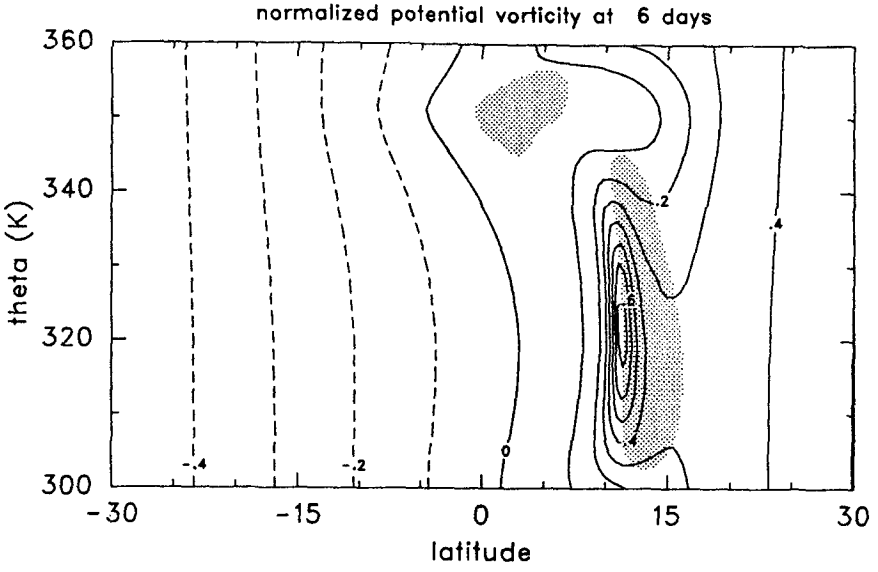


FIG. 2. Results at $T = 3$ days for an ITCZ located at 10°N. The upper left panel shows isolines of σ^*/σ_0 (i.e., potential pseudodensity measured in units of σ_0) in (ϕ, θ) space. Note that the convection of the ITCZ generates a lower tropospheric region of low potential pseudodensity and an upper tropospheric region of high potential pseudodensity. The upper right panel shows isolines of $P\sigma_0/(2\Omega)$ (i.e., potential vorticity measured in units of $2\Omega/\sigma_0$). The stippling indicates regions where the poleward isentropic gradient of potential vorticity is reversed. The bottom panel shows pressure (nearly horizontal lines) in kPa and zonal balanced wind in m s^{-1}. Solid wind contours indicate westerly flow, dashed contours easterly flow, with a contour interval of 1 m s^{-1}. These wind and mass fields are in (ϕ, θ) space and are associated with the potential pseudodensity and potential vorticity fields shown in the top two panels.





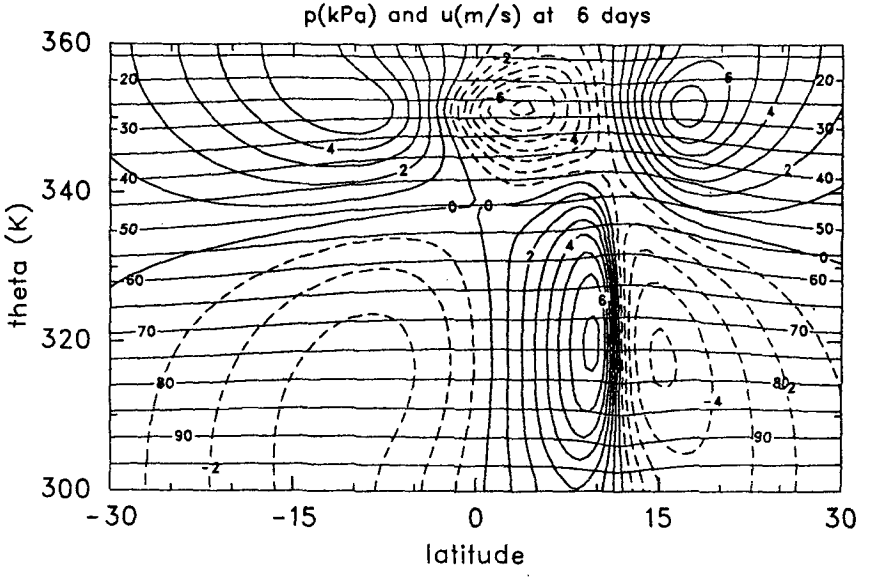


FIG. 3. As in Fig. 2 but for $T = 6$ days.

This result is related to the meridional parcel displacements forced by the convection. Since Φ is a conservative quantity and $\dot{\theta}$ is known, and since the actual latitude $\phi(\Phi, \Theta)$ is part of the solution of the invertibility problem, meridional parcel displacements or trajectories are easy to construct. Two sets of such trajectories from the initial time to 3 days and from 3 to 6 days are shown in Fig. 4, along with the $\dot{\theta}$ field. Away from the ITCZ, $\dot{\theta} = 0$ and parcel trajectories are along isentropic surfaces. In a lower tropospheric layer bounded by two isentropes, mass is removed in the ITCZ, and there is a shift in parcel positions toward the ITCZ. The largest shifts are on the cross-equatorial side, because $\hat{f}P$ is smallest there. Corresponding shifts away from the ITCZ occur in an upper tropospheric layer bounded by two isentropes. As the heating proceeds, $\hat{f}P$ becomes larger in the lower troposphere near the ITCZ. This increasing resistance to motion along isentropic surfaces, coupled with the fixed $\dot{\theta}$ field, causes the depth of the ITCZ inflow to deepen with time.

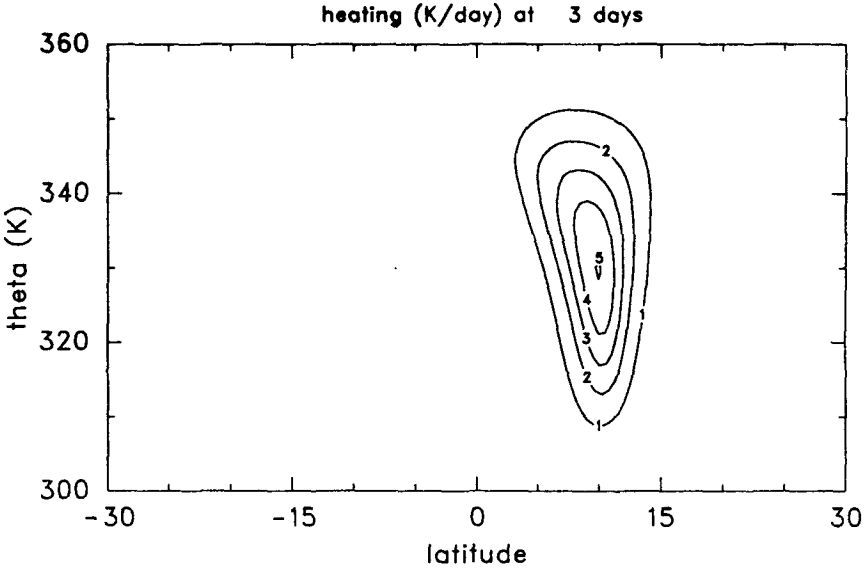
The anisotropic response or enhancement of the cross-equatorial Hadley cell has interesting effects on the potential vorticity field. To see this, consider the -0.1 and the 0.4 potential vorticity lines in Figs. 2 and 3. These P lines mark chains of fluid parcels beginning approximately equal distances from the ITCZ. At 6 days the -0.1 line is more distorted than the 0.4 line. This is a direct result of the fact that the meridional circulation associated with the cross-equatorial Hadley cell is more intense than the meridional circulation associated with the Hadley cell north of the ITCZ. The $P = 0$ curve marks the chain of fluid particles which started at rest on the equator. Regardless of the hemisphere into which these particles move, they must acquire a westerly flow, since they move closer to the axis of the earth's rotation. Thus, the $P = 0$ line bends more than the $u = 0$ line, so that it lies in the lower tropospheric westerlies north of the equator and the upper tropospheric westerlies south of the equator.

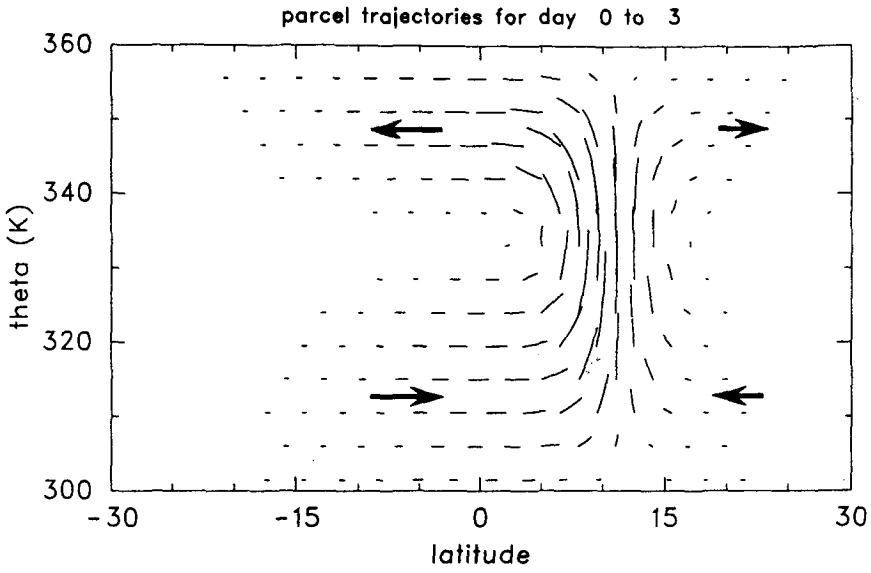
The convective modification of the P field occurs within a background state that has a northward increase of P. As convection continues, the gradient of P becomes locally reversed in the lower troposphere poleward of the ITCZ and in the upper troposphere equatorward of the ITCZ. These regions of reversed isentropic poleward gradient of potential vorticity are indicated by stippling in Figs. 2 and 3. Such features develop quickly and are consistent with observations made by Burpee (1972) and Reed et al. (1977) in their studies of the origins of easterly waves in the lower troposphere of the north African region. According to Charney and Stern (1962), Eliassen (1983) and Shepherd (1989), such zonal flows (i.e., those with a reversal in the meridional gradient of the potential vorticity) satisfy the necessary condition for combined barotropic–baroclinic instability. Thus, it would appear that ITCZ convection alone can lead to the generation of unstable zonal flows. This may be the cause of periodic breakdowns of the ITCZ. The precise role of the upper-level potential vorticity gradient reversal remains intriguing and is complicated by the fact that it can come into close proximity with the inertially unstable (Stevens 1983) region where $\hat{f}P < 0$. Is it possible that at upper levels on the equatorward side of the ITCZ there are generated waves which take the form of equatorially trapped modes propagating into the stratosphere and playing a role in the QBO?

Results at 6 days for an ITCZ located at 15°N are shown in Fig. 5. Comparing Fig. 5 with Fig. 3 we note that, except for the latitudinal shift, the σ^* fields are essentially identical. However, the potential vorticity, zonal wind and mass fields are different, with the ITCZ at 15°N producing a potential vorticity anomaly, neighboring zonal winds and isobaric surface deviations considerably larger than those produced by the ITCZ at 10°N. These differences can be interpreted as follows. Since $D\sigma^*/Dt = -\sigma^* \partial\dot{\theta}/\partial\Theta$ and the initial σ^* is constant, the time evolution of σ^* for ITCZs at different latitudes is essentially identical except for the meridional shift. Since $DP/Dt = P\partial\dot{\theta}/\partial\Theta$ and the initial P increases to the north, the material rate of change of P is larger for an ITCZ at 15°N. An alternate interpretation is that, since $P = (2\Omega \sin\Phi)/\sigma^*$, identical σ^* anomalies shifted from 10°N to 15°N result in P anomalies which are approximately 50% larger for the ITCZ at 15°N.

All results shown so far have an apparent heat source with $\sin^2(\pi Z)$ vertical structure, which leads to internal potential vorticity anomalies. Figure 6 shows results at 6 days for an ITCZ at 10°N when the apparent heat source has a $\sin(\pi Z)$ vertical structure. In this case $\partial\dot{\theta}/\partial\Theta$ does not vanish at the top and bottom boundaries, and the extrema in σ^*, P and u occur on the boundaries. It is debatable which vertical profile of apparent heat source is more realistic. Most diagnostic studies (e.g., Yanai et al. 1973, Fig. 10) would tend to favor the $\sin(\pi Z)$ case in the lower troposphere and the $\sin^2(\pi Z)$ case in the upper troposphere. However, in some physical situations where stratiform precipitation plays an important role, both of the vertical profiles used here would need significant modification. For example, the apparent heat source in a region of stratiform precipitation is characterized by heating due to net condensation at upper levels and cooling due to evaporation at lower levels. The corresponding $\partial\dot{\theta}/\partial\Theta$ then leads to a midtropospheric maximum in potential vorticity (Hertenstein and Schubert 1991).

The results shown in Figs. 2–6 were all obtained with a constant initial σ^*. For comparison with observations let us now consider an initial σ^* that depends on θ but not ϕ. This represents an initial condition with no zonal flow and with a mean tropical temperature profile. The mean temperature profile is taken from the GATE mean sounding (Fulton and Schubert 1985, Table 4). Beginning with this initial σ^*, we obtain the results shown in Fig. 7 at $T = 6$ days. The results





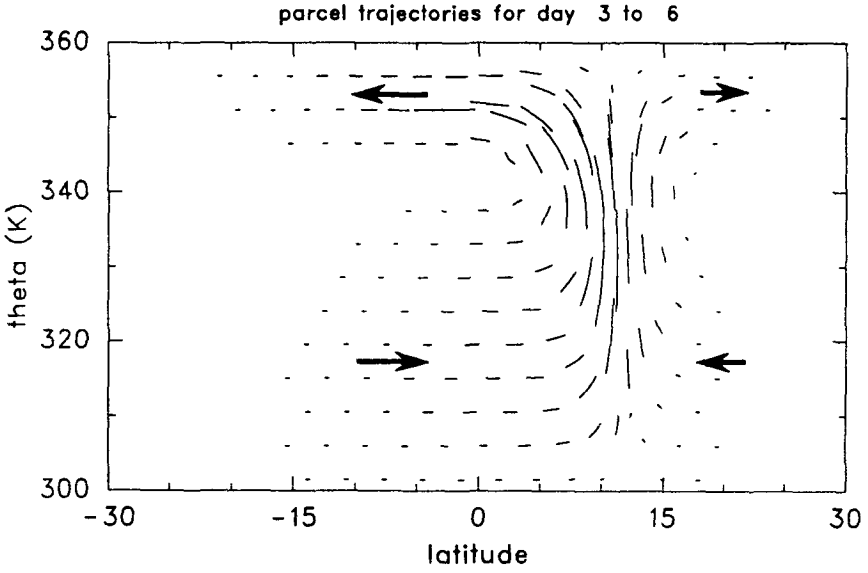
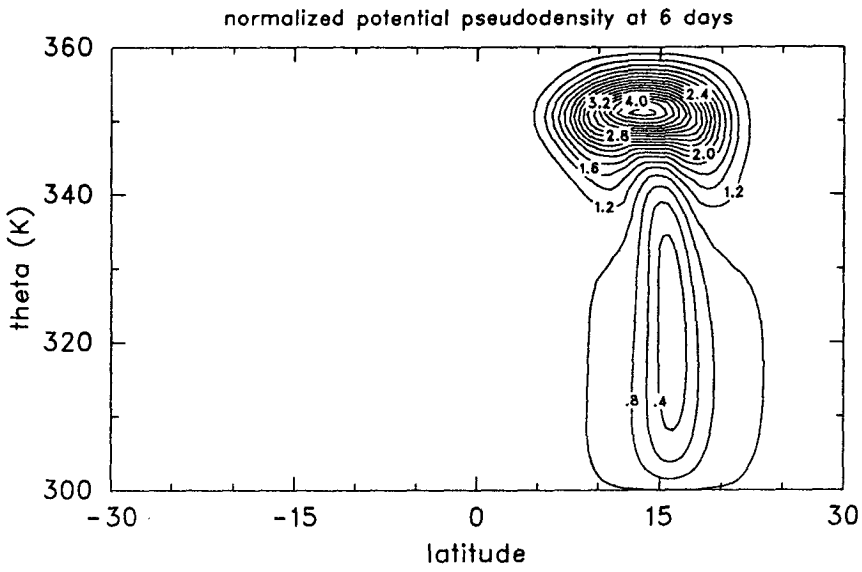
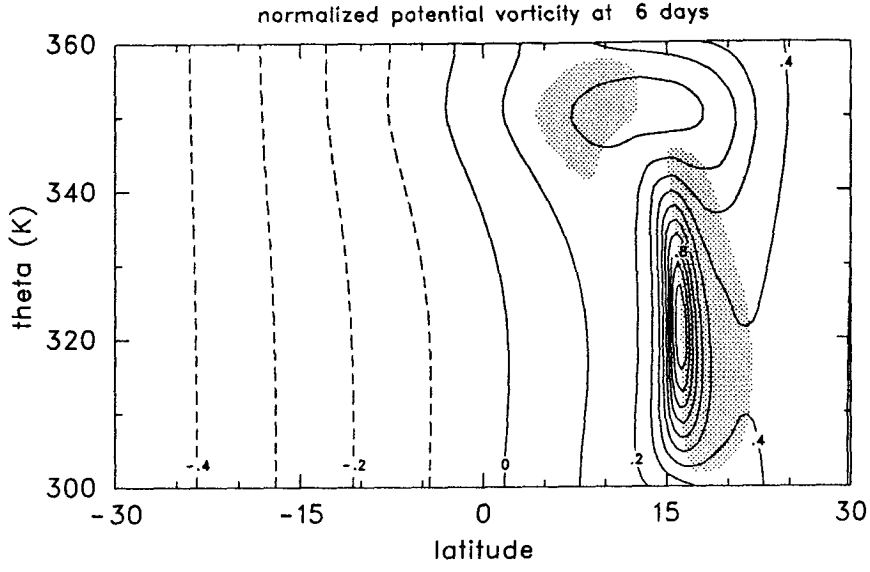


FIG. 4. The upper left panel shows the heating function $\dot{\theta}(\phi, \theta)$ in K day^{-1} at 3 days. The distortion of the $\dot{\theta}$ field from a true Gaussian function results from the transformation to actual latitude ϕ. The other two panels show parcel trajectories from the initial time to 3 days and from 3 days to 6 days.





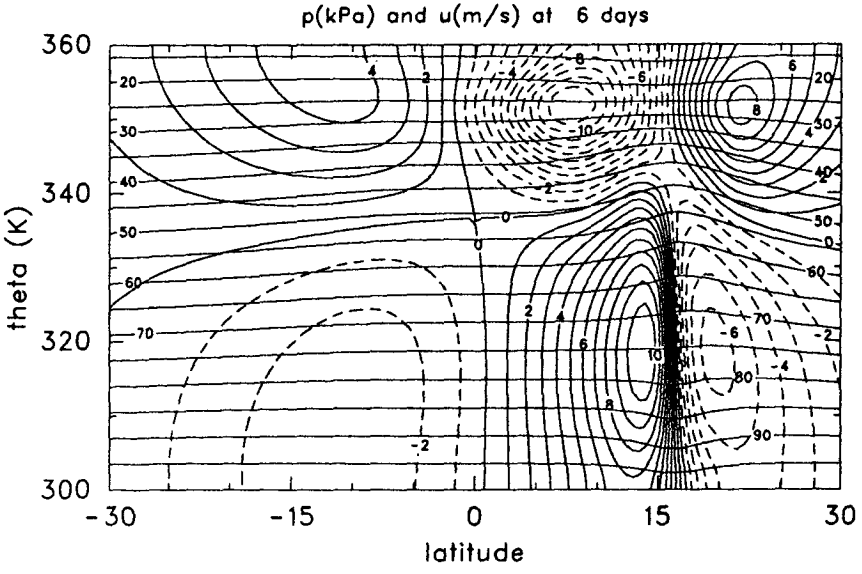


FIG. 5. As in Figs. 2 and 3 but for $T = 6$ days with an ITCZ at 15°N. In comparing this figure with Fig. 3, we note that an ITCZ at 15°N produces a potential vorticity anomaly and associated zonal winds approximately 50% stronger than for an ITCZ at 10°N.

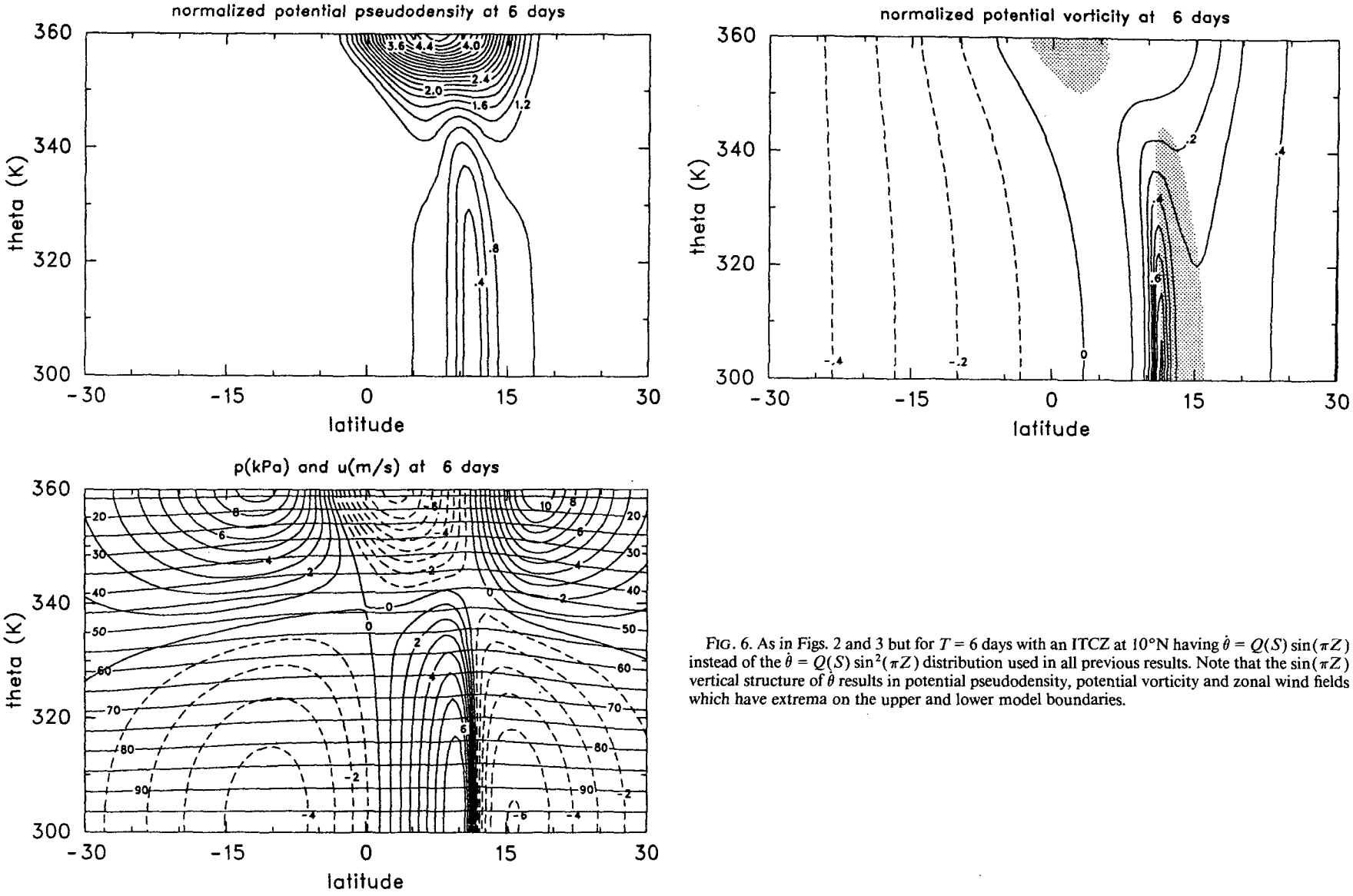


FIG. 6. As in Figs. 2 and 3 but for $T = 6$ days with an ITCZ at 10°N having $\dot{\theta} = Q(S)\sin(\pi Z)$ instead of the $\dot{\theta} = Q(S)\sin^2(\pi Z)$ distribution used in all previous results. Note that the $\sin(\pi Z)$ vertical structure of $\dot{\theta}$ results in potential pseudodensity, potential vorticity and zonal wind fields which have extrema on the upper and lower model boundaries.

are qualitatively similar to those presented previously, but the potential pseudodensity and potential vorticity anomalies occur in a more realistic background state which includes low values of σ^* and high values of P in the stratosphere. The top panel in Fig. 8 shows an expanded, Northern Hemisphere view of the potential vorticity for the same experiment at $T = 3$ days. The bottom panel in Fig. 8 is an isentropic coordinate adaptation of Burpee's (1972) August mean cross section of potential vorticity at 5°E. During August the ITCZ rainfall maximum stretches across Africa at approximately 11°N (Burpee 1972, Fig. 2). Thus, both the observations and the model show two regions of reversed potential vorticity gradient, one poleward of the ITCZ at low levels and one equatorward of the ITCZ at upper levels. In a further analysis of the characteristics of North African easterly waves during the summers of 1968 and 1969, Burpee (1974, Figs. 7 and 8) showed that the maximum amplitude of the surface meridional wind and surface pressure oscillations occur at approximately 18°N, i.e., near the northern edge of the region of reversed potential vorticity gradient. From the model simulations it is clear that the instability of the lower tropospheric easterly jet can be explained by ITCZ convection alone; there is no need to invoke surface heating in the Saharan region. Once this is accepted, there is no reason to regard the African region as particularly unique, and we should expect this same instability and breakdown of the easterly jet in all regions with a well-defined ITCZ.

6. Concluding remarks

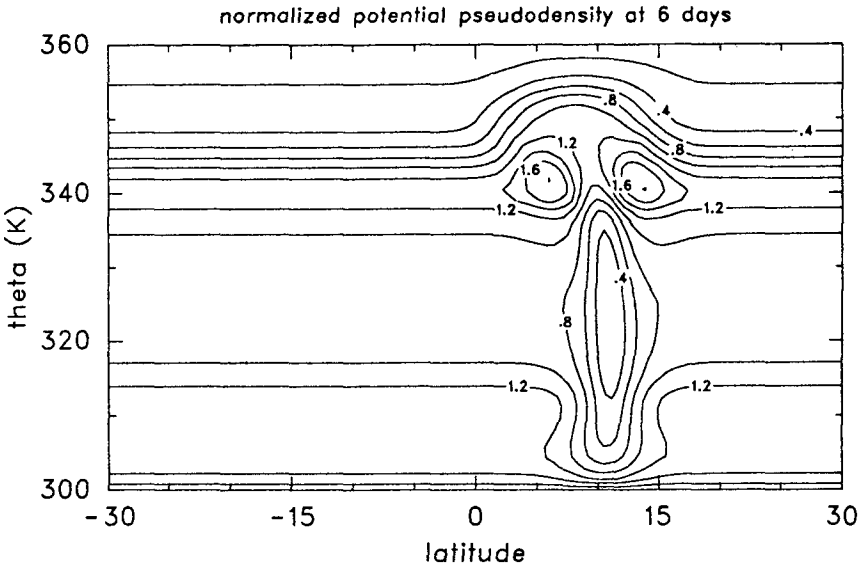
In the preceding analysis we have made extensive use of the concepts of potential vorticity and potential pseudodensity. This raises the question of whether one of these variables is preferable to the other. We believe the answer to this question depends on the context of the discussion. For example, let us consider the problem of how to define the tropopause. Operational definitions of the tropopause often involve criteria based on abrupt changes of lapse rate. Dynamicists, especially those concerned with tropopause folding, often prefer criteria based on abrupt changes of potential vorticity. The potential vorticity approach, while more dynamically justifiable than the lapse rate approach, fails in tropical regions, as can be seen from the difficulty involved in defining a tropopause near the equator based on the potential vorticity field shown in the upper right panel of Fig. 7. However, the potential pseudodensity field shown in the upper left panel of Fig. 7 has a behavior that allows it to serve as the basis of a dynamical definition of the tropopause at all latitudes. In other discussions, such as those involving combined barotropic–baroclinic instability, potential vorticity may be preferable. Thus, it may be useful to regard potential vorticity and potential pseudodensity as dual variables, with the preference of one over the other depending on what aspect of the dynamics is being discussed.

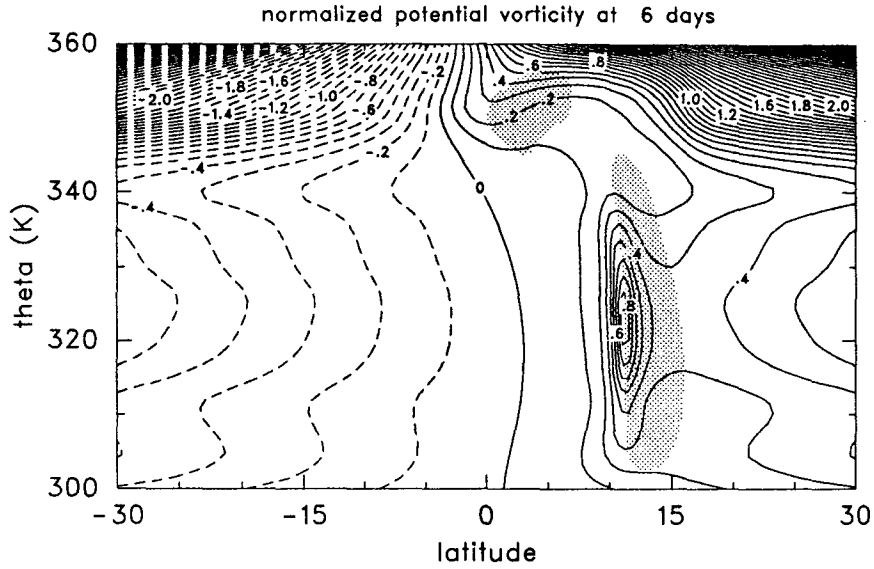
An alternate interpretation of the potential vorticity dynamics occurring in convective situations has recently been provided by Haynes and McIntyre (1987, 1990). It is based on rewriting (3.2) as

$$\frac{\partial(\sigma P)}{\partial t} + \frac{\partial}{a\cos\phi\,\partial\phi}\left\{\left(\sigma P v - \dot{\theta}\frac{\partial u}{\partial\theta}\right)\cos\phi\right\} = 0, \tag{6.1}$$

and regarding P as the mixing ratio of a notional potential vorticity substance (PVS), i.e., P = PVS/mass or σP = PVS/volume of (ϕ, θ)-space. According to (6.1) the total advective and nonadvective flux ($\sigma P v - \dot{\theta}\partial u/\partial\theta$) is along the isentrope, and an isentropic surface is semipermeable in the sense that it is permeable to mass but impermeable to potential vorticity. Multiplying (6.1) by $a\cos\phi$ and integrating from pole to pole and over some layer bounded by two isentropic surfaces, we obtain $\int\int P\sigma a\cos\phi d\phi d\theta = 0$, which states that the total potential vorticity substance in this layer is always zero. In a certain sense, the positive potential vorticity in the Northern Hemisphere cancels the negative potential vorticity in the Southern Hemisphere, although, of course, the zero potential vorticity line is not always exactly at the equator. According to the Haynes–McIntyre interpretation, the impermeability of isentropic surfaces to potential vorticity, coupled with the withdrawal of mass from the lower isentropic layers causes concentration of potential vorticity there, while the addition of mass to the upper isentropic layers causes dilution of potential vorticity there.

There are several ways to extend the analysis presented here. One obvious possibility is to include the effects of friction in (2.1). Then, (3.9) includes the additional term $\partial(\sigma^*\dot{\Phi}\cos\Phi)/(\cos\Phi\partial\Phi)$, and it becomes possible to investigate flows which are thermally and frictionally controlled, including steady state flows that are produced when the vector $(\sigma^*\dot{\Phi}\cos\Phi, \sigma^*\dot{\theta}\cos\Phi)$ is nondivergent in the (Φ, Θ) plane. A second extension of the present analysis becomes apparent by noting that, although the $\dot{\theta}$ field chosen in section 5 simulates the apparent heat source associated with the ITCZ, it neglects the apparent heat source associated with trade cumulus convection (Nitta and Esbensen 1974). This additional heat source tends to produce a low potential pseudodensity trade inversion layer with a high potential pseudodensity region below it. Since warmer sea-surface temperatures lead to deeper trade cumuli, the increase in sea-surface temperature toward the ITCZ should cause this potential pseudodensity couplet to slope upward toward the ITCZ. By the invertibility principle, variations in potential pseudodensity along isentropic surfaces induce zonal flows. Thus, shallow convection may be important for two reasons—the moistening of low-level air flowing toward the ITCZ and the production of zonal flow associated with the sloping potential pseudodensity couplet. Solutions resulting from trade cumulus forcing might help





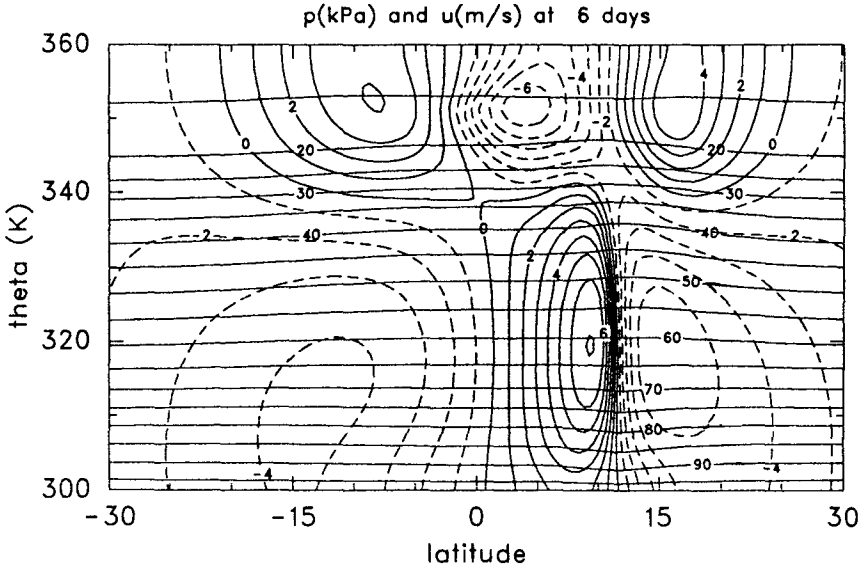


FIG. 7. As in Fig. 3 but for a model solution initialized with a temperature profile taken from the GATE mean sounding.

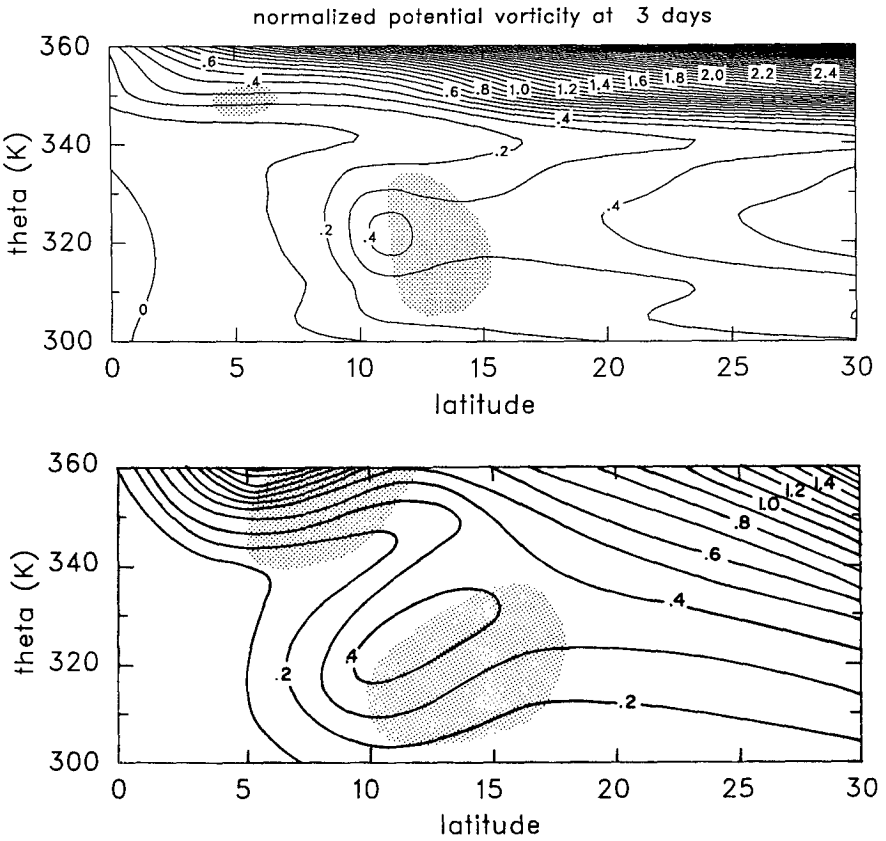


FIG. 8. The top panel shows an expanded, Northern Hemisphere view of the potential vorticity for the same experiment as in Fig. 7 but at $T = 3$ days. The bottom panel is an isentropic coordinate adaptation of Burpee's (1972) August mean cross section of potential vorticity at 5°E.

us better understand the importance of the parameterization of shallow convection in general circulation models.

Since we have bypassed the parameterization of moist physical processes, we cannot claim to have developed a closed theory of the Hadley circulation. However, even though the present work and the work of Thorpe (1985, 1986), Schubert and Alworth (1987), and Schubert et al. (1989) do not constitute closed theories, their value may be that they force us to take a somewhat different view of convectively driven circulations, with primary attention focused on the potential pseudodensity or potential vorticity fields. Since this is precisely the view often taken in midlatitude baroclinic wave and frontogenesis studies using semigeostrophic theory, we have moved closer to the goal of a unified dynamical framework for the study of midlatitude and tropical phenomena. A natural extension of the present work would be a fully three-dimensional theory that has (3.9) and (4.1) as a special two-dimensional case. Steps in this direction have recently been taken by Shutts (1989) and Magnusdottir and Schubert (1990, 1991) using three-dimensional generalizations of semigeostrophic theory with a variable Coriolis parameter and by Stevens et al. (1990) using the long-wave approximation.

Acknowledgments. We wish to thank Scott Fulton, Chungu Lu, Gudrun Magnusdottir and Gerald Taylor for their helpful comments on the solution of the invertibility principle. This work was supported by the National Science Foundation, under grants ATM-8814541 and ATM-8918725, and the Office of Naval Research, under Grant N000014-88-K-0214. Acknowledgement is made to the National Center for Atmospheric Research, which is sponsored by the National Science Foundation, for much of the computing support used in this research.

APPENDIX

Numerical Methods for Solution of the Invertibility Principle

We wish to use an iterative technique to solve the discretized version of (4.1). Because of the anisotropy associated with the variable Coriolis parameter, the approach here is somewhat different than that of Fulton (1989) for a related f-plane invertibility problem. In

particular, we do not use point relaxation but rather a more efficient procedure based on Newton's method and line relaxation (Fulton et al. 1986).

Using the same nondimensional vertical coordinate Z introduced in section 5, and nondimensionalizing $\mathcal{M}$ by $c^2 = \alpha R(\Theta_T - \Theta_B)$, σ^* by $\sigma_0 = (p_B - p_T)/(\Theta_T - \Theta_B)$, Γ by $\Gamma_0 = R/p_B$, Π by c_p, and p by p_B, the invertibility relation (4.1) becomes

$$\frac{\partial s}{\partial S}\frac{\partial^2 \mathcal{M}}{\partial Z^2} + \frac{\epsilon S s(1-S^2)}{(1-s^2)^2}\left(\frac{\partial s}{\partial Z}\right)^2 + \Gamma\sigma^* = 0, \quad \text{(A.1a)}$$

$$\frac{\epsilon S}{2}\left(\frac{s^2 - S^2}{1-s^2}\right) + \frac{\partial \mathcal{M}}{\partial S} = 0, \quad \text{(A.1b)}$$

while the boundary conditions become

$$\kappa\alpha\frac{\partial \mathcal{M}}{\partial Z} = (1-\alpha)^\kappa \quad \text{at} \quad Z = 1, \quad \text{(A.1c)}$$

$$\beta\frac{\partial \mathcal{M}}{\partial Z} - \mathcal{M} + \frac{\epsilon}{8}\frac{(s^2 - S^2)^2}{1-s^2} = 0 \quad \text{at} \quad Z = 0, \quad \text{(A.1d)}$$

$$s = 1 \quad \text{at} \quad S = 1, \quad \text{(A.1e)}$$

$$s = -1 \quad \text{at} \quad S = -1, \quad \text{(A.1f)}$$

where $\Gamma = \Pi^{(\kappa-1)/\kappa}$, $\Pi = \kappa\alpha\partial\mathcal{M}/\partial Z$, $\epsilon = 4\Omega^2 a^2/c^2$ (Lamb's parameter), $\alpha = (p_B - p_T)/p_B$ and $\beta = \Theta_B/(\Theta_T - \Theta_B)$. Note that we have retained the same symbols for nondimensional quantities. For the results presented in this paper we have used $\Omega = 7.292 \times 10^{-5}$ rad s^{-1}, $a = 6371$ km, $R = 287$ J kg^{-1} K^{-1}, $\Theta_B = 300$ K, $\Theta_T = 360$ K, $p_B = p_0 = 100$ kPa and $p_T = 12.5$ kPa, so that $c = 122.7$ m s^{-1} and $\epsilon = 57.3$.

In order to discretize the invertibility principle (A.1), consider the grid defined by $S_j = j\Delta S$ ($j = -J, -J+1, \cdots, J$) with $\Delta S = 1/J$ and $Z_k = k\Delta Z$ ($k = -1, 0, \cdots, K+1$) with $\Delta Z = 1/K$. Let us define $\mathcal{M}$ at all even j points and s at all odd j points. By choosing J odd, the poles $S = \pm 1$ are s points at which the boundary conditions (A.1e, f) are imposed. We now define

$$x_{j,k} = \begin{cases} \mathcal{M}_{j,k}, & \text{if } j \text{ even} \\ s_{j,k}, & \text{if } j \text{ odd,} \end{cases} \quad \text{(A.2)}$$

and use standard centered differences to discretize (A.1a) at the $\mathcal{M}$ points and (A.1b) at the s points. Then the discrete versions of (A.1a,b) are

$$G_{j,k} = 0, \quad \text{(A.3a)}$$

where

$$G_{j,k} = \Gamma_{j,k}\sigma^*_{j,k}2\Delta S(\Delta Z)^2 + (x_{j+1,k} - x_{j-1,k})(x_{j,k+1} - 2x_{j,k} + x_{j,k-1}) + \frac{\epsilon\Delta S S_j(1 - S_j^2)(x_{j+1,k} + x_{j-1,k})}{[1 - \frac{1}{4}(x_{j+1,k} + x_{j-1,k})^2]^2}\left[\frac{1}{4}(x_{j+1,k+1} + x_{j-1,k+1} - x_{j+1,k-1} - x_{j-1,k-1})\right]^2$$

for j even,

$$G_{j,k} = \epsilon\left(\frac{x_{j,k}^2 - S_j^2}{1 - x_{j,k}^2}\right)S_j\Delta S + x_{j+1,k} - x_{j-1,k}$$

for j odd, $\Gamma_{j,k} = \Pi_{j,k}^{(\kappa-1)/\kappa}$, and $\Pi_{j,k} = \kappa\alpha(\mathcal{M}_{j,k+1} - \mathcal{M}_{j,k-1})/(2\Delta Z)$. The boundary conditions are

$$x_{j,K+1} = x_{j,K-1} + \frac{2\Delta Z(1-\alpha)^\kappa}{\kappa\alpha}, \quad \text{(A.3b)}$$

$$x_{j,-1} = x_{j,1} - \frac{2\Delta Z}{\beta}x_{j,0} + \frac{\epsilon\Delta Z}{8\beta} \times \left[\frac{(x_{j-1,0}^2 - S_{j-1}^2)^2}{1 - x_{j-1,0}^2} + \frac{(x_{j+1,0}^2 - S_{j+1}^2)^2}{1 - x_{j+1,0}^2}\right], \quad \text{(A.3c)}$$

$$x_{J,k} = 1, \quad \text{(A.3d)}$$

$$x_{-J,k} = -1. \quad \text{(A.3e)}$$

The interior equations (A.3a) are applied for $j = -J+1, \cdots, J-1$ and $k = 0, 1, \cdots, K$, and the boundary conditions (A.3b–c) use centered differences with ghost points at $k = -1$ and $k = K+1$. For the results presented here we have chosen $J = 81$ and $K = 40$.

Suppose we have an estimate of x on a row of interior points $j = -J+1, \cdots, J-1$ and on surrounding rows, but this estimate does not satisfy (A.3). Holding the values of x on surrounding rows fixed, we wish to simultaneously update the entire row of x's so that (A.3a) is satisfied for the entire row. Then, omitting the subscript k, we have $2J-1$ nonlinear equations $G_j = 0$, which can be written in the vector form $\mathbf{G}(\mathbf{x}) = 0$, where $\mathbf{x}$ is a vector consisting of the $2J-1$ values of x_j. Newton's method for this system is

$$\mathcal{J}(\mathbf{x}^{\text{new}} - \mathbf{x}^{\text{old}}) + \mathbf{G}(\mathbf{x}^{\text{old}}) = 0, \quad \text{(A.4)}$$

where $\mathcal{J}$ is the Jacobian matrix of the system. In particular $\mathcal{J}$ is a tridiagonal matrix whose j, j' entry is $\mathcal{J}_{j,j'} = \partial G_j/\partial x_{j'}$. Zebra relaxation is used to update $\mathbf{x}$ by (A.4) at all interior lines. One sweep of zebra relaxation consists of: first, simultaneously solving the tridiagonal systems for all white (even) lines with new values of the unknown replacing the old ones; then using these updated values, the same procedure is repeated for all black (odd) lines. After updating all interior points, the ghost points are updated using (A.3b–c). This linear relaxation procedure on a single grid is not particularly efficient since smooth error components are slow to be eliminated. However, it could be optimized by incorporation into a multigrid procedure such as that described by Fulton (1989) for the f-plane invertibility problem. Even though several hundred sweeps are required to obtain satisfactory convergence, each solution

shown here was obtained in about ten minutes on a SUN Sparcstation 1.

REFERENCES

Burpee, R. W., 1972: The origin and structure of easterly waves in the lower troposphere of North Africa. *J. Atmos. Sci.*, **29**, 77–90.

——, 1974: Characteristics of North African easterly waves during the summers of 1968 and 1969. *J. Atmos. Sci.*, **31**, 1556–1570.

Charney, J. G., and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.

Eliassen, A., 1983: The Charney-Stern theorem on barotropic–baroclinic instability. *Pure Appl. Geophys.*, **121**, 563–572.

Fulton, S. R., 1989: Multigrid solution of the semigeostrophic invertibility relation. *Mon. Wea. Rev.*, **117**, 2059–2066.

——, P. E. Ciesielski and W. H. Schubert, 1986: Multigrid methods for elliptic problems: A review. *Mon. Wea. Rev.*, **114**, 943–959.

——, and W. H. Schubert, 1985: Vertical normal mode transforms: Theory and application. *Mon. Wea. Rev.*, **113**, 647–658.

Hack, J. J., and W. H. Schubert, 1990: Some dynamical properties of idealized thermally-forced meridional circulations in the tropics. *Meteor. Atmos. Phys.*, **44**, No. 1–4, 101–118.

——, ——, D. E. Stevens and H.-C. Kuo, 1989: Response of the Hadley circulation to convective forcing in the ITCZ. *J. Atmos. Sci.*, **46**, 2957–2973.

Haynes, P. H., and M. E. McIntyre, 1987: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.*, **44**, 828–841.

——, and ——, 1990: On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.*, **47**, 2021–2031.

Held, I. M., and A. Y. Hou, 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37**, 515–533.

Hertenstein, R. F. A., and W. H. Schubert, 1991: Potential vorticity anomalies associated with squall lines. *Mon. Wea. Rev.*, **119**, in press.

Hoskins, B. J., and F. P. Bretherton, 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. *J. Atmos. Sci.*, **29**, 11–37.

——, M. E. McIntyre and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.

Lindzen, R. S., and A. Y. Hou, 1988: Hadley circulations for zonally averaged heating centered off the equator. *J. Atmos. Sci.*, **45**, 2416–2427.

Magnusdottir, G., and W. H. Schubert, 1990: On the generalization of semi-geostrophic theory to the β-plane. *J. Atmos. Sci.*, **47**, 1714–1720.

——, and ——, 1991: Semigeostrophic theory on the hemisphere. *J. Atmos. Sci.*, **48**, 1449–1456.

Nitta, T., and S. Esbensen, 1974: Heat and moisture budget analyses using BOMEX data. *Mon. Wea. Rev.*, **102**, 17–28.

Reed, R. J., D. C. Norquist and E. E. Recker, 1977: The structure and properties of African wave disturbances as observed during phase III of GATE. *Mon. Wea. Rev.*, **105**, 317–333.

Schneider, E. K., 1977: Axially symmetric steady state models of the basic state for instability and climate studies. Part II. Nonlinear calculations. *J. Atmos. Sci.*, **34**, 380–296.

——, and R. S. Lindzen, 1977: Axially symmetric steady state models of the basic state for instability and climate studies. Part I. Linear calculations. *J. Atmos. Sci.*, **34**, 253–279.

Schubert, W. H., and B. T. Alworth, 1987: Evolution of potential vorticity in tropical cyclones. *Quart. J. Roy. Meteor. Soc.*, **113**, 147–162.

——, S. R. Fulton and R. F. A. Hertenstein, 1989: Balanced atmospheric response to squall lines. *J. Atmos. Sci.*, **46**, 2478–2483.

Shepherd, T. G., 1989: Nonlinear saturation of baroclinic instability. Part II: Continuously stratified fluid. *J. Atmos. Sci.*, **46**, 888–907.

Shutts, G. J., 1980: Angular momentum coordinates and their use in zonal, geostrophic motion on a hemisphere. *J. Atmos. Sci.*, **37**, 1126–1132.

——, 1989: Planetary semi-geostrophic equations derived from Hamilton's principle. *J. Fluid Mech.*, **208**, 545–573.

Stevens, D. E., 1983: On symmetric stability and instability of zonal mean flows near the equator. *J. Atmos. Sci.*, **40**, 882–893.

——, H.-C. Kuo, W. H. Schubert and P. E. Ciesielski, 1990: Quasi-balanced dynamics in the tropics. *J. Atmos. Sci.*, **47**, 2262–2273.

Thorpe, A. J., 1985: Diagnosis of balanced vortex structure using potential vorticity. *J. Atmos. Sci.*, **42**, 397–406.

——, 1986: Synoptic scale disturbances with circular symmetry. *Mon. Wea. Rev.*, **114**, 1384–1389.

Yanai, M., S. Esbensen and J.-W. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627.