# The Surface Branch of the Zonally Averaged Mass Transport Circulation in the Troposphere

ISAAC M. HELD

*NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey*

TAPIO SCHNEIDER

*Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey*

(Manuscript received 10 February 1998, in final form 23 July 1998)

ABSTRACT

The near-surface branch of the overturning mass transport circulation in the troposphere, containing the equatorward flow, is examined in isentropic and geometric coordinates. A discussion of the zonal momentum balance within isentropic layers shows that the equatorward flow at a given latitude is confined to isentropic layers that typically intersect the surface at that latitude. As a consequence of mass transport within the surface mixed layer, much of the equatorward flow occurs in layers with potential temperatures below the mean surface potential temperature.

In the conventional transformed Eulerian mean formulation for geometric coordinates, the surface branch of the overturning circulation is represented in an unrealistic manner: streamlines of the residual circulation do not close above the surface. A modified residual circulation is introduced that is free from this defect and has the additional advantage that its computation, unlike that of the conventional residual circulation, does not require division by the static stability, which may approach zero in the planetary boundary layer. It is then argued that cold air advection by the residual circulation is responsible for the formation of surface inversions at all latitudes in idealized GCMs with weak thermal damping. Also included is a discussion of how a general circulation theory for the troposphere must be built upon a theory for the near-surface meridional mass fluxes.

## 1. Introduction

The mean meridional mass transport within isentropic layers in the troposphere is characterized by poleward flow in the upper troposphere and equatorward flow near the ground. Cross-isentropic rising in low latitudes, corresponding to diabatic heating, and cross-isentropic sinking in high latitudes, corresponding to diabatic cooling, balance the divergence and convergence of the horizontal mass transport, forming an overturning circulation. The near-surface branch of this overturning circulation, containing the equatorward flow, is the focus of the present paper.

In section 2, we discuss from the perspective of isentropic coordinates how the range of potential temperatures at which the equatorward flow occurs relates to the distribution of instantaneous surface potential temperatures. We then turn to the transformed Eulerian mean, or residual, circulation (Andrews and McIntyre 1976; Edmon et al. 1980), which provides an approximation in geometric coordinates to the zonal-mean isentropic mass circulation. In its conventional form, the residual circulation need not be closed above the surface; rather, its streamlines can intersect the ground. To conserve mass, one must think of at least part of the equatorward mass transport as occurring within an infinitesimally thin layer along the surface. A more realistic representation of the near-surface flow results from redefining the residual circulation in a way similar to that used in an oceanic context by Treguier et al. (1997), which, in turn, is a special case of a generalized residual circulation defined by Andrews and McIntyre (1978). In section 3, we introduce this modified residual circulation for the atmosphere, which approximates the zonal-mean isentropic mass circulation in the same sense as does the conventional residual circulation but which, in contrast to the latter, has all streamlines closed above the ground.

Since the resulting equatorward flow in the modified circulation is distributed over a layer of finite thickness, it can be viewed as advecting the potential temperature. This advection may play a quantitatively important role in the formation of surface inversions, which frequently

*Corresponding author address:* Dr. Isaac M. Held, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University, P.O. Box 308, Princeton, NJ 08542.
E-mail: ih@gfdl.gov

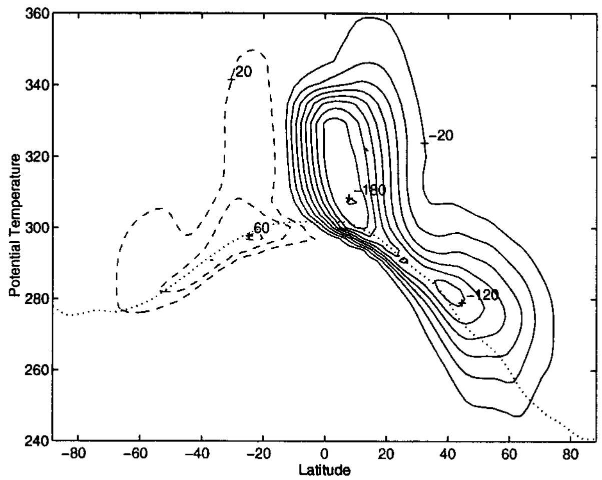


Fig. 1. Isentropic mass transport streamfunction $\Psi_\Theta$ [$10^9$ kg s$^{-1}$] for Jan (50-yr mean of GCM integration); (solid lines) clockwise and (dashed lines) counterclockwise circulation. The dotted line represents the median surface potential temperature.

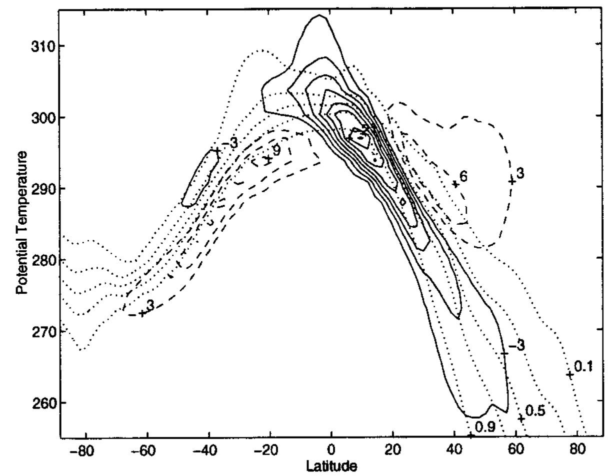


Fig. 2. The meridional mass transport $\partial_\Theta \Psi_\Theta$ [$10^9$ kg s$^{-1}$ K$^{-1}$] near the surface. The dotted lines represent the 10%, 30%, 50%, 70%, and 90% isolines of the cumulative surface potential temperature distribution.

occur during winter in high latitudes and are commonly attributed to strong radiative cooling of the ground. Idealized general circulation models (GCMs) suggest, however, that besides radiative cooling, advection of cold air by the residual circulation may also contribute to the formation of wintertime inversions. As described in section 4, in an idealized GCM with sufficiently weak surface thermal damping, polar air is advected equatorward by the near-surface residual circulation, creating a strong surface inversion not just at high latitudes, but at all latitudes. Strong thermal damping is required to prevent the formation of this global advectively generated inversion.

The discussion of the overturning circulation's surface branch suggests a conceptual framework within which theories of the tropospheric mean circulation can be elaborated. Such a framework is discussed in section 5. The conclusions are summarized in section 6.

## 2. The isentropic perspective

### *a. Dynamics of the mass transport circulation*

The mean meridional mass flux between the isentropes $\Theta$ and $\Theta + \delta\Theta$ is given by $V(\Theta)\delta\Theta = \overline{vH}^{\Theta}\delta\Theta$, where $v$ denotes the meridional velocity and the thickness $H \equiv -g^{-1}\partial p/\partial\Theta$ plays the role of a density in isentropic coordinates; the symbol $\overline{(\,)}^{\Theta}$ stands for a zonal average along isentropes. The mean meridional mass circulation can be represented by a streamfunction, $\Psi_\Theta$, defined such that

$$V \equiv \frac{1}{2\pi a \cos(\theta)} \frac{\partial \Psi_\Theta}{\partial \Theta},$$

with $\theta$ denoting latitude and $a$ the radius of the earth.

Figure 1 shows the time mean of the streamfunction $\Psi_\Theta$ of the zonally averaged isentropic overturning circulation in an atmospheric GCM. The GCM, the atmospheric component of a climate model currently in use at the Geophysical Fluid Dynamics Laboratory, is a $\sigma$-coordinate spectral model with R30 resolution in the horizontal and with 14 vertical levels. The streamfunction in Fig. 1 represents the average over all Januarys of a 50-yr model integration with prescribed seasonally varying sea surface temperatures.

The streamfunction in the GCM is in rough qualitative agreement with the observed streamfunctions presented in Johnson (1989, Fig. 8) and Karoly et al. (1997). In particular, the GCM reproduces the separation of the flow into two distinct overturning cells in the winter hemisphere: a tropical Hadley cell dominated by the mass transport due to the ageostrophic meridional flow and a midlatitude cell associated with extratropical eddies.

Figure 2 displays the surface branch of the overturning circulation and the surface potential temperature distribution in greater detail. Contours represent the time average of the zonally integrated meridional mass transport per unit potential temperature, $\partial_\Theta \Psi_\Theta = [2\pi a \cos(\theta)]V$. Included in this figure are the estimated 10%, 30%, 50%, 70%, and 90% probability isolines of the surface potential temperature distribution. The 10% isoline marks that surface potential temperature that, at a given latitude, is exceeded only 10% of the time. The other probability isolines have an analogous interpretation. The 50% curve, the median, approximates the mean surface potential temperature.

From Fig. 2, we see that equatorward mass transport only occurs in isentropic layers that, at the latitude in question, are frequently interrupted by the surface; the mean poleward mass transport is confined to layers that are rarely or never interrupted by the surface. Closer inspection shows that the equatorward mass flux attains its maximum at potential temperatures that are lower than the mean surface potential temperature. We can rationalize these observations by considering the momentum balance within isentropic layers.

In a statistically steady state and in the absence of dissipation, the zonal momentum budget reads (Andrews et al. 1987, chap. 3)

$$\overline{v}^*\overline{(f + \zeta)}^\Theta = -\overline{H}^\Theta\overline{\hat{v}\hat{P}}^* + \overline{Q\partial_\Theta u}^\Theta, \qquad (1)$$

where the absolute vorticity $f + \zeta$ is the sum of the Coriolis parameter $f$ and the relative vorticity $\zeta$, $P = (f + \zeta)/H$ is the potential vorticity (PV), and $Q = D\Theta/Dt$ the diabatic heating rate. The symbol $\overline{()}^* \equiv \overline{H()}^\Theta/\overline{H}^\Theta$ denotes a mass-weighted zonal average along isentropic layers and $(\hat{\ }) \equiv () - \overline{()}^*$ deviations therefrom. In the interior of the extratropical troposphere, the vertical advection of zonal momentum $\overline{Q\partial_\Theta u}^\Theta$ and the term $\overline{\zeta}^\Theta\overline{v}^*$ are both smaller than $f\overline{v}^*$ by a factor on the order of the Rossby number, so $f\overline{v}^*$ must be balanced by the eddy PV flux term $-\overline{H}^\Theta\overline{\hat{v}\hat{P}}^*$. The meridional mass flux $V = \overline{H}^\Theta\overline{v}^*$ must therefore be poleward wherever the eddy potential vorticity flux is southward.

In isentropic layers that are not interrupted by the ground, the thickness $H$ and with it the potential vorticity $P = (f + \zeta)/H$ are well-defined. In any such layer, we expect the mean PV gradient to be positive (that is, PV increasing northward) or possibly close to zero in regions where PV has been homogenized by sufficiently strong eddy stirring. The eddy PV flux associated with eddy stirring is usually downgradient and hence southward. Exceptions to this rule of thumb occur only when the radiative forcing is such that it reverses the natural (planetary) PV gradient, or in cases in which some other mechanism besides the downgradient transport of PV dominates the generation of potential enstrophy (Rhines and Holland 1979). The observational data presented in Edmon et al. (1980) reveal that, at least in winter, the average eddy PV flux in the interior of the midlatitude troposphere is indeed southward.

Combining this argument for southward eddy PV fluxes with the above relation between the eddy PV flux and the meridional mass flux, we expect poleward mass fluxes in uninterrupted isentropic layers, in which PV is well defined. The equatorward return flow must then be confined to interrupted layers near the surface.

As suggested by the form of the transformed Eulerian mean equations [see Andrews et al. (1987); Edmon et al. (1980); and section 3 of the present paper], the equatorward mass transport in the interrupted layers is related to the poleward eddy heat flux near the surface. In making the connection between the near-surface mass flux and the eddy heat flux, it is convenient to define at each latitude $\theta$ the minimum potential temperature $\Theta_I(\theta)$ above which we can be reasonably confident that isentropic layers will not be interrupted by the surface. We denote the instantaneous pressure at this isentropic interface as $p_I$ and refer to the air below it as the *surface layer*. The zonal momentum balance of the air in the surface layer can be approximated as

$$D + C + S \approx 0, \qquad (2)$$

where $D$ is the form drag on the upper boundary of the surface layer plus any mountain torque at the surface, $C$ is the Coriolis force due to the meridional mass transport within the surface layer, and $S$ is the surface stress. The convergence of horizontal eddy momentum fluxes also contributes to the zonal momentum budget, but since it is observed to be small near the surface, we neglect it here.

The poleward mass flux in the surface layer can be written as

$$g^{-1}\overline{\tilde{v}(p_s - p_I)} = g^{-1}\overline{\tilde{v}_g(p_s - p_I)} + g^{-1}\overline{\tilde{v}_a(p_s - p_I)}. \qquad (3)$$

Here $\tilde{v}$ represents the meridional velocity's vertical average across the surface layer. This average meridional velocity has been decomposed into a geostrophic part, $\tilde{v}_g$, and an ageostrophic part, $\tilde{v}_a$. The overbar indicates an average over $x$ at a given time, and $p_s$ is the surface pressure. Accordingly, decomposing the Coriolis force contribution $C$ to the zonal momentum budget into two parts, $C = C_g + C_a$, one can show that the Coriolis force $C_g$ associated with the geostrophic mass flux balances the form drag on the surface layer: $C_g + D = 0$. More generally, on every isentrope $\Theta$ it holds that

$$(f/g)\overline{v_g\partial_\Theta p}^\Theta - \partial_\Theta\overline{(p\partial_x z)}^\Theta = 0, \qquad (4)$$

where $v_g$ is the geostrophic part of the local meridional velocity. The appendix contains a derivation of this result, which is related to but differs slightly from the derivation in Andrews (1983).

The remaining terms in the momentum balance are the zonal surface stress $S$ and the Coriolis force $C_a$ due to the ageostrophic mass transport: $C_a + S \approx 0$. The ageostrophic transport in (3) can thus be viewed as the Ekman drift associated with the surface stress. In the region of surface westerlies in midlatitudes, this Ekman drift is poleward; however, the total mass transport, the sum of the Ekman drift and the geostrophic mass flux, is equatorward. The geostrophic mass flux must therefore dominate the Ekman drift. Since, in the extratropics, the surface stress determining the Ekman drift approximately balances the vertically integrated eddy momentum flux convergence, the partial cancellation of the geostrophic mass flux by the Ekman drift can be neglected to the extent that the eddy momentum flux convergence is negligible compared with the form drag. From the perspective of wave activity conservation, this means that the Ekman drift is negligible to the extent that the wave activity generated near the surface is dissipated at the same latitude rather than being transported to other latitude zones (e.g., Held and Hoskins 1985). In the atmosphere's surface layer, the Ekman drift is not completely negligible compared with the form drag and the associated geostrophic mass flux, but it is small enough to be ignored in a qualitative first approximation.

Setting the surface layer thickness $H_s$ equal to the value of $H$ near the surface and assuming that its variations are small, we can approximate the eddy-induced variations in the surface layer mass per unit area by

$$g^{-1}(p_s' - p_I') \approx -H_s\Theta_s', \tag{5}$$

where $\Theta_s'$ is the near-surface eddy potential temperature and primes denote deviations from the zonal mean. Only fluctuations in the surface potential temperature appear on the right-hand side of (5) because $\Theta_I$ is, by definition, constant along latitude circles. Neglecting variations in the geostrophic velocity across the surface layer, and using the fact that the zonal average of the geostrophic flow vanishes, the geostrophic mass flux can be written as

$$g^{-1}\overline{\tilde{v}_g(p_s - p_I)} = g^{-1}\overline{\tilde{v}_g(p_s' - p_I')} \approx - H_s\overline{\tilde{v}_g\Theta_s'}. \tag{6}$$

We expect downgradient (poleward) eddy potential temperature fluxes near the surface partly for the same reasons we expect downgradient eddy potential vorticity fluxes in interior layers: potential temperature is a conserved quantity whose variance tends to cascade irreversibly to small scales where it is dissipated. Direct heat exchange with the surface also damps the temperature variance. As implied by (6), the poleward eddy potential temperature flux leads to equatorward mass fluxes in the surface layer.

If there is a mixed layer adjacent to the surface, within which $\partial\Theta/\partial p \approx 0$, the thickness $H_s$ is no longer well defined and the estimate [(6)] must be modified. The total mass flux in the surface layer can be decomposed into a sum of the mass flux within the mixed layer and the mass flux between the top of the mixed layer and the top of the surface layer. Let $p_m$, with $p_m > p_I$, denote the pressure at the top of the mixed layer. Continuing to neglect the vertical structure of the geostrophic flow and noting that $\Theta_s$ is the potential temperature both at the top of the mixed layer and at the surface, we can approximate the surface layer mass flux as

$$\begin{aligned} &g^{-1}\overline{\tilde{v}_g(p_s - p_m)} + g^{-1}\overline{\tilde{v}_g(p_m - p_I)} \\ &\quad \approx g^{-1}\overline{\tilde{v}_g(p_s - p_m)} - H_m\overline{\tilde{v}_g\Theta_s'}, \end{aligned} \tag{7}$$

where $H_m$ is the value of the thickness $H$ above the mixed layer. If we can ignore the correlations between the depth $p_s - p_m$ of the mixed layer and the vertically averaged geostrophic flow $\tilde{v}_g$, then the first term on the right-hand side of (7) is negligible and (7) takes a form analogous to (6), with $H_s$ replaced by $H_m$. The mass flux in the surface layer is still proportional to the eddy heat flux near the surface, with a constant of proportionality dependent on the static stability *above* the mixed layer. If correlations between the mixed layer depth and the meridional geostrophic flow are not negligible, then an additional term is present that is not directly related to the eddy heat flux.

### *b. The distribution of the mass flux within the surface layer*

While the presence of a mixed layer does not affect the total mass flux in the surface layer if its depth is not correlated with the meridional flow, its presence does affect the distribution of the mass flux among the isentropic layers that form the surface layer. In fact, it appears that the mixed layer is the key ingredient that results in the equatorward mass flux being concentrated within isentropic layers that have potential temperatures lower than the mean surface potential temperature. To make this case, we construct a kinematic model for the vertical distribution of the mass flux within the surface layer.

We first ignore the presence of a mixed layer and assume that the thickness $H$ of each infinitesimal isentropic layer is a time-independent constant whenever the layer is present. (This is equivalent to assuming that the static stability is a constant.) Additionally, we assume that the meridional velocity in the surface layer has negligible vertical structure and is proportional to the deviation $\Theta_s' = \Theta_s - \overline{\Theta}_s$ of the surface potential temperature $\Theta_s$ from its mean $\overline{\Theta}_s$, so that southerlies are always warm and northerlies are always cold; that is, $\tilde{v}_g = \alpha\Theta_s'$ with a positive constant $\alpha$. The instantaneous mass flux in a layer with potential temperature $\Theta$ is thus $\alpha H\Theta_s' = \alpha H(\Theta_s - \overline{\Theta}_s)$ whenever $\Theta$ is greater than the surface potential temperature $\Theta_s$, and zero otherwise. Now suppose that we are given the probability density $\pi(\Theta_s)$ of the surface potential temperature $\Theta_s$ (see Fig. 2), with

$$\int_{-\infty}^{\infty} \pi(\Theta_s)\, d\Theta_s = 1. \tag{8}$$

The average mass flux in the layer with potential temperature $\Theta$ is then

$$V(\Theta) = \alpha H \int_{-\infty}^{\Theta} (\Theta_s - \overline{\Theta}_s)\pi(\Theta_s)\, d\Theta_s. \tag{9}$$

If, as is roughly the case in Fig. 2, the probability distribution $\pi(\Theta_s)$ is symmetric about the mean surface potential temperature $\overline{\Theta}_s$, then the mass flux $V(\Theta)$ is also distributed symmetrically within the surface layer, with equal contributions from isentropic layers above and below the mean surface potential temperature. More specifically, if $\pi(\Theta_s)$ is a Gaussian, then $V(\Theta)$ is also a Gaussian. For $\Theta \ll \overline{\Theta}_s$, $V$ decays to zero because layers with potential temperatures much below the mean surface potential temperature are rarely present; for $\Theta_I \geq \Theta \gg \overline{\Theta}_s$, $V$ decays because the layer is often present, so that poleward and equatorward mass fluxes cancel. Such a symmetric distribution of the mass flux about the mean surface potential temperature does not fit the results in Fig. 2 very well.

To improve this descriptive model, we add a mixed layer of mass per unit area, $h = \Delta p_{\mathrm{mixed}}/g$. The mixed layer supplies a flux, $\tilde{v}_g h = \alpha(\Theta_s - \overline{\Theta}_s)h$, to the vertically integrated mass transport. With this additional contribution, the mass flux [(9)] now becomes

$$V(\Theta) = \alpha \int_{-\infty}^{\Theta} [H + h\delta(\Theta_s + \epsilon - \Theta)] \times (\Theta_s - \overline{\Theta}_s)\pi(\Theta_s)\, d\Theta_s. \quad (10)$$

(The small number $\epsilon$ is included to make it clear that the $\delta$-function lies within the limits of integration; after carrying out the integration, $\epsilon$ can be set to zero.) If $\pi(\Theta)$ is symmetric about the mean surface potential temperature $\overline{\Theta}_s$, the mixed layer contribution to the net mass flux, proportional to $(\Theta - \overline{\Theta}_s)\pi(\Theta)$, is antisymmetric about $\overline{\Theta}_s$, with equatorward flow in layers below the mean and poleward flow in layers above the mean.

We neglect, as above, the vertical structure of the geostrophic flow in the surface layer, so that the ratio of the mass flux in the mixed layer to the mass flux in the remainder of the surface layer is equal to the ratio of the masses per unit area in these two parts of the surface layer. The ratio of the contribution from the mixed layer $\delta$-function to that of the first term in (10) is hence given by

$$\Delta p_{\text{mixed}}/\Delta p, \quad (11)$$

where the pressure difference $\Delta p$ is the depth of the surface layer portion lying above the mixed layer. Since the surface layer corresponds to the layer of frequently interrupted isentropes, the mass per unit area in the surface layer portion above the mixed layer can be estimated as $\Delta p/g \approx H\Delta\Theta$, where $\Delta\Theta$ is the standard deviation of the surface potential temperature.

A standard deviation of $\Delta\Theta = 10$ K and a static stability of 3 K km$^{-1}$ results in a surface layer depth of roughly 3 km, which is of the same order as, but substantially larger than, typical mixed layer depths. This is consistent with our GCM analysis. The mixed layer contribution is not so large that the mass fluxes change sign and become poleward everywhere in the upper half of the surface layer, as they would if the mixed layer term were dominant. Yet it is large enough to skew the mass flux distribution toward potential temperatures below the average surface potential temperature. We have thus arrived at a qualitative explanation for the distribution of the mass transport in Fig. 2.

## 3. The transformed Eulerian mean circulation

Consider a Boussinesq fluid for simplicity. The zonal mean potential temperature equation reads

$$\partial_t\overline{\Theta} + \overline{v}\partial_y\overline{\Theta} + \overline{w}\partial_z\overline{\Theta} + \partial_y\overline{v'\Theta'} + \partial_z\overline{w'\Theta'} = \overline{Q}, \quad (12)$$

where $Q$ represents all diabatic effects. The overbar now indicates a zonal average along a surface of constant height $z$. We can write the mean meridional circulation in terms of a streamfunction, $\Psi$, such that

$$\overline{v} = \partial_z\Psi, \qquad \overline{w} = -\partial_y\Psi, \quad (13)$$

with $\Psi = 0$ at the ground.

Making the standard change of variables (Andrews and McIntyre 1976),

$$\Psi_* \equiv \Psi + \Psi_e, \quad (14)$$

with

$$\Psi_e \equiv -\frac{\overline{v'\Theta'}}{\partial_z\overline{\Theta}}, \quad (15)$$

and defining $(\overline{v}_*, \overline{w}_*)$ to be the mean flow associated with $\Psi_*$, the potential temperature equation can be rearranged, without approximation, into its transformed Eulerian mean form,

$$\partial_t\overline{\Theta} + \overline{v}_*\partial_y\overline{\Theta} + \overline{w}_*\partial_z\overline{\Theta} + \partial_z T_* = \overline{Q}, \quad (16)$$

where

$$T_* \equiv \overline{v'\Theta'}\frac{\partial_y\overline{\Theta}}{\partial_z\overline{\Theta}} + \overline{w'\Theta'}. \quad (17)$$

The eddy forcing $\partial_z T_*$ and the horizontal advection $\overline{v}_*\partial_y\overline{\Theta}$ are ignored in quasigeostrophic theory. In what follows, we will argue why this is generally not a good approximation near the earth's surface.

Since the horizontal velocity in the boundary layer increases only logarithmically with height, eddy winds at a height of a few meters above the surface are of similar magnitude as the interior, geostrophic winds. In midlatitudes, the eddy heat flux contribution [(15)] to the residual streamfunction therefore remains significant down to very low altitudes. The poleward eddy heat flux at the top of the planetary boundary layer is somewhat larger than that at an altitude of a few meters, causing there to be some equatorward return flow in the residual circulation. According to Oort and Rasmusson (1971), however, the magnitude of the midlatitude eddy heat flux at 1000 mb still is roughly one-half of its maximum value near 850 mb. Yet the residual streamfunction has to vanish at the ground, which entails that a large part of the equatorward return flow occurs in a thin layer immediately above the surface. In this layer, the residual circulation transports potential temperature equatorward through the advection term $\overline{v}_*\partial_y\overline{\Theta}$ in (16). Since the layer in which the advection occurs is thin, the resulting cooling rate is large and must be balanced by a strong warming due to the term $\partial_z T_*$ in (16). Unlike anything appearing in the untransformed Eq. (12), this way of reorganizing the heat balance thus results in a large cancellation of terms near the surface.

This shortcoming of the conventional transformed Eulerian mean formulation becomes manifest in models that use a geophysical drag law boundary condition in place of a no-slip condition. Variables at the surface of such a model are intended to mimic those actually occuring at some reference height (say 10 m). Since the eddy heat flux at the surface of such a model is nonzero, the residual circulation $\Psi_*$ does not vanish at the ground. Streamlines of the residual circulation intersect the surface, and we speak of the circulation as being closed by a "$\delta$-function mass flux" at the ground.

Another drawback of the transformed balance [(16)] near the surface is that its computation requires division by the static stability $\partial_z\overline{\Theta}$. In a turbulent mixed layer near the ground, the static stability approaches zero, which may lead to a poorly defined residual circulation.

Andrews and McIntyre (1978) note that singling out the horizontal direction in the eddy flux component of the residual circulation [(15)] is arbitrary; residual circulations can be defined with eddy contributions proportional to the eddy heat flux in an arbitrary direction in the meridional-vertical plane. This freedom in the definition of the residual circulation can be exploited to obtain a residual streamfunction that represents the near-surface mass flux in a more realistic manner than does the conventional residual streamfunction. With a term involving the vertical eddy heat flux in place of the horizontal eddy heat flux in (15), the residual circulation can be redefined as

$$\Psi_*^\dagger \equiv \Psi + \Psi_e^\dagger, \tag{18}$$

with

$$\Psi_e^\dagger \equiv \frac{\overline{w'\Theta'}}{\partial_y\overline{\Theta}}. \tag{19}$$

Since $w'$ vanishes at the boundary, this modified circulation also vanishes at the boundary, even in models with a drag law boundary condition. The corresponding transformed mean equation reads

$$\partial_t\overline{\Theta} + \overline{v}_*^\dagger\partial_y\overline{\Theta} + \overline{w}_*^\dagger\partial_z\overline{\Theta} + \partial_y T_*^\dagger = \overline{Q}, \tag{20}$$

where

$$T_*^\dagger \equiv \overline{w'\Theta'}\frac{\partial_z\overline{\Theta}}{\partial_y\overline{\Theta}} + \overline{v'\Theta'}. \tag{21}$$

In terms of the *mixing slope S* and the *slope of the mean isentropes I,*

$$S \equiv \frac{\overline{w'\Theta'}}{\overline{v'\Theta'}}, \qquad I \equiv -\frac{\partial_y\overline{\Theta}}{\partial_z\overline{\Theta}}, \tag{22}$$

the relation between the eddy parts of the conventional and the modified residual streamfunction can be expressed as

$$\Psi_e^\dagger \equiv (S/I)\Psi_e, \tag{23}$$

To the extent that the interior of the atmosphere is adiabatic, we expect $S \approx I$ and, therefore, $\Psi_e \approx \Psi_e^\dagger$ and $\Psi_* \approx \Psi_*^\dagger$. In a statistically steady state, $S \approx I$ implies that the left-hand side of the eddy potential temperature variance budget

$$\begin{aligned}\partial_t\left(\overline{\Theta'^2}/2\right) &+ \overline{v'\Theta'}\partial_y\overline{\Theta} + \overline{w'\Theta'}\partial_z\overline{\Theta} \\ &= \overline{\Theta'Q'} - (\overline{v}\partial_y + \overline{w}\partial_z)\left(\overline{\Theta'^2}/2\right) - \partial_y\left(\overline{v'\Theta'^2}/2\right) \\ &\quad - \partial_z\left(\overline{w'\Theta'^2}/2\right)\end{aligned} \tag{24}$$

is approximately zero. Hence, assuming $S \approx I$ is tantamount to neglecting (i) the advection of variance by the mean meridional circulation $(\overline{v}, \overline{w})$ and by the eddies themselves and (ii) the production or destruction of eddy variance through diabatic effects. Because of latent heat release, however, diabatic variance modification is not negligible in the atmosphere; thus, the mixing slope is often subtantially larger than the slope of mean isentropes. If the atmosphere were more adiabatic, one would expect that either $\overline{v}_*$ or $\overline{v}_*^\dagger$, when multiplied by the mean thickness $\overline{H} = -g^{-1}\overline{\partial p/\partial\Theta}$, would be a good approximation of the isentropic mass flux per unit potential temperature, which was considered in section 2. A full discussion of this approximate relationship in the oceanic context can be found in McIntosh and McDougall (1996) and McDougall and McIntosh (1996).

Near the surface, if the horizontal heat flux remains significant, the damping of temperature variance must be such as to create a mixing slope that is closer to the horizontal than the isentropic slope. In this region, $\Psi_*$ and $\Psi_*^\dagger$ differ, but neither is an approximation of the isentropic mass transport streamfunction. If we assume that $S \approx I$ at some height, $Z$, above the planetary boundary layer, then the total transport due to eddies below height $Z$ is

$$-\left.\frac{\overline{v'\Theta'}}{\partial_z\overline{\Theta}}\right|_Z \tag{25}$$

with either definition of the residual circulation. Therefore, as far as an intuitive connection with the isentropic mass transport is concerned, there seems to be no particular advantage in using $\Psi_*$ rather than $\Psi_*^\dagger$ as the residual circulation. However, the modified circulation $\Psi_*^\dagger$ has the distinct advantage of providing an equatorward return flow that is smoothly distributed through the planetary boundary layer, the precise form of the return flow depending on the vertical structure of the mixing slope $S$. Also, the cooling due to equatorward flow in the modified circulation is of a more modest size. The term $\partial_y T_*^\dagger$ is typically much smaller in amplitude and more slowly varying than $\partial_z T_*$. We can rewrite $T_*^\dagger$ in the form

$$T_*^\dagger \equiv \overline{v'\Theta'}(1 - S/I). \tag{26}$$

If the interior is approximately adiabatic in the sense that $S \approx I$, then this term contributes only in the planetary boundary layer, where the eddy heat flux convergence is reduced, as one moves away from the surface, by the factor $(1 - S/I)$.

Following Treguier et al. (1997), we can estimate the relative magnitudes of the terms $\overline{v}_*^\dagger\partial_y\overline{\Theta}$ and $\partial_y T_*^\dagger$ in (20). In midlatitudes the eddy contribution [(25)] dominates the total mass transport in the surface layer, so an order-of-magnitude estimate for the total transport is given by

$$\delta\overline{v}_*^\dagger \approx \frac{\overline{v'\Theta'}}{\partial_z\overline{\Theta}}, \tag{27}$$

where $\delta$ is the depth of the layer containing the return flow. Near the surface, where $S/I$ is small, the term involving the eddy heat flux convergence scales like

$$\partial_y T_*^\dagger \approx \overline{v'\Theta'}/L, \tag{28}$$

where $L$ is the scale of variation of the eddy heat flux. Therefore, the advection $\overline{v}_*^\dagger \partial_y \overline{\Theta}$ dominates $\partial_y T_*^\dagger$ if

$$LI/\delta \gg 1, \tag{29}$$

where the isentropic slope $I$ is evaluated at the top of the surface layer. In the atmosphere, we typically have $LI/H_T \approx 1$, where $H_T$ is the depth of the troposphere, so (29) should be well obeyed. The scaling [(29)] implies that the mean cooling due to equatorward mass flow near the surface is typically balanced by diabatic air mass modification rather than convergence of horizontal eddy heat fluxes.

Note also that the computation of $\Psi_*^\dagger$ does not require division by the static stability, so the modified streamfunction remains well defined even within a well-mixed boundary layer. Of course, we now have a problem in regions where the meridional gradient $\partial_y \overline{\Theta}$ vanishes or is small. While this is not an issue in midlatitudes near the surface, it does occur in the Tropics and near the tropopause. One way to avoid this problem is to define the residual streamfunction to be whichever of $\Psi_*$ and $\Psi_*^\dagger$ has the smaller magnitude. Equivalently, we multiply the eddy part $\Psi_e$ of the conventional residual streamfunction by $S/I$ where $S/I < 1$ and leave $\Psi_e$ unmodified where $S/I > 1$. Despite its arbitrariness, this hybrid residual circulation is as closely related to the isentropic mass transport as is $\Psi_*$ or $\Psi_*^\dagger$ individually. One could also choose to make the transition from $\Psi_*$ to $\Psi_*^\dagger$ more smoothly as a function of $S/I$, so as to avoid discontinuous derivatives. Note that if an eddy flux closure theory provides an estimate of $S/I$ that remains bounded and in which, in particular, the numerator $S$ vanishes wherever the denominator $I$ vanishes, there would be no difficulty with the use of $\Psi_*^\dagger$ throughout the domain.

Figure 3a shows the conventional residual streamfunction and Fig. 3b the modified streamfunction, both computed as time mean streamfunctions from the same GCM for which the isentropic overturning was displayed in Fig. 1. In the time mean streamfunctions, the eddy terms (15) and (19) are defined as deviations from the temporal and zonal mean so as to include the fluxes due to stationary eddies. In order to compute the streamfunctions from the $\sigma$-coordinate GCM, the Boussinesq equations used in the preceding analysis have been generalized to $\sigma$ coordinates. The overbar must now be interpreted as denoting a zonal and temporal average on a $\sigma$ surface, in which the quantity being averaged is weighted by the instantaneous surface pressure. In Fig. 3b, we have used the construction described in the preceding paragraph, in which the time average of the conventional residual circulation $\Psi_*$ is replaced by the time average of the modified residual circulation $\Psi_*^\dagger$ where

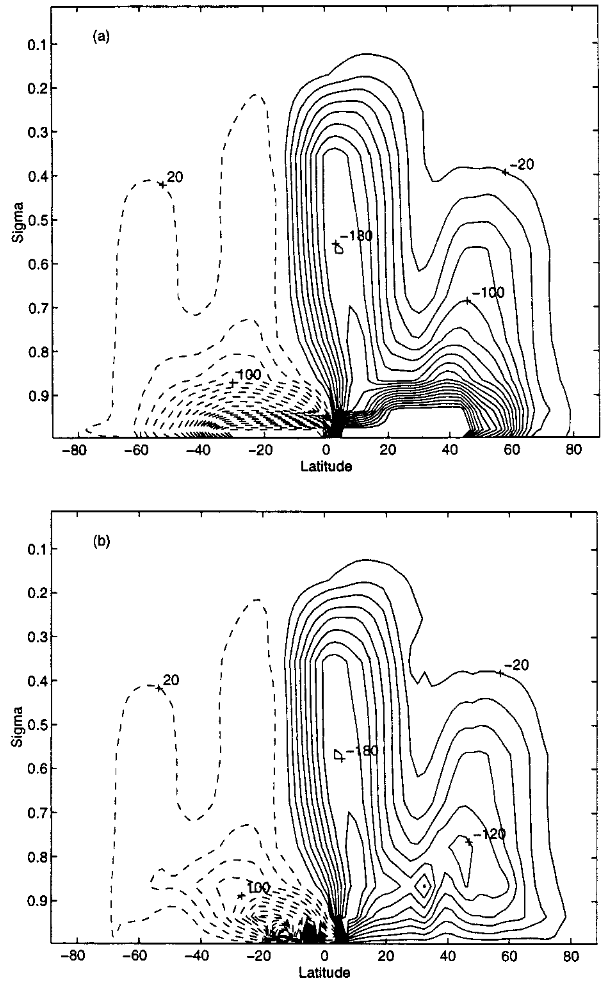


FIG. 3. Residual mean meridional streamfunctions [$10^9$ kg s$^{-1}$] corresponding to the isentropic mass circulation in Fig. 1: (a) conventional transformed Eulerian mean streamfunction $\Psi_*$ and (b) modified residual streamfunction (hybrid of $\Psi_*$ and $\Psi_*^\dagger$).

$S/I < 1$. As a consequence, the circulation is modified mainly in the lower troposphere, in the region pictured in Fig. 4. Evidently, the modified circulation in Fig. 3b provides a more realistic representation of the return flow near the surface.

## 4. Surface inversions

Whereas wintertime inversions near the surface are commonly attributed to strong radiative cooling, experiments with the idealized dry GCM described by Held and Suarez (1994, hereafter HS94) suggest a different interpretation. In this model, the atmosphere is forced by Newtonian relaxation to a prescribed temperature field. Here we modify the thermal damping, using a long damping time of 40 days ($k_a = k_s = 1/40$ day$^{-1}$ in the notation of HS94) everywhere in the model domain. The model is run at T42 horizontal resolution

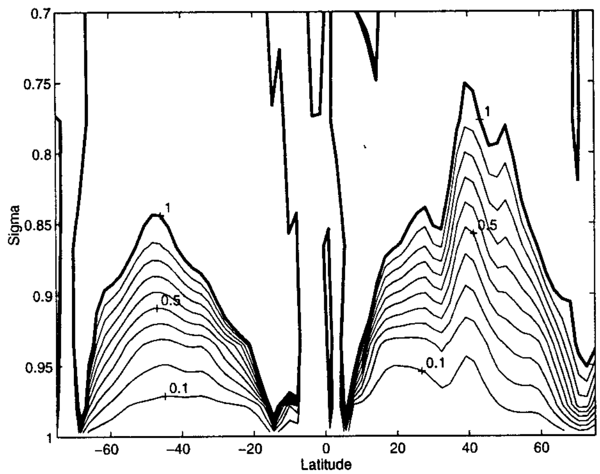


FIG. 4. Ratio $S/I$ of the mixing slope $S$ over the isentropic slope $I$. Contours are shown only where $S/I < 1$.

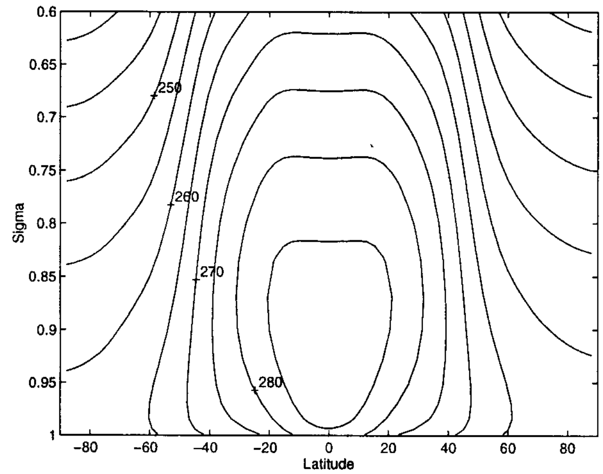


FIG. 5. 800-day average of the zonal mean temperature [K] in an idealized GCM.

with 45 $\sigma$ levels in the vertical, unequally spaced so as to better resolve the flow near the surface.

Figure 5 shows the resulting climatic temperature distribution. A strong inversion forms near the surface, not just near the poles, but at all latitudes. HS94 increase the strength of the thermal damping near the surface to prevent the formation of these unrealistic surface inversions. The scaling arguments of the previous section suggest that we can interpret the inversions as resulting from cold air advection by the modified residual circulation. There is no tendency for the inversions to be weaker in midlatitudes, where the Eulerian mean flow near the surface is poleward; this reinforces the value of thinking in terms of a transformed temperature equation. (Since the static stability near the surface is very high in this integration, the value of the conventional residual streamfunction $\Psi_*$ at the boundary is nearly zero, so the conventional streamfunction $\Psi_*$ and the modified streamfunction $\Psi_*^\dagger$ are almost identical. Therefore, the surface inversion can also be interpreted as resulting from cold air advection by the conventional residual circulation.)

## 5. Implications for general circulation theory

The preceding discussion has underscored the value of the distinction between isentropic layers that are rarely interrupted by the surface and those that are often interrupted by the surface. Particularly in midlatitudes, this conceptual decomposition of the tropospheric dynamics into a near-surface and an interior component suggests a general circulation theory analogously divided into two intertwined parts.

First, *given the equatorward mass flux in the surface layer,* one could try to develop a theory describing how the tropospheric temperatures adjust so as to provide the diabatic fluxes and poleward flow required to close the circulation. One can think of the horizontal divergence of the surface mass flux as resulting in mass sources or sinks for different isentropic layers, which, in a steady state, must be balanced in the interior. Since the eddy PV flux distribution in the interior essentially determines the mass flux, this part of the midlatitude circulation problem effectively reduces to developing a theory for the vertical structure of the eddy PV flux in the interior, given the vertical integral of this flux. This part of the theory must predict the structure of the extratropical tropopause and the troposheric static stability. In theories for the wind-driven oceanic thermocline, Ekman pumping near the surface similarly drives the interior flow. However, a fundamental difference is that the surface mass flux in the atmosphere is dominated by the eddy contribution, whereas over most of the ocean surface the mean flow presumably dominates. [But see Marshall (1997) and Treguier et al. (1997) for a discussion of analogous horizontal eddy mass fluxes near the surface in the oceanic context.]

Second, one requires a theory for the surface mass flux itself. Ignoring the eddy momentum fluxes as a first approximation, one needs to establish theories for the eddy heat fluxes near the surface and the static stability above the planetary boundary layer. The static stability would presumably be determined by the interior tropospheric dynamics. The near-surface eddy heat fluxes are closely tied to the near-surface temperature gradients. Since these temperature gradients are strongly maintained by large-scale thermal forcing, eddies typically interact with a well-defined environmental gradient that varies on larger scales than the eddies themselves, thus rendering diffusive eddy closures a viable option near the surface. The diffusivity in such a closure theory would likely depend on the surface thermal gradients and on aspects of the internal tropospheric structure, such as the static stability and the depth of the troposphere.

## 6. Concluding remarks

The isentropic mean meridional mass transport and the residual mean meridional circulation have been computed from atmospheric and model data on numerous occasions; they also have been explicitly compared (e.g., Karoly et al. 1997). Whereas other studies have traditionally focused on the poleward flow in the upper-tropospheric branches of these overturning circulations, we have examined the equatorward flow in the surface branches.

The isentropic perspective is in many respects the more straightforward. Consistent with expectations based on the zonal momentum budget, one finds value in the distinction between isentropic layers that are typically uninterrupted and those that are typically interrupted by the surface. The uninterrupted layers contain the poleward flow; the interrupted layers contain the equatorward flow, most of which occurs in isentropic layers with potential temperatures that are lower than the mean surface potential temperature.

In the conventional transformed Eulerian mean circulation, the near-surface branch of the mean meridional mass transport circulation is represented in an unrealistic manner: some streamlines of the residual streamfunction intersect the ground unless a no-slip condition is imposed. We introduce a modified residual circulation whose streamlines do close above the surface, so that the near-surface equatorward flow can be interpreted as advecting the temperature field. This modified residual circulation approximates the isentropic overturning in the same sense as does the conventional residual circulation.

Experiments with an idealized GCM, which has only weak thermal damping, suggest that it may be misleading to regard surface inversions as a consequence of local radiative stabilization alone. The fact that the idealized GCM produces a strong surface inversion at all latitudes indicates that mean equatorward transport of cold air may also be important. Support for this view comes from scaling arguments, which show that cold air advection by the modified residual circulation cannot be balanced by horizontal mixing with warmer air along the surface; rather, diabatic airmass modification associated with vertical heat fluxes is required.

The near-surface equatorward mass transport can be understood as determined by the eddy heat fluxes near the surface and the static stability of the troposphere. The poleward mass flux in the upper troposphere is essentially determined by the vertical structure of the potential vorticity flux. This distinction between the dominant force balances in the surface layer and in the interior suggests a general circulation theory that is analogously divided into two intertwined parts: (i) a theory (diffusive, perhaps) for the surface mass flux, the convergence or divergence of which provides an inflow into or outflow from the tropospheric interior; and (ii) a theory for how the tropospheric temperatures adjust so as to be consistent with the diabatic fluxes and poleward flow required to close the circulation.

*Acknowledgments*. We wish to express our thanks to Robert Hallberg, Paul Kushner, and an anonymous reviewer, whose comments on an earlier draft of this paper led to significant improvements in the final version.

## APPENDIX

### Form Drag and the Geostrophic Mass Flux

Consider a surface of constant potential temperature $\Theta$, which may intersect the ground at the latitude in question. The geostrophic, hydrostatic, and thermal wind relations in isentropic coordinates are

$$fv_g \equiv \partial_x M, \qquad \partial_\Theta M = \Pi, \qquad f\partial_\Theta v_g = \partial_x \Pi, \quad \text{(A1)}$$

where $\Pi \equiv c_p(p/p_*)^\kappa$ is the Exner function and $M \equiv c_pT + gz = \Theta\Pi + gz$ the Montgomery streamfunction. Where the surface potential temperature is higher than $\Theta$, we use the artifice of defining $p(\Theta)$ and $z(\Theta)$ to equal the surface pressure and height, respectively. This convention assures that the thickness $H$ vanishes in the ''massless'' regions on the $\Theta$ surface that lie ''below the ground.'' The geostrophic velocity $v_g(\Theta)$ in these regions is determined by definition (A1) so that the thermal wind equation is still satisfied. The geostrophic meridional mass flux per unit potential temperature is $Hv_g = -g^{-1}v_g\partial_\Theta p$, with

$$v_g\partial_\Theta p = \partial_\Theta(v_g p) - p\partial_x\Pi/f \quad \text{(A2)}$$

from the thermal wind equation. Averaging over $x$ along the isentrope $\Theta$ and using the convention described above for the massless part of the domain, we have (since $\Pi$ is a function of $p$)

$$\overline{v_g\partial_\Theta p}^\Theta = \partial_\Theta\overline{(v_g p)}^\Theta. \quad \text{(A3)}$$

But geostrophy implies

$$fv_g p = p\partial_x M = p\partial_x(\Theta\Pi + gz) = \Theta p\partial_x\Pi + gp\partial_x z, \quad \text{(A4)}$$

so that, averaging over $x$ as before,

$$\overline{v_g p}^\Theta = (g/f)\overline{p\partial_x z}^\Theta. \quad \text{(A5)}$$

Therefore,

$$(f/g)\overline{v_g\partial_\Theta p}^\Theta = \partial_\Theta\overline{(p\partial_x z)}^\Theta. \quad \text{(A6)}$$

Within the hydrostatic approximation, the Coriolis force due to the geostrophic mass flux (the left-hand side of the above equation) thus balances the form drag (the right-hand side) exactly.

## REFERENCES

Andrews, D. G., 1983: A finite-amplitude Eliassen–Palm theorem in isentropic coordinates. *J. Atmos. Sci.*, **40,** 1877–1883.

——, and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: The generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.,* **33,** 2031–2048.

——, and M. E. McIntyre, 1978: Generalized Eliassen–Palm and Charney–Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres. *J. Atmos. Sci.,* **35,** 175–185.

——, J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. International Geophysical Series, Vol. 40, Academic Press, 489 pp.

Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, 1980: Eliassen–Palm cross sections for the troposphere. *J. Atmos. Sci.,* **37,** 2600–2616.

Held, I. M., and B. J. Hoskins, 1985: Large-scale eddies and the general circulation of the troposphere. *Advances in Geophysics,* Vol. 28A, Academic Press, 3–31.

——, and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.,* **75,** 1825–1830.

Johnson, D. R., 1989: The forcing and maintenance of global monsoonal circulations: An isentropic analysis. *Advances in Geophysics,* Vol. 31, Academic Press, 43–304.

Karoly, D. J., P. C. McIntosh, P. Berrisford, T. J. McDougall, and A. C. Hirst, 1997: Similarities of the Deacon cell in the Southern Ocean and Ferrel cells in the atmosphere. *Quart. J. Roy. Meteor. Soc.,* **123,** 519–526.

Marshall, D., 1997: Subduction of water masses in an eddying ocean. *J. Mar. Res.,* **55,** 201–222.

McDougall, T. J., and P. C. McIntosh, 1996: The temporal-residual-mean velocity. Part I: Derivation and the scalar conservation equations. *J. Phys. Oceanogr.,* **26,** 2653–2665.

McIntosh, P. C., and T. J. McDougall, 1996: Isopycnal averaging and the residual mean circulation. *J. Phys. Oceanogr.,* **26,** 1655–1660.

Oort, A. H., and E. M. Rasmusson, 1971: Atmospheric circulation statistics. Prof. Paper 5, U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, 323 pp.

Rhines, P. B., and W. R. Holland, 1979: A theoretical discussion of eddy-driven mean flows. *Dyn. Atmos. Oceans,* **3,** 289–325.

Treguier, A. M., I. M. Held, and V. D. Larichev, 1997: On the parameterization of quasi-geostrophic eddies in primitive equation ocean models. *J. Phys. Oceanogr.,* **27,** 567–580.