STRATOSPHERE-TROPOSPHERE EXCHANGE

James R. Holton[1]
Peter H. Haynes[2] and Michael E. McIntyre[2]
Anne R. Douglass[3] and Richard B. Rood[3]
Leonhard Pfister[4]

Abstract. In the past, studies of stratosphere-troposphere exchange of mass and chemical species have mainly emphasized the synoptic- and small-scale mechanisms of exchange. This review, however, includes also the global-scale aspects of exchange, such as the transport across an isentropic surface (potential temperature about 380 K) that in the tropics lies just above the tropopause, near the 100-hPa pressure level. Such a surface divides the stratosphere into an "overworld" and an extratropical "lowermost stratosphere" that for transport purposes need to be sharply distinguished. This approach places stratosphere-troposphere exchange in the framework of the general circulation and helps to clarify the roles of the different mechanisms involved and the interplay between large and small scales. The role of waves and eddies in the extratropical overworld is emphasized. There, wave-induced forces drive a kind of global-scale extratropical "fluid-dynamical suction pump," which withdraws air upward and poleward from the tropical lower stratosphere and pushes it poleward and downward into the extratropical troposphere. The resulting global-scale circulation drives the stratosphere away from radiative equilibrium conditions. Wave-induced forces may be considered to exert a nonlocal control, mainly downward in the extratropics but reaching laterally into the tropics, over the transport of mass across lower stratospheric isentropic surfaces. This mass transport is for many purposes a useful measure of global-scale stratosphere-troposphere exchange, especially on seasonal or longer timescales. Because the strongest wave-induced forces occur in the northern hemisphere winter season, the exchange rate is also a maximum at that season. The global exchange rate is not determined by details of near-tropopause phenomena such as penetrative cumulus convection or small-scale mixing associated with upper level fronts and cyclones. These smaller-scale processes must be considered, however, in order to understand the finer details of exchange. Moist convection appears to play an important role in the tropics in accounting for the extreme dehydration of air entering the stratosphere. Stratospheric air finds its way back into the troposphere through a vast variety of irreversible eddy exchange phenomena, including tropopause folding and the formation of so-called tropical upper tropospheric troughs and consequent irreversible exchange. General circulation models are able to simulate the mean global-scale mass exchange and its seasonal cycle but are not able to properly resolve the tropical dehydration process. Two-dimensional (height-latitude) models commonly used for assessment of human impact on the ozone layer include representation of stratosphere-troposphere exchange that is adequate to allow reasonable simulation of photochemical processes occurring in the overworld. However, for assessing changes in the lowermost stratosphere, the strong longitudinal asymmetries in stratosphere-troposphere exchange render current two-dimensional models inadequate. Either current transport parameterizations must be improved, or else, more likely, such changes can be adequately assessed only by three-dimensional models.

1. INTRODUCTION

Dynamical, chemical, and radiative coupling between the stratosphere and troposphere are among the many important processes that must be understood for prediction of global change. Of special significance is the transport of trace chemical species, natural and anthropogenic, between the stratosphere and the troposphere. For instance, anthropogenic species transported from the troposphere into the stratosphere initiate much of the chemistry responsible for stratospheric ozone depletion, according to a large body of evidence [e.g., *World Meteorological Organization (WMO)*, 1995]. Conversely, downward transport from the stratosphere not only constitutes the main removal mechanism for many stratospheric species, including those involved in ozone depletion, but also represents a significant input of ozone and other reactive species into the tropospheric chemical system

[1]Department of Atmospheric Sciences, University of Washington, Seattle.
[2]Centre for Atmospheric Science at the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England.
[3]Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland.
[4]NASA Ames Research Center, Moffett Field, California.

Reviews of Geophysics, 33, 4 / *November 1995*
pages 403–439
8755-1209/95/95RG-02097$15.00
Paper number 95RG02097

[e.g., *Levy et al.*, 1980]. Chemical effects from stratosphere-troposphere exchange (STE) can, in turn, influence the radiative flux balance in the troposphere and lower stratosphere and can do so in more than one way [e.g., *Ramaswamy et al.*, 1992; *Toumi et al.*, 1994]. STE can therefore have a significant role in the radiative forcing of global climate change.

It is now widely appreciated that the dynamics of the troposphere and the stratosphere are, in principle, inseparable [e.g., *Hoskins et al.*, 1985]. Indeed, one might pose the question, Does it make any sense at all to distinguish between the troposphere and the stratosphere? If not, then STE would just be part of the overall pattern of transport, with no particular significance on its own. However, there are strong reasons why the distinction between troposphere and stratosphere will remain useful. One of them is that vertical transport of air and chemical species through the depth of the troposphere can occur on timescales as short as a few hours via moist convection and on timescales of days via baroclinic eddy motions in middle latitudes. By contrast, vertical transport through a similar altitude range in the stratosphere takes months, indeed a year or more in the lower stratosphere, and this vertical transport must be accompanied by radiative heating or cooling.

There are important implications therefore both for chemistry and for radiative flux balances. The difference between the vertical transport timescales in the stratosphere and troposphere is part of what lies behind, for instance, the rapid increase in ozone mixing ratio and the rapid decrease in water vapor mixing ratio with altitude observed just above the tropopause. We can often identify "tropospheric" and "stratospheric" air parcels by their differing chemical compositions. If only for this reason, it seems likely that we shall continue to distinguish between troposphere and stratosphere and to regard quantitative measures of the transport between them as an interesting aspect of transport.

The stratosphere and troposphere are sometimes regarded as isolated boxes and the exchange problem as merely one of estimating a single mass transport rate between the two boxes. However, the relatively long timescales for vertical transport within the stratosphere and the inhomogeneity of the stratospheric photochemical environment, as regards both actinic radiative fluxes and chemical tracer distributions, compel us to take a less simplified viewpoint. STE cannot, in reality, be thought of in terms of slow transport between two well-mixed boxes; single-number measures of exchange are therefore, by themselves, of limited utility. More useful measures of STE must concern not only the transport across the tropopause but also the rate at which tropospheric material is supplied to and removed from the regions in the stratosphere in which there are chemical sources and sinks, for whichever chemical species are of interest. It is necessary therefore to take into account the species' photochemical sensitivities at different altitudes and latitudes, and the global-scale circulation including the spatiotemporal structure of transport within the stratosphere. Such considerations will be given emphasis in this review, which will also try to elucidate the relation between the global-scale circulation and the processes, some of them on small scales, that may effect or affect transport across the tropopause itself.

What, in any case, do we mean by the tropopause? By convention of the World Meteorological Organization (WMO), the tropopause is defined as the lowest level at which the temperature lapse rate decreases to 2 K km^{-1} or less and the lapse rate averaged between this level and any level within the next 2 km does not exceed 2 K km^{-1}. Thus the tropopause is characterized by an increase in the static stability in moving from the troposphere to the stratosphere; this shows even in a monthly mean at fixed altitudes and latitudes (Figure 1). However, this conventional "thermal" definition of the tropopause obscures the fact that the tropopause often behaves, chemically speaking, as if it were a material surface to varying degrees of approximation. The quasi-material behavior can be rationalized by noting, first, that the tropical tropopause corresponds roughly to an isentropic or stratification surface (whose potential temperature $\theta \approx 380$ K in the annual mean) and, second, that the extratropical tropopause corresponds roughly to a surface of constant potential vorticity (PV), provided that the PV is defined in the conventional way (see glossary following section 9). With the conventional definition of PV, the extratropical tropopause is found from observation to be remarkably close to the 2-PVU potential vorticity surface, where PVU denotes the standard potential vorticity unit (1 PVU = 10^{-6} m^2 s^{-1} K kg^{-1}). The tropopause slopes downward and poleward, intersecting isentropic surfaces as suggested by the heavy curve in Figure 1, and is usually distinguished on each such intersecting isentropic surface by a band of sharp gradients of PV, with values typically near 2 PVU, whose instantaneous configuration may well be strongly deformed, indeed irreversibly deformed, by synoptic-scale baroclinic eddies and other extratropical weather-related disturbances.

Figure 2 and Plate 1 (also Plate 2 below) illustrate how large these deformations of the instantaneous tropopause can be on a typical intersecting isentrope, 320 K in this case. Figure 2a shows a coarse-grain view of the isentropic distribution of PV, estimated from a state-of-the-art operational weather analysis. Figure 2b shows the corresponding satellite image of the 5.7- to 7.1-μm water vapor absorption band, which is sensitive to total water vapor above about the 500-hPa surface. Darker areas correspond to regions of low water vapor and provide a finer-scale view of tropopause deformations that bring dry stratospheric

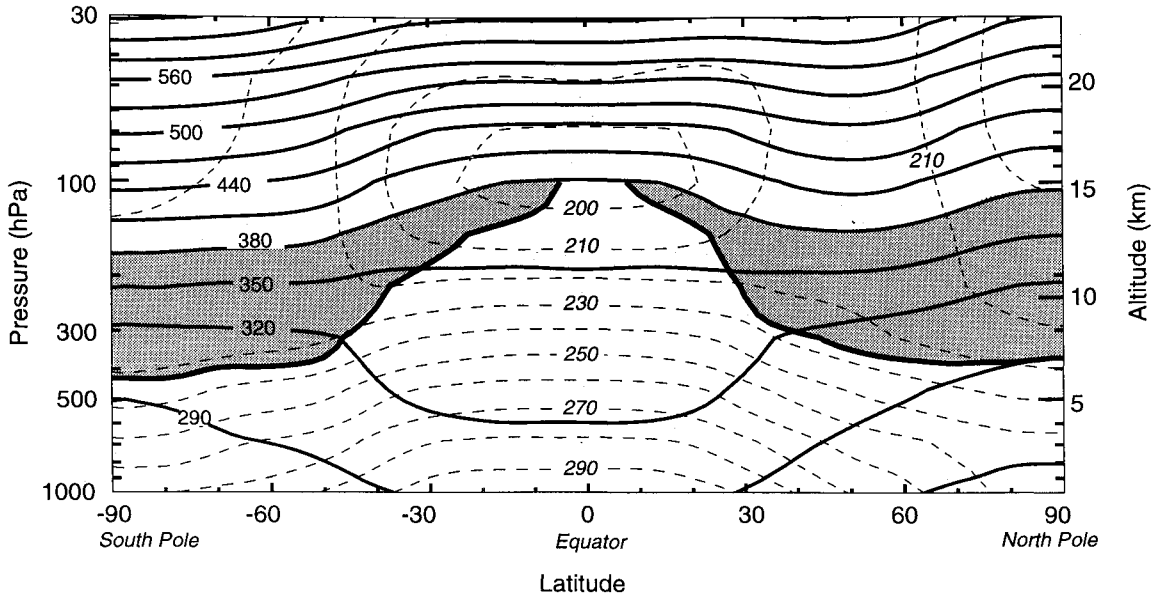

Figure 1. Latitude-altitude cross section for January 1993 showing longitudinally averaged potential temperature (solid contours) and temperature (dashed contours). The heavy solid contour (cut off at the 380-K isentrope) denotes the 2-PVU potential vorticity contour, which approximates the tropopause outside the tropics. Shaded areas denote the "lowermost stratosphere," whose isentropic or potential temperature surfaces span the tropopause. Data are from UKMO analyses [*Swinbank and O'Neill*, 1994]. In these data the tropical tropopause occurs near $\theta \approx 380$ K, which is somewhat higher than the mean tropical tropopause potential temperature ($\theta \approx 370$ K) determined from radiosonde data by *Reid and Gage* [1981]. Polar tropopauses are far more variable and less well defined than might be suggested here. Figure courtesy of C. Appenzeller.

air equatorward and downward than is available from the coarse-grain PV analysis alone. Plate 1 shows an even finer scale view of tropopause deformations as derived from an accurate simulation of an advected passive chemical tracer [*Appenzeller et al.*, 1995]. The tracer follows the instantaneous tropopause, defined here as the 2-PVU contour, fairly closely over 4 days. The coloring marks air that had PV values >1 PVU initially; most of this air, with initial PV values ≳2 PVU, stays in the stratosphere. Fine details apart, it is clear that the instantaneous tropopause on the 320-K isentrope is making north-south excursions of 3000 km or more; such excursions are commonly observed, especially near the ends of storm tracks, as in this example, and especially in what meteorologists call "blocking" or, more generally, "low zonal index" situations.

It is generally thought that the climatological-mean structure and position of the tropopause, indicated approximately in Figure 1, is determined both by radiative-convective adjustment (especially in the tropics where deep convection clamps tropospheric temperatures close to a moist adiabat [e.g., *Emanuel et al.*, 1994]) and also by the action of synoptic-scale baroclinic eddies in the extratropics. The latter idea fits well with the suggestion of *McIntyre and Palmer* [1984] that such eddies can sharpen the isentropic gradients of PV marking the extratropical tropopause, by mixing "PV substance" (see glossary) along the tropospheric portions of isentropic surfaces. Such mixing does, of course, involve irreversible deformation of PV contours on isentropic surfaces, as glimpsed in Figure 2 and Plate 1, and erodes or entrains stratospheric material into the troposphere. *Lindzen* [1994] argues from the notion of "baroclinic adjustment" that such PV mixing on isentropes spanning the troposphere should have a self-limiting character, tending not only to sharpen the tropopause but also to push it to higher altitudes and hence to exert a negative or stable feedback control over its position in middle and high latitudes. The idea that tropospheric baroclinic eddies influence extratropical tropopause height was first suggested by *Held* [1982]. However, a quantitative overall picture of how different processes combine to maintain the observed climatological-mean tropopause is still lacking.

Given the quasi-material nature of the tropopause, it follows that reversible deformations of its shape, however large and rapid, are of little significance to transport in themselves. As was already suggested, what is significant is the irreversible transport across the tropopause, more precisely, across some representative mean position of the tropopause, that may be associated, for instance, with the synoptic-scale baroclinic eddies or with smaller-scale processes exemplified in Figure 2 and Plate 1 or with global-scale

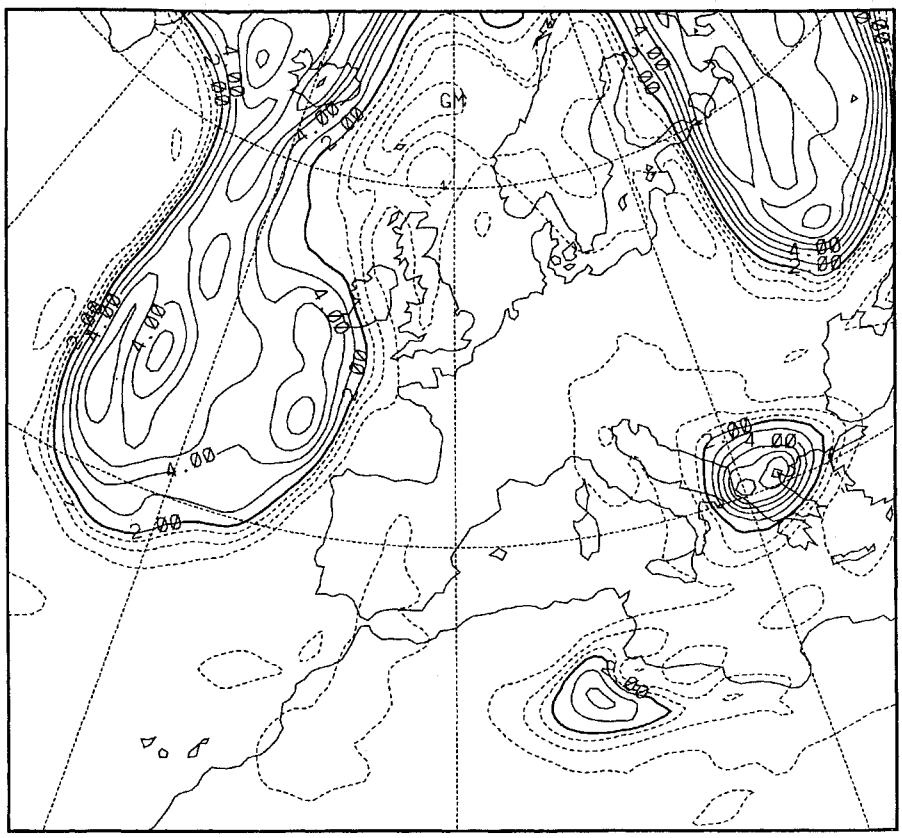

Figure 2a. Isentropic contours of PV on the 320-K surface for May 14, 1992, at 1200 UT, calculated from European Centre for Medium-Range Weather Forecasts (ECMWF) operational analyses. The instantaneous tropopause appears as the band of dotted contours (0.5, 1.0, and 1.5 PVU) and the first solid contour (2 PVU); contours for 3 and 4 PVU and higher are also shown solid. (From *Appenzeller et al.* [1995].)

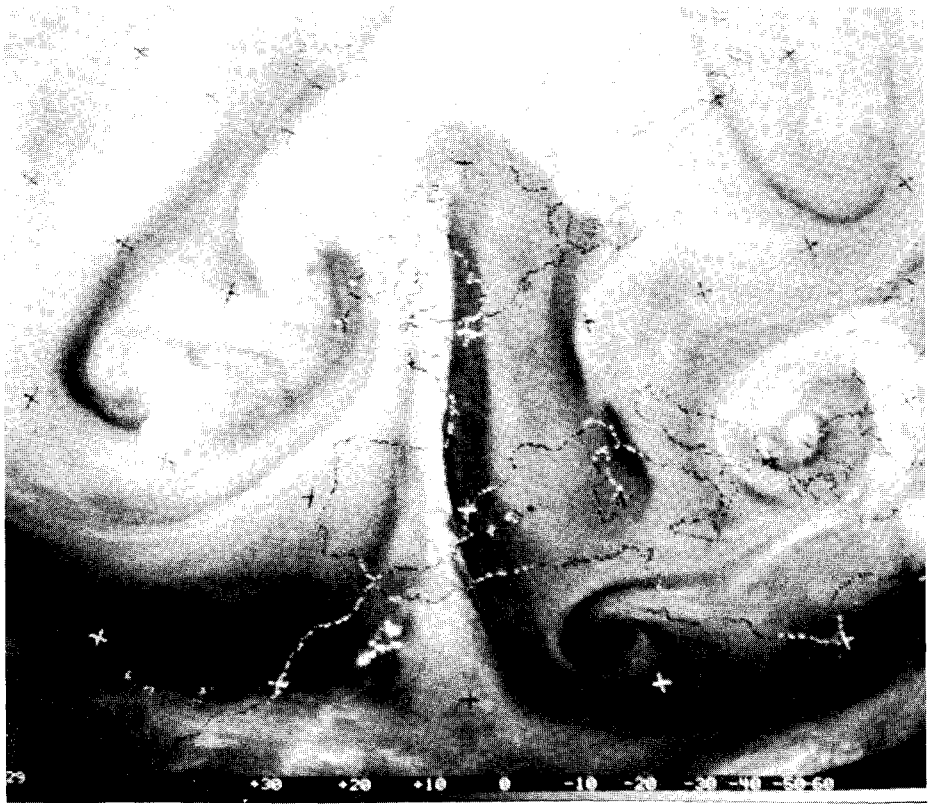

Figure 2b. Meteosat 5.7- to 7.1-μm water vapor image for the same time as Figure 2a. (From *Appenzeller et al.* [1995].)

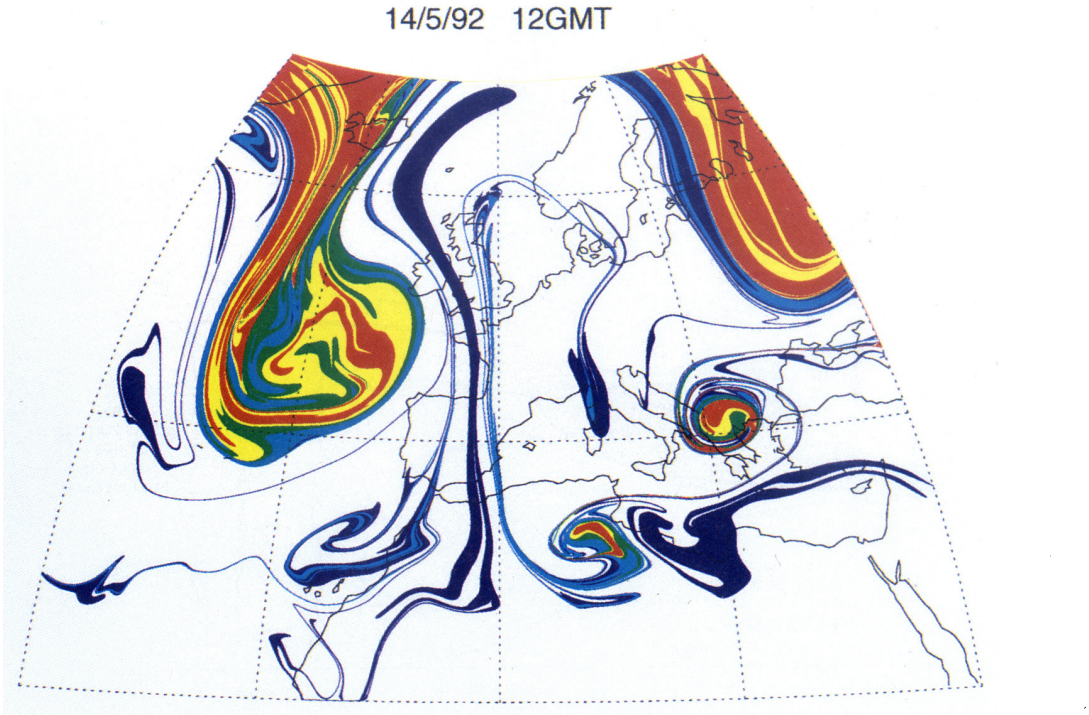

Plate 1. Passive tracer on 320-K isentrope for the same time as Figure 2, May 14, 1992, 1200 UT, estimated from ECMWF-analyzed winds using a high-precision Lagrangian tracer-advection algorithm ("contour advection") [see *Norton*, 1994; *Waugh and Plumb*, 1994]. The tracer contours were initialized 4 days earlier to coincide, at that time, with the 1-, 2-, 3-, 4-, and 5-PVU contours on a smoothed isentropic map of PV showing no small-scale structure. Values corresponding to 1 PVU or higher are colored. (From *Appenzeller et al.* [1995].)

diabatic ascent or descent. But how do events near the tropopause fit into a global-scale picture of general circulation and transport? The next section sketches what is involved and gives the plan of the rest of the review.

2. STE AND THE GLOBAL CIRCULATION, GENERAL APPROACH, AND PLAN OF THE REVIEW

Figure 3 is an attempt to summarize part of what is involved. It is useful to distinguish first of all between, on the one hand, transport along isentropic surfaces, of the kind illustrated in Plate 1, which may occur adiabatically (wavy arrows in Figure 3), and on the other hand transport across isentropic surfaces, which may require diabatic processes, including small-scale three-dimensionally turbulent processes. Since the tropopause intersects the isentropes, transport can occur in either way and is likely to occur in both ways. In the region of the upper troposphere and lower stratosphere consisting of isentropic surfaces that intersect the tropopause, air and chemical constituents can be irreversibly transported as adiabatic eddy motions lead to large latitudinal displacements of the tropopause, as in Figure 2 and Plate 1, followed by irreversible mixing on small scales. The darker shading in Figure 3 shows the region within the lower stratosphere most directly affected by these eddy transport effects. This region will be referred to as the "lowermost stratosphere" or, when clarity demands it, as the "lowermost extratropical stratosphere." It is also the stratospheric part of what *Hoskins* [1991] calls the "middleworld."

For chemical purposes especially, the lowermost stratosphere must be sharply distinguished from the tropospheric part of the middleworld, where moist convection strongly transports material across isentropic surfaces. Equally, the lowermost stratosphere must be distinguished from the rest of the stratosphere, it being the only part of the stratosphere accessible from the troposphere via transport along isentropic surfaces. The term "lowermost stratosphere" will help keep both distinctions clear.

Now air in the region where isentropic surfaces lie entirely in the stratosphere, referred to by Napier Shaw and Hoskins as the "overworld," cannot reach the troposphere without first slowly descending across isentropic surfaces, a process that must be accompanied by diabatic cooling. Conversely, air in the "underworld," where isentropic surfaces lie entirely in the troposphere, cannot reach the stratosphere without first rising across isentropic surfaces. The isentropic surface bounding the overworld and the lowermost

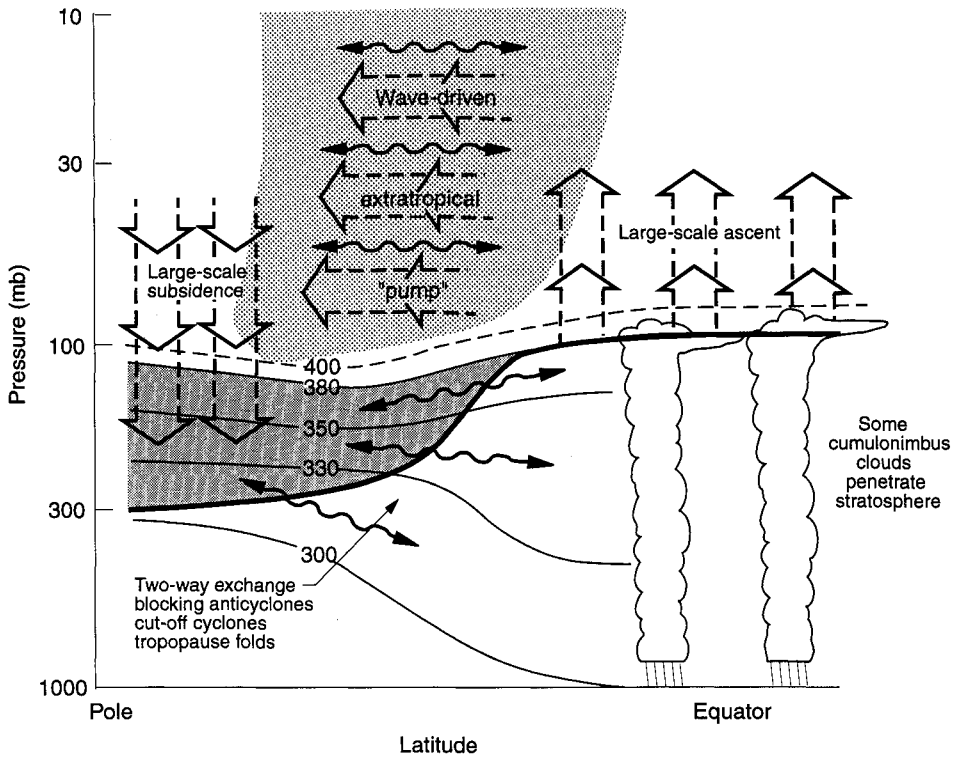

Figure 3. Dynamical aspects of stratosphere-troposphere exchange. The tropopause is shown by the thick line. Thin lines are isentropic or constant potential temperature surfaces labeled in kelvins. The heavily shaded region is the "lowermost stratosphere," where isentropic surfaces span the tropopause and isentropic exchange by tropopause folding occurs. The region above the 380-K surface is the "overworld," in which isentropes lie entirely in the stratosphere. Light shading in the overworld denotes wave-induced forcing (the extratropical "pump"). The wavy double-headed arrows denote meridional transport by eddy motions, which include tropical upper tropospheric troughs and their cutoff cyclones, as well as their midlatitude counterparts including folds [see *Hoskins et al.*, 1985, Figure 2a]. Not all eddy transports are shown, and the wavy arrows are not meant to imply any two-way symmetry. The broad arrows show transport by the global-scale circulation, which is driven by the extratropical pump (see section 3). This global-scale circulation is the primary contribution to exchange across isentropic surfaces (e.g., the 400-K surface) that are entirely in the overworld.

stratosphere generally has a potential temperature around 380 K, depending on cloud top heights (see Figure 3). In this review the 380-K isentrope will be used to designate the lower boundary of the overworld, but the reader should keep in mind that the actual boundary is somewhat variable.

It is useful then to distinguish transport in the overworld from transport in the lowermost stratosphere. Transport in the lowermost stratosphere requires consideration of the details of synoptic-scale and small-scale processes and how these link to the overworld. Horizontal mixing can be especially significant in the lowermost stratosphere and, most of all, as was already mentioned, during blocking events, when meridional motions are enhanced. Thus exchange between the troposphere and the lowermost stratosphere can be significantly faster than exchange between the overworld and the lowermost stratosphere.

The distinction between the overworld and the lowermost stratosphere is therefore of prime importance. It should be clear that owing to this distinction it is not essential (indeed, it may be misleading) to measure STE by the transport across the tropopause. For many purposes, transport across another control surface may be more relevant and more effectively evaluated. For example, consider the case of a chemical species (e.g., methane (CH_4), nitrous oxide (N_2O), or one of the chlorofluorocarbons (CFCs)) that has a tropospheric source and a stratospheric sink, with the sink being above 18 km or so, in the overworld. The transport across the 380-K isentrope (see Figure 3), which, as will be argued in sections 3 and 4, can largely be understood as part of the global-scale circulation of the overworld, is then a perfectly acceptable measure of exchange, indeed often more relevant because of the higher location of the photochemical sink.

The same applies to a species that has a stratospheric source and a largely tropospheric sink. In this context, details of the transport across the tropopause, involving processes such as tropopause folding and overshooting tropical cumulus anvils, are largely irrelevant. On the other hand, such details are important

for determining the locations and intermittency of STE and for calculating measures of STE relevant to species that have appreciable sources in the part of the stratosphere below the 380-K isentrope (and therefore in the lowermost stratosphere), for example, from aircraft emissions.

Indeed, recent developments in atmospheric chemistry suggest that the lowermost stratosphere and upper troposphere are more important for global chemical processes than was formerly thought. For example, some of the heterogeneous chemistry responsible for observed ozone depletion occurs in the lowermost stratosphere, and much of the tropospheric nonurban photochemical ozone production occurs in the upper troposphere [*WMO*, 1995]. Understanding and modeling the chemistry of these regions requires that we know the temporal and spatial distribution of trace chemical transport in the upper troposphere and lowermost stratosphere, both along and across isentropic surfaces. The wavy arrows in Figure 3 are not, of course, meant to suggest the complete picture, even for transport along isentropes. The "exchange" need not be symmetric; there may well be stronger transports along isentropes within the lowermost stratosphere, as well as outside it.

Our first aim then is to place STE in the framework of the general circulation and to clarify the roles of the different mechanisms involved on various scales. Sections 3 and 4 begin by reviewing our understanding of global-scale transport, considering first its main mechanism via a relevant idealized model and then the application to the real atmosphere. Section 5 considers quantitative estimates of global transport made on the basis of our theoretical understanding and of global-scale dynamical and chemical observations. Section 6 reviews synoptic-scale and smaller-scale aspects of exchange in the extratropics. Section 7 considers the tropics, where certain aspects of exchange, principally concerned with water vapor, clearly do depend crucially on small-scale details. The representation of both global-scale and small-scale aspects of exchange in models is considered in section 8. Finally, a summary and conclusions are given in section 9, followed by a glossary of technical terms.

3. GLOBAL-SCALE DYNAMICAL ASPECTS OF EXCHANGE: THE EXTRATROPICAL PUMP

A prime requirement is to understand how the processes suggested in Figure 3 fit together mechanistically. One prerequisite to such understanding is to recognize the existence of nonlocal dynamical effects in the atmosphere.

Recognition of such nonlocal effects is crucial to understanding any fluid-dynamical system that supports fast waves, in the appropriate sense, that is, waves whose travel times are short in comparison with other timescales of interest. For example, to understand why air moves toward the inlet of any suction pump, one needs to recognize that the travel times of acoustic waves are short in this sense and that there is a corresponding nonlocal effect. Mass conservation has a key role and, for practical purposes, acts instantaneously, as a nonlocal constraint. The time delay between turning on the pump and the establishment of the flow toward the suction tube is of the order of an acoustic propagation time, usually negligible in comparison with all other timescales of interest. This nonlocal picture is to be contrasted with statements like "air moves toward the suction tube because the pressure-gradient force pushes it," which miss the point that the pressure gradient adjusts nonlocally, to fit in with the mass-conservation constraint. The pressure-gradient force, whether written on the right or the left of the momentum equation, cannot usefully be regarded as causing the flow. The nonlocal character of the adjustment can be expressed mathematically via an elliptic (see glossary) partial differential equation for the pressure, formed by taking the divergence of the momentum equation.

There is a fundamentally similar nonlocal effect having direct relevance to the STE problem, namely, the effect of the extratropical stratosphere and mesosphere upon the tropical stratosphere. This depends on the fact that the global-scale travel times of acoustic waves and large-scale gravity waves are short in comparison with other timescales of interest. The effect has been demonstrated, indeed has been intensively studied, in the dynamical literature beginning with the pioneering work of *Eliassen* [1951] and *Dickinson* [1968]. These and many other studies have shown that the extratropical stratosphere and mesosphere act persistently upon the tropical lower stratosphere as a kind of global-scale fluid-dynamical suction pump, driven by certain eddy motions. The distinction between tropics and extratropics arises from the Earth's rotation, on which the extratropical pumping action depends.

One reason why we can be confident about the sign implied by the word "suction" is that statistically speaking, the most important of the eddy motions in question, so-called "breaking Rossby waves" and related potential-vorticity-transporting motions, have a one-signed or ratchet-like character related to the sense of the Earth's rotation. (For detailed examples and historical remarks, see, for instance, *McIntyre* [1993, and references therein]). These eddy effects have a strong tendency to add up and to give a persistently one-way pumping action, whose strength varies seasonally and interannually and in which air is gradually withdrawn from the tropical stratosphere and pushed poleward and ultimately downward. Where air parcels are being pulled upward (as in the tropical stratosphere), adiabatic cooling pulls temperatures below their radiative values; where air parcels are being

pushed downward (as happens most strongly in high latitudes in winter and spring), adiabatic warming pushes temperatures above their radiative values.

This mechanically pumped global-scale circulation is often referred to by dynamicists as the "wave-driven circulation." Its existence not only explains why observed temperatures differ from the radiative values toward which they would otherwise tend to relax [e.g., *Dickinson*, 1969; *Fels*, 1985] but is also an essential part of what controls the tropospheric lifetimes of long-lived trace chemicals such as CFCs, nitrous oxide, and methane that have photochemical sinks in the stratosphere. The main uncertainty today is not the physical reality of the extratropical pumping action, but its precise strength and its seasonal and interannual variability, and certain nonlinear aspects of the resulting circulation dynamics that bear on the precise way the pumping action reaches into the tropics.

The nonlocal effects in question operate just as robustly in a zonally symmetric, that is, longitude-independent, model atmosphere as they do in the real atmosphere. We can gain more detailed insight into the pumping action, including the way in which the implied adiabatic warming and cooling interacts with radiation on different timescales, by considering a thought experiment performed on a zonally symmetric atmosphere in which the one-signed Rossby-wave effects are parameterized simply as a retrograde or westward wave-induced force. This force plus the Earth's rotation is the essential cause of the pumping, which can be thought of as a "quasi-gyroscopic" effect in the sense that pushing a ring of air westward has a tendency to make it move poleward. The relevant theoretical formulations were first put forward in the papers by Eliassen and Dickinson already referred to, and *McIntyre* [1992] provides an elementary introduction highlighting the underlying physical principles with some illustrative thought experiments.

A suitable set of model equations for such a zonally symmetric atmosphere is as follows, with the wave-induced force per unit mass denoted by $\bar{G}$:

$$\frac{\partial \bar{u}}{\partial t} - 2\Omega \sin\phi \bar{v}^* = \bar{G} \tag{1}$$

$$2\Omega \sin\phi \frac{\partial \bar{u}}{\partial z} + \frac{R}{aH}\frac{\partial \bar{T}}{\partial \phi} = 0 \tag{2}$$

$$\frac{\partial \bar{T}}{\partial t} + \bar{w}^*\left(\frac{HN^2}{R}\right) = \bar{Q}_s + \bar{Q}_l(\bar{T}) \tag{3}$$

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}(\bar{v}^* \cos\phi) + \frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0 \bar{w}^*) = 0 \tag{4}$$

Here ϕ is latitude and z is log-pressure height (approximately corresponding to geometric altitude), Ω is the angular velocity of the Earth, R is the gas constant for dry air, $H \cong 7$ km is the nominal (constant) density scale height used in the log-pressure coordinates, a is the radius of the Earth, and ρ_0 is a nominal basic state density proportional to $\exp(-z/H)$. In (3), N^2 is the square of the buoyancy frequency (a measure of static stability), defined as

$$N^2 \equiv \frac{R}{H}\left(\frac{dT_0}{dz} + \frac{\kappa T_0}{H}\right)$$

where T_0 is a reference temperature dependent only on z and $\kappa \equiv R/c_p$, where c_p is the specific heat of air at constant pressure. Finally, $\bar{Q}_s$ and $\bar{Q}_l(\bar{T})$ are short-wave and long-wave radiative heating rates, respectively, $\bar{u}$ is the zonal wind component, $\bar{T}$ is the temperature, and $\bar{v}^*$ and $\bar{w}^*$ are the latitudinal and vertical velocity components, respectively. These quantities may, as will be explained later, be interpreted as certain zonally averaged quantities in the real atmosphere or in three-dimensional atmospheric models (that is why the overbars are added). For the moment, however, $\bar{u}$, $\bar{T}$, $\bar{v}^*$, and $\bar{w}^*$ are simply taken as the zonally symmetric fields of the thought experiment, in which we shall imagine, furthermore, that the wave-induced force $\bar{G}$ is prescribed. The equations will then describe a pumping action of the kind in question: a global-scale manifestation of fluid-dynamical nonlocalness.

For the sake of mathematical simplicity, these particular model equations use the approximations of so-called quasi-geostrophic theory [e.g., *Andrews et al.*, 1987]. The quasi-geostrophic approximations are adequate for gaining qualitative insight into the mechanism of the pumping action and the timescales involved. Studies in which the approximations are relaxed, using the more accurate "primitive equations," are known to give qualitatively the same results [e.g., *Haynes et al.*, 1991]. Equation (1) is the zonal momentum equation. It states that the wave-induced force is balanced by the zonal flow acceleration minus the Coriolis force due to the latitudinal motion $\bar{v}^*$. Equation (2) is the thermal wind equation. It states that the vertical shear of the zonal wind is proportional to the latitudinal gradient of the temperature. This important coupling between the wind and temperature fields follows from the assumption that the zonal flow is in hydrostatic and geostrophic balance. This is part of how the theory includes the all-important nonlocal effects. Equation (3) is the thermodynamic energy equation. It states that the net diabatic heating is balanced by the sum of adiabatic cooling by the vertical motion and the temperature tendency. In (3) we distinguish between the part of the diabatic heating that is largely independent of temperature (more or less corresponding to the short-wave radiative heating) and the part that depends strongly, and nonlocally, on the temperature field (more or less corresponding to the long-wave radiative heating and responsible for the relaxational nature of the total

diabatic heating). Equation (4) is the mass-continuity or mass-conservation equation. It relates the latitudinal and vertical components, respectively, of the velocity in the latitude-height plane. This velocity field is often referred to as the "meridional circulation."

Equations (1)–(4) may be regarded as predictive equations for the "unknowns" $\partial\bar{u}/\partial t$, $\partial\bar{T}/\partial t$, $\bar{v}^*$, and $\bar{w}^*$. Given the wave-induced force, which appears solely as the $\bar{G}$ term on the right-hand side of (1), and the contributions to diabatic heating, $\bar{Q}_s$ and $\bar{Q}_l$, they may be combined to give a single equation for one of the unknown quantities. In order to simplify the calculations and to focus on the timescales of importance, it is convenient to follow *Garcia* [1987] and to assume that the time dependence is harmonic, with constant frequency σ. (Since the problem is linear, more general time dependencies can be deduced by Laplace transformation.) Thus we write $\bar{G} = \mathrm{Re}\,(\hat{G}e^{i\sigma t})$ and $\bar{Q}_s = \mathrm{Re}\,(\hat{Q}e^{i\sigma t})$ and characterize the response by the single variable $\hat{w}$ such that the vertical velocity $\bar{w}^* = \mathrm{Re}\,[\hat{w}(\phi, z)e^{i\sigma t}]$. As was emphasized earlier, it is important that the relaxational dependence of the diabatic heating field on the temperature be taken into account. It is convenient to parameterize this dependence by assuming $\bar{Q}_l$ to be given by Newtonian cooling with constant timescale α^{-1}, that is, $\bar{Q}_l = -\alpha\bar{T}$, with $\bar{T}$ now the temperature anomaly. With the above assumptions, (1)–(4) can be combined into a single equation for $\hat{w}$ of the following form:

$$\frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial(\rho_0\hat{w})}{\partial z}\right) + \left(\frac{i\sigma}{i\sigma+\alpha}\right)\frac{N^2}{4\Omega^2a^2\cos\phi}\frac{\partial}{\partial\phi}\left(\frac{\cos\phi}{\sin^2\phi}\frac{\partial\hat{w}}{\partial\phi}\right) = \left(\frac{i\sigma}{i\sigma+\alpha}\right)\frac{(R/H)}{4\Omega^2a^2\cos\phi}\frac{\partial}{\partial\phi}\left(\frac{\cos\phi}{\sin^2\phi}\frac{\partial\hat{Q}}{\partial\phi}\right) + \frac{1}{2\Omega a\cos\phi}\frac{\partial}{\partial\phi}\left(\frac{\cos\phi}{\sin\phi}\frac{\partial\hat{G}}{\partial z}\right) \quad (5)$$

This equation provides a suitable model for studying nonlocal control of the meridional circulation by the wave-induced force and by the short-wave heating $\bar{Q}_s$, which we also, of course, take to be prescribed.

Since the operator acting on $\hat{w}$ on the left-hand side of (5) is elliptic (see glossary) when $\sigma/\alpha \neq 0$, it follows that a wave-induced force, or a short-wave heating, localized to a particular region, will give rise to a response that extends away from that region. Conversely, $\hat{w}$ in a given region will be affected by changes in the wave-induced force or the short-wave heating elsewhere. This is how the mathematics expresses the nonlocalness. We can say that $\hat{w}$ and therefore $\bar{w}^*$ are "under nonlocal control."

For reasons to emerge (e.g., Figure 4), we shall be especially interested in the downward and sideways control exerted by a localized forcing $\bar{G}$ upon the vertical mass flux $\rho_0\bar{w}^*$. What is the spatial structure of the response to a localized forcing? Inspection of (5) shows that all aspects of the response depend on the parameter σ/α. As $\sigma/\alpha \to \infty$, the adiabatic limit studied by *Eliassen* [1951] and extended to a spherical geometry by *Plumb* [1982] is approached. A value for α of 1/(20 days) is often taken as representative of the lower stratosphere. The adiabatic limit will therefore hold for timescales of less than a few days. This limit is appropriate for transient fluctuations in $\bar{G}$, such as those induced by Rossby waves in stratospheric sudden warmings. In that case, $i\sigma/(i\sigma + \alpha) \to 1$, and since away from the forcing region the two terms on the left-hand side in (5) must balance, simple scaling then implies that the vertical scale Δz in $\rho_0\bar{w}^*$ is related to the latitudinal scale $a\Delta\phi$ by

$$\Delta z \sim \max\left[\frac{2\Omega\sin\phi}{N}a\Delta\phi, \frac{4\Omega^2\sin^2\phi}{N^2}\frac{(a\Delta\phi)^2}{H}\right]$$

where the left-hand part of the expression applies when $\Delta z < H$ and the right-hand part applies when $\Delta z > H$. Note that when the left-hand part applies, the ratio of vertical scale to latitudinal scale equals the ratio of the Coriolis frequency to the buoyancy frequency, which in middle latitudes is of the order of 10^{-2}. An example of this adiabatic limit is shown in Figure 4a. The arrowed curves are the mass-flux streamlines of the response to a simple forcing, consisting only of a westward force $\bar{G} < 0$ applied in the shaded region. Note that the mass-flux pattern extends mainly downward and sideways and that the right-hand part of the above expression for Δz is relevant. The pumping action immediately reaches to a great depth proportional to the square of the latitudinal scale of the pattern, the density scale height H being far smaller than the vertical scale of the pattern. For practical purposes we may think of the bottom of Figure 4a as the bottom of the stratospheric overworld, around 380 K or 15 km above the Earth's surface. The pumping action reaches sideways across the tropics and into the opposite hemisphere, in agreement with observations of large-scale transient events in the real atmosphere [e.g., *Dunkerton et al.*, 1981; *Randel*, 1993].

If $\alpha > \sigma$, that is, for timescales of a month or more when $\alpha \approx$ 1/(20 days), the characteristic scales of the response are related by

$$\Delta z \sim \max\left[\left(\frac{\alpha}{\sigma}\right)^{1/2}\frac{2\Omega\sin\phi}{N}a\Delta\phi, \left(\frac{\alpha}{\sigma}\right)\frac{4\Omega^2\sin^2\phi}{N^2}\frac{(a\Delta\phi)^2}{H}\right] \quad (6)$$

The square of the latitudinal scale is again relevant to the response at the largest scales, and for any given

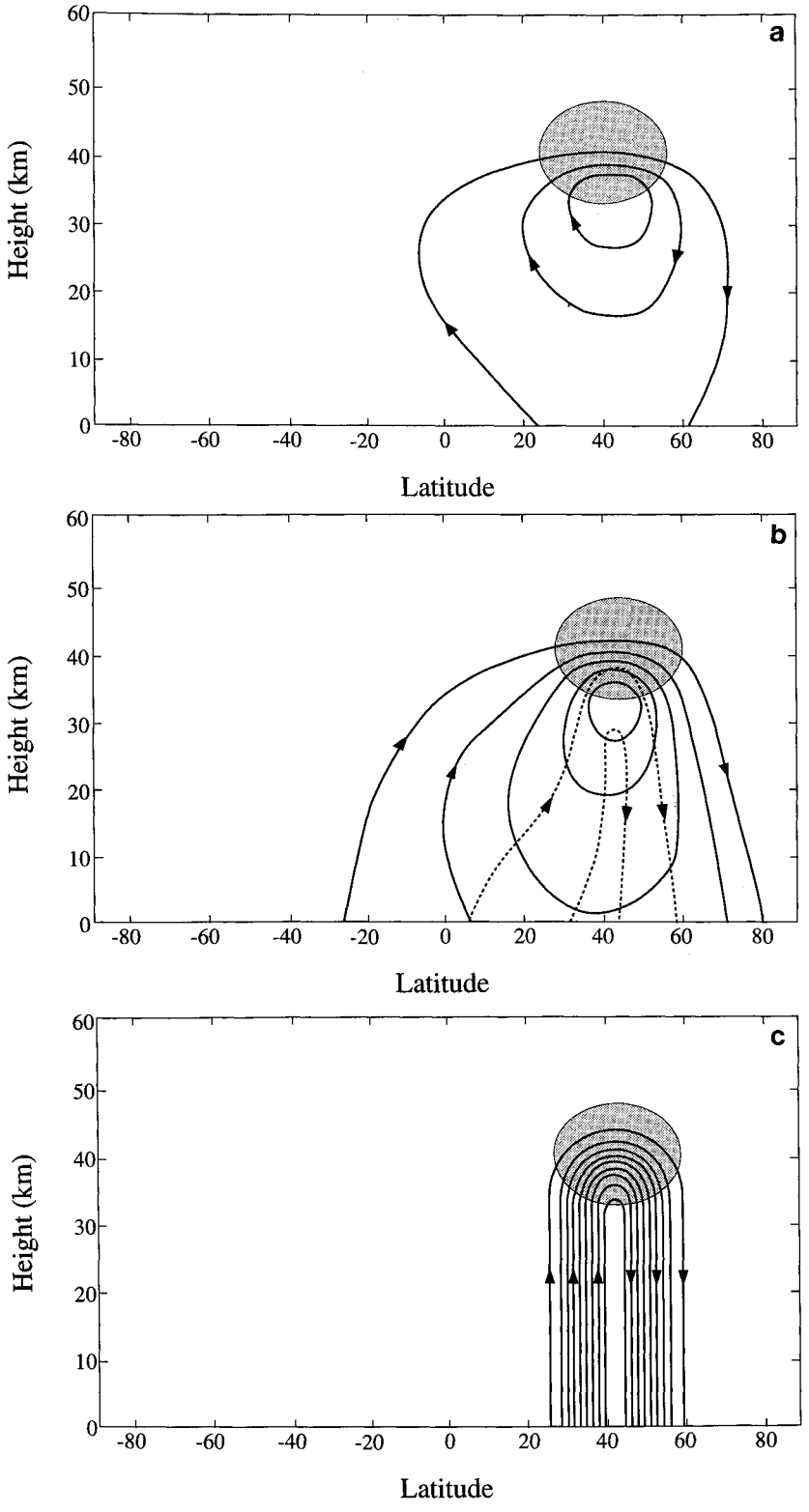

Figure 4. Idealized numerical experiments on the response of the mean stratospheric zonally symmetric circulation in the latitude-height plane to a westward force applied in the shaded region. Contours are streamlines, with the same contour interval used in each panel. (a) Adiabatic response for $\sigma/\alpha \gg 1$. (b) Response for $\sigma/\alpha = 0.34$, corresponding to annual frequency and 20-day radiative damping timescale; the solid and dashed contours show the response that is in phase and 90° out of phase with the forcing, respectively. (c) Steady state response ($\sigma/\alpha \ll 1$). Note that the magnitude of the response in the mass stream function increases as σ/α decreases and that it is given quantitatively by (7) for Figure 4c. The model is unbounded below, so the ordinate has arbitrary origin but can be regarded, for instance, as showing approximate vertical distance above the lower boundary of the overworld.

latitudinal scale the response extends still deeper vertically than is the case in the adiabatic limit. This occurs because as the forcing timescale increases relative to the radiative damping timescale, that is, as σ/α decreases, the importance of the coefficient of the latitudinal derivative term on the left-hand side of (5) diminishes relative to the vertical derivative term. The response must therefore have a deeper vertical scale if these two terms are to balance (as they must away from the forcing region). Such deepening is seen in Figure 4b, where $\sigma/\alpha = 0.34$, corresponding to annual frequency and a radiative damping time of 20 days. Also, the sideways control reaches still farther into and across the tropics.

In the limit as the timescale becomes very large, only the vertical derivative term remains on the left-hand side of (5). The equation then changes its character to hyperbolic, and the streamlines away from the forcing are lines of constant latitude (Figure 4c). In the same limit the heating term on the right-hand side of (5) vanishes. It follows that in this limit the mean circulation is controlled entirely by the wave-induced force, even if $Q \neq 0$. An imposed short-wave heating cannot drive a persistent meridional circulation but is simply balanced by an adjustment in the long-wave temperature-dependent part of the heating. Radiative equilibrium will hold everywhere except below the region in which there is a wave-induced force. There, adiabatic heating and cooling due to the vertical circulation drive the temperature away from radiative equilibrium.

This steady state limit, $\sigma/\alpha \to 0$, may also be analyzed by considering the momentum equation (1) directly. In the steady state the acceleration $\partial\bar{u}/\partial t = 0$, and the wave-induced force is balanced by the Coriolis force. It follows from combining (1) with (4) and requiring that $\rho_0\bar{w}^* \to 0$ as $z \to \infty$, that $\bar{w}^*$ is given by

$$\bar{w}^* = -\frac{1}{\rho_0 a \cos\phi}\frac{\partial}{\partial\phi}\left[\frac{\cos\phi}{2\Omega\sin\phi}\int_z^\infty \rho_0\bar{G}\,dz'\right] \quad (7)$$

Thus in the steady state limit the nonlocal control of the mass flux expressed by (5) has become a purely downward control, as illustrated in Figure 4c. A special case of this was noted by *Dickinson* [1968]. The vertical velocity at some location is controlled by the distribution of wave-induced forcing (and its latitudinal derivative) above that location, as is explicitly shown by expression (7). This "downward control" principle is discussed by *Haynes et al.* [1991], together with its generalization beyond quasi-geostrophic theory.

The downward control principle is valid only in the steady state limit, and as Figure 4 reminds us, the approach to the limit is not uniform in latitude. Noting that in the steady state limit the second term on the left-hand side of (5) vanishes, we deduce that the value of σ (and hence the timescale) for which the response approaches steadiness is proportional to the radiative time α^{-1} and is given by

$$\sigma^{-1} \geq \alpha^{-1}\frac{N^2}{4\Omega^2\sin^2\phi(a\Delta\phi)^2}\min\left[\Delta z^2, H\Delta z\right] \quad (8)$$

In (8), $a\Delta\phi$ and Δz can take on different interpretations. If they are assumed to represent the scales of the forcing itself, then the above timescale is that to attain steady state in the region where the forcing is applied. On the other hand, if Δz is taken to be the vertical distance from a localized forcing of horizontal scale $a\Delta\phi$, then the timescale is that for the steady state to be attained at the location in question. Note that the expression is then consistent with the notion that the response is diffusive, with timescale proportional to $(\Delta z)^2$, if $\Delta z \ll H$, and propagating, with timescale proportional to Δz, if $\Delta z \gg H$ (see *Haynes et al.* [1991] for more details). It follows from the above that the timescale to reach a steady state decreases as either the latitudinal scale ($a\Delta\phi$) or the latitude itself increases. Conversely, in the tropics the timescale to reach a steady state becomes very long. For this and other reasons to do with nonlinear effects in the tropics, the steady state limit, and therefore the downward control principle, is of direct practical interest only in the extratropics.

To gain some quantitative insight, we might take Δz to be equal to H and, again, the radiative timescale to be 20 days. If the steadiness criterion is to be satisfied for variations with the annual frequency, in the sense that the left-hand side of (8) is at least a factor of 2 greater than the right-hand side, then at 45° latitude the horizontal scale must be at least 1100 km, while at 20° it increases to 2000 km.

When the zonally symmetric model considered here is extended to apply to a zonally averaged circulation in the real atmosphere, the downward control principle (equation (7)) provides a framework in which to interpret the wave driving of the meridional circulation and also a simple quantitative recipe for calculating the strength of this circulation, given observational estimates of the zonal mean force due to Rossby waves and to gravity waves. Since the climatological distributions of zonal mean Rossby-wave and gravity-wave forces in the stratosphere and mesosphere appear to be hemispheric in scale (Figure 5), it is possible that estimates of the seasonally and zonally averaged extratropical meridional circulation based on the downward control principle could be fairly reliable, at least in the solstice seasons. This has been verified to a remarkable degree with observational data and with output from a general circulation model; see section 5.

What is especially remarkable is the apparent success of such calculations even in the subtropics, beyond their domain of validity. As was noted earlier, as the equator is approached, the timescale to achieve the steady state, in which the wave-induced force $\bar{G}$ comes into balance with the Coriolis force associated with the meridional flow, becomes longer and longer. Thus if the wave-induced force has nonzero frequency, there must always be a low-latitude region in which downward control does not apply but in which the wave-induced force is balanced instead by an acceleration of the zonal wind. Such a response is observed, for example, in the quasi-biennial zonal wind oscillation in the lower equatorial stratosphere and the semiannual oscillation in the upper equatorial stratosphere and mesosphere (see, for example, *Andrews et al.* [1987, chapter 8]). To estimate the width of this low-latitude region, it is appropriate to let $\sin\phi \approx \Delta\phi$ in (8), which may then be rewritten as

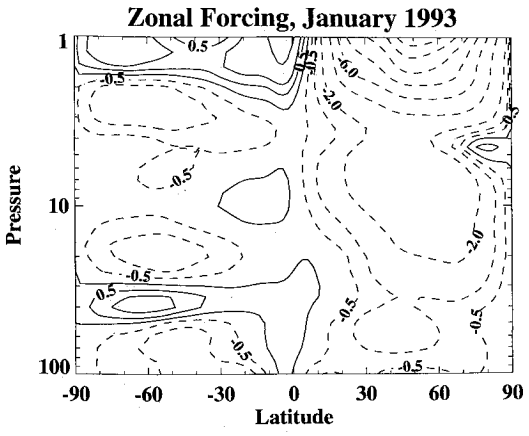

Figure 5. Wave-induced zonal force per unit mass, $\bar{G}$, in units of meters per second per day, for January 1993, deduced indirectly from the momentum balance (equation (1)) using a radiatively derived meridional circulation and the observed zonal mean wind. Contours are at 2 m s^{-1} d^{-1} intervals, with additional contours of ± 1 and ± 0.5 m s^{-1} d^{-1}. The eastward force centered near 60°S and 40 hPa may be an artifact caused by the ill-conditionedness in calculating radiative heating rates from observed temperatures, especially in the summertime lower stratosphere. (After *Rosenlof* [1995]).

$$a\Delta\phi \geq \left(\frac{\sigma}{\alpha}\right)^{1/4}\left(\frac{N^2a^2}{4\Omega^2}\min\left[\Delta z^2, H\Delta z\right]\right)^{1/4} \qquad (9)$$

The right-hand side of (9) may be interpreted as giving the minimum horizontal distance from the equator for the steady state downward control regime to hold. For the values of α, σ, and Δz taken earlier, this distance is about 1900 km, though a more accurate theory may well give a somewhat larger value through nonlinear effects; see below. These considerations, taken together with results like Figure 4b, strongly indicate that the most effective location for wave-induced forcing, if it is to control the annual variation of upwelling at the equator (see *Yulaeva et al.* [1994] and section 5), is in the subtropics. On the annual timescale the upwelling response to forcing at high latitudes will decay substantially at low latitudes (evident from Figure 4b), whereas the response to a force at a latitude equatorward of that given by the right-hand side of (9) appears primarily as an acceleration, $\partial\bar{u}/\partial t$, rather than an induced meridional circulation. This in turn suggests a crucial role for the position of the subtropical edge of the "surf zone" in the wintertime middle stratosphere, where Rossby-wave breaking can lead to a significant contribution to $\bar{G}$.

A second limitation on the downward control principle (equation (7)) at low latitudes is the neglect of relative angular momentum gradients, that is, of latitudinal and vertical advection acting on gradients of $\bar{u}$ in the quasi-geostrophic version of the zonally symmetric model equations given above. At low latitudes, such advection becomes increasingly important. These extra terms may be incorporated into the theory most conveniently if (1) is replaced by the angular momentum form:

$$\frac{\partial\bar{m}}{\partial t} + \frac{\bar{v}^*}{a}\frac{\partial\bar{m}}{\partial\phi} + \bar{w}^*\frac{\partial\bar{m}}{\partial z} = (a\cos\phi)\bar{G} \qquad (10)$$

where $\bar{m} = a\cos\phi(\bar{u} + a\Omega\cos\phi)$ is the absolute angular momentum. (Again, the notation $\bar{m}$ is used because the equation also applies for zonally averaged quantities when the flow is zonally asymmetric.) Although $\bar{m}$ is dominated by the contribution from the Earth's rotation, the latitudinal gradient of $\bar{m}$, which appears in (10), is influenced by the contribution from the relative motion $\bar{u}$, particularly at low latitudes. The extra terms in (10) can be incorporated to give a nonlinear generalization of the downward control expression (equation (7)), which then becomes an integral along contours of constant absolute angular momentum. Further details are given by *Haynes et al.* [1991]. The full expression has been used by *Rosenlof and Holton* [1993] in their estimate of $\bar{w}^*$ from observations.

An important aspect of the downward control principle that is weakened in the tropics is that the change in $\bar{w}^*$ due to a change in wave-induced forcing is no longer given entirely by appropriate modification of the $\bar{G}$ appearing in (7) or its nonlinear generalization. Rather, since a significant part of the changed forcing may cause a zonal flow acceleration, the resulting change in the angular momentum distribution must also be taken into account. At low latitudes, then, the problem of determining the meridional circulation given the wave forcing is fully nonlinear; the meridional circulation and the angular momentum distribution have to be determined simultaneously. It can be argued, from considerations of angular momentum advection, that the effective value of Ω in (9) is reduced by such nonlinearities, extending the reach of extratropical pumping action into and across the tropics even in the steady state.

Another possibility arising from nonlinearity in the tropics is when the zonal wind distribution may be such that the angular momentum contours become closed in the interior of the atmosphere (that is, there is an interior maximum in angular momentum). It is then possible to have a free, closed meridional circulation that is everywhere parallel to the angular momentum contours. Since the advection of angular momentum in (10) then vanishes, such a circulation can exist in the absence of a wave-induced force. This regime was investigated for the tropospheric Hadley circulation by *Schneider* [1977] and *Held and Hou* [1980]. *Dunkerton* [1989] investigated the extent to which a similar regime may be relevant to the equatorial middle atmosphere, particularly at the solstices, and modeled the formation of a low-latitude free meridional circulation cell associated with the cross-

equatorial gradient of radiative heating in the middle and upper stratosphere. However, the relevance of this nearly inviscid free circulation to the time-averaged upwelling in the tropical lower stratosphere is doubtful. There is little resemblance, even at low latitudes, between this free circulation and the circulation calculated diabatically from observations [e.g., *Solomon et al.*, 1986; *Rosenlof*, 1995], particularly with regard to the latter's broad region of tropical upwelling in the lower stratosphere. *Dunkerton* [1991] showed that the addition of a wintertime extratropical wave force to his model enhances the circulation, giving systematic upwelling in low latitudes and extending the downwelling circulation well into the winter hemisphere. Thus while a clear mathematical theory of the connection between the time-averaged extratropical force and the time-averaged tropical upwelling remains to be completed, the wave-induced forcing appears to make the dominant contribution to driving the circulation, including exchange across the base of the overworld. In other words, it seems clear for practical purposes that the mass transport across any quasi-horizontal surface, even the tropical upwelling, is primarily controlled by the extratropical wave-driven pumping above that surface.

4. GLOBAL-SCALE DYNAMICAL ASPECTS OF EXCHANGE: APPLICATION TO THE REAL ATMOSPHERE

The previous section has explained how nonlocal control of the meridional circulation arises in the dynamics of a zonally symmetric model atmosphere. How is this relevant to transport of trace chemicals in the real atmosphere or in realistic numerical models, where "waves and eddies," or departures from zonal symmetry, are often very strong?

As was already suggested, the most direct relevance is to exchange across the 380-K surface, and there are two basic points. The first is that the strongest departures from zonal symmetry are usually found in the extratropics. The second is that in the overworld, there is still a sharp distinction between tropics and extratropics, evident from the observation that many chemical species of interest in the overworld have different tropical and extratropical mixing ratios (recent references among many being *Trepte and Hitchman* [1992] and *Randel et al.* [1993]). This suggests at once that horizontal eddy exchange between the tropical and extratropical overworlds is weak, as was also noted long ago from classic observations of nuclear bomb test debris. Recent new evidence is mentioned in section 7. Fluid-dynamical model studies [e.g., *Norton*, 1994; *Polvani et al.*, 1995; *Chen et al.*, 1994] show that this is explicable in terms of a subtropical "eddy transport barrier" against transport along isentropes. As explained by *Juckes and McIntyre* [1987] in connection with the similar but tighter eddy transport barriers found at the edges of the winter polar vortices, such barriers work via a combination of dynamical effects involving both horizontal shear on the one hand and isentropic gradients of PV or "Rossby-wave elasticity" on the other [*Hoskins et al.*, 1985, section 6].

The subtropical eddy transport barrier is not a perfect barrier, especially in the layer between 380 and 480 K [e.g., *Trepte et al.*, 1993; *Boering et al.*, 1995], but is still the main reason why the mean-circulation picture suggested in Figure 3 is relevant to global-scale chemical transport. It is the strong correlation between the global-scale vertical velocity and the chemical mixing ratios on either side of the barrier that make the mean circulation chemically significant. As has been emphasized by R. A. Plumb (personal communication, 1994), the net transport by the mean meridional circulation, or by any circulation, of any chemical into the overworld would vanish if the mixing ratio were uniform on, say, the 380-K surface.

The usual approach has been to focus on the height-latitude structure of dynamical and chemical tracer fields by dividing all fields into a longitudinally averaged part (the "zonal mean") and an "eddy" part. There are a number of alternative ways in which the zonal mean may be computed (see glossary). The separation between zonal mean and eddy contributions to transport differs according to the type of averaging employed. These differences are manifested, for example, by the striking difference between the conventional Eulerian-mean circulation measured in pressure or log-pressure coordinates and the Eulerian-mean circulation measured in isentropic coordinates (i.e., with potential temperature as a vertical coordinate [e.g., *Tung*, 1982]. Of the two, only the latter, often referred to as the "zonal-mean diabatic circulation," or diabatic circulation for short, provides a qualitatively correct approximation to the advective transport of tracers. Indeed, in the high-latitude winter stratosphere the conventional Eulerian-mean circulation and the diabatic circulation are in opposite senses. This is one of the reasons why the use of isentropic coordinates greatly facilitates conceptual modeling of transport.

The isentropic coordinate formulation also has the important conceptual advantage of clearly separating transport along isentropic surfaces from transport across isentropic surfaces. One qualitative advantage of such coordinates is that eddy dispersion, which is primarily along isentropic surfaces, appears nearly horizontal. This is true to the extent that the eddies are quasi-adiabatic [e.g., *Tung*, 1982] and is in accord with direct evidence from model simulations [*Plumb and Mahlman*, 1987]. The zonally symmetric mean diabatic circulation, on the other hand, generally provides a useful approximation to the cross-isentropic trans-

port and is thus particularly relevant for questions of STE.

On the other hand, pressure coordinates have some advantages for purposes of data analysis. For this reason a modified version of the Eulerian-mean equations in pressure coordinates, the "transformed Eulerian mean" (TEM), is often used in practice. The TEM provides a measure of the global-scale meridional circulation that is similar to the diabatic circulation of the isentropic coordinate formulation. The TEM equations have exactly the same structure as the model equations (1)–(4), except that zonally averaged quantities replace zonally symmetric quantities. Furthermore, eddy effects do indeed appear, to good approximation, as wave-induced forces in this formulation. The Rossby-wave part of such forces involves "breaking," as well as dissipation by infrared radiative relaxation, and hence involves eddy dispersion. Thus we may apply insights gained from the zonally symmetric model to deduce that the global-scale meridional circulation, and therefore the zonal mean transport of tracers across isentropic surfaces, is under nonlocal control by the zonally averaged wave-induced forces.

It follows therefore that on sufficiently long timescales the transport across a lower stratospheric control surface, including, for example, the 380-K isentrope, is largely controlled by wave-induced forces in the extratropical middle atmosphere, which in turn are believed to be mainly due to Rossby and gravity waves originating in the extratropical troposphere. From observations and modeling it appears that wave-induced forces in both the stratosphere and mesosphere are important contributors to the distribution of wave-induced forcing (Figure 5). On the whole, the Rossby-wave force makes the most important contribution to transport across lower stratospheric control surfaces, though there are important variations with latitude and season, the mesospheric gravity wave force being particularly important in the southern hemisphere winter [e.g., *Garcia and Boville*, 1994]. Furthermore, gravity wave forces in the summer hemisphere and middle and upper stratosphere, though not well known, are almost certainly an essential part of how the rising branch of the stratospheric circulation varies seasonally.

Note well the implication of this picture for the causes of upwelling in the tropical lower stratosphere. The broad-scale upwelling mass flux averaged over the tropics is largely independent of local processes. These include both the large-scale tropospheric heating field due to moist convection and the mass injection by cumulonimbus turrets penetrating the lower stratosphere. The approximate independence of the broad-scale upwelling from such heating and cumulonimbus effects is sometimes regarded with surprise, but it is an inescapable consequence of fluid-dynamical nonlocalness as manifested by extratropical pumping.

Thus *Yulaeva et al.* [1994] have argued that the annual cycle in lower stratospheric temperatures (and by implication, in the upwelling) is primarily due to the interhemispheric contrast in wintertime stratospheric Rossby-wave activity and not to annual cycles in sea surface temperatures and tropical convective activity. This has interesting implications for STE, both for the global-scale upwelling rate and also for the dehydration mechanism, which requires sufficiently low tropopause temperatures (and which will be examined in more detail in section 7).

Refinements to this picture of nonlocal control of the global-scale circulation by $\bar{G}$ must take account of the following:

1. The detailed latitude-height distribution of $\bar{G}$ depends on wave refraction, wave breaking, and radiative damping of waves in the middle atmosphere. Although broad-scale tropical stratospheric upwelling, for instance, is largely controlled by the extratropical $\bar{G}$ field, there are feedbacks on $\bar{G}$ itself through anything that affects the extratropical wave fields, including tropospheric wave sources.

2. A limited role for transient responses to heating in the stratosphere and troposphere may need to be recognized. The time-dependent linear theory of symmetric circulations reviewed in the previous section does predict some penetration of the circulation induced by tropospheric heating into the lowest part of the stratosphere but is probably more relevant to the seasonally varying signal [*Dickinson*, 1971b] than to the longtime average [*Dickinson*, 1971a]. Indeed, indications from the steady state nonlinear theories are that tropospheric thermal forcing cannot alone drive any upwelling in the lower stratosphere (R. A. Plumb, personal communication, 1994). The vertical penetration distance, for a circulation forced from below by heating that varies on the timescale of the annual cycle, can be estimated from the requirement that the two sides of (6) balance. If the latitudinal scale of the tropospheric heating is taken to be 1000 km, it follows that the vertical penetration distance is only 2 km or so. All this suggests that, while there will certainly be some influence of tropospheric heating on local upwelling rates in the very lowest part of the tropical stratosphere, the primary control, determining what is taken up into the stratospheric overworld, is exerted by extratropical stratospheric pumping.

3. A more accurate picture of the role of eddies in cross-isentropic transport needs to be developed. Even in the isentropic formulation the effect of the eddies on transport is not purely dispersive, nor is the cross-isentropic transport purely advective. This is inevitable with any kind of Eulerian mean. *Mahlman et al.* [1980], in their pioneering study, have noted that there is an advective part of the eddy transport in the isentropic-coordinate formulation (not included in the mean diabatic circulation) that plays an important role in the steepening of tracer surfaces relative to isentropic surfaces. Indeed, they suggest that in the winter lower stratosphere this contribution, which is caused

by anomalous diabatic heating and cooling directly associated with the eddies, is numerically comparable to that from the mean circulation itself. Correspondingly, there may also be a dispersive effect of the eddies across, as well as along, isentropic surfaces [*Mahlman et al.*, 1984]. Evidence from model simulations suggests that such dispersion has an effective vertical diffusion coefficient K_z of no greater than a few times 10^{-1} m^2 s^{-1} [*Plumb and Mahlman*, 1987]. Local vertical mixing due to clear-air-turbulence "pancakes" is thought to be typically of a similar order [*Dewan*, 1981]. Simple scaling shows that for a vertical velocity scale W and a length scale D, the ratio of vertical diffusion to vertical advection is $K_z/(WD)$. Since cross-isentropic advective transport is associated with vertical velocities of a few times 10^{-4} m s^{-1}, it is to be expected that the advective transport will dominate ($K_z/(WD) \ll 1$) on vertical length scales $D \gg 1$ km. In this sense it is justified to assume that the cross-isentropic transport is dominated by the advective transport on sufficiently large vertical scales. However, the dispersive effects could be important for the evolution of thin layers of volcanic aerosol or for aircraft emissions.

In this section the emphasis has been on STE as measured by global-scale upwelling or downwelling across a suitable control surface such as the 380-K isentropic surface. We have argued that such upwelling or downwelling is best viewed as controlled in a nonlocal way by extratropical eddy forcing in the form of wave-induced forces. As will be seen in the next section, this viewpoint not only is qualitatively insightful, but allows useful quantitative estimates of such transport, including many aspects of its seasonal and latitudinal variation [e.g., *Rosenlof*, 1995].

The foregoing picture does not imply that at every instant the flux of a constituent downward into the lowermost stratosphere need equal the flux across the extratropical tropopause into the upper troposphere, taking into account synoptic- and smaller-scale dynamical events (cutoff cyclones, tropopause folds, etc.; see Figure 3). It is quite clear that these two fluxes need not be equal. If we extend downward the zonal-mean, eddy formalism described earlier to include the upper troposphere, then with realistic eddy statistics the transport across the tropopause itself appears primarily as an eddy dispersion along isentropes. While the strength of the meridional circulation cannot be regarded as wholly independent of the eddy dispersion [e.g., *Holton*, 1986], it is a highly nonlocal function of such eddy dispersion at many altitudes, as a result of the pumping mechanism illustrated above. It follows that there need be no simple relationship between the flux of tracers across a lower control surface representative of the tropopause itself and the flux across an upper quasi-horizontal control surface at 380 K or so. Nonetheless, if the two were badly out of balance for a significant period of time, then any tracer whose sources were solely above the upper control surface would be flushed out of, or build up in, the region between the two control surfaces until the ambient concentrations were such that the fluxes balanced.

It is only on timescales long enough for the system to be in a statistical steady state that fluxes across the two control surfaces must be the same, assuming that there is no source or sink in the intervening region. At the current accuracy of observational estimation, that is probably all that can be said. On the other hand, if observational estimates were accurate for shorter timescales, then it would not be surprising to find imbalances between the two fluxes, reflecting, for example, seasonal variation of chemical tracer amounts in the lowermost stratosphere or seasonal variation in the mean position of the extratropical tropopause itself [e.g., *Robinson*, 1980].

5. OBSERVATIONAL ESTIMATES OF GLOBAL-SCALE EXCHANGE

We have seen in section 3 that the downward control principle (based on the zonal mean momentum budget; see (7)) can be used to provide a reasonable estimate of the extratropical meridional mean mass flow when the time constant of the motion is much longer than the radiative timescale. As an alternative, the zonally averaged thermodynamic energy equation (3) can be solved for $\bar{w}^*$ if the zonal mean potential temperature distribution and diabatic heating distribution are known. This calculation is, however, subject to considerable uncertainty in the lower stratosphere, where the net radiative heating is a small residual of several large terms. The downward control calculation also suffers problems probably owing to unresolved gravity wave forcing in the momentum budget and the time-dependent departures in the tropics and subtropics illustrated by Figure 4b.

Rosenlof and Holton [1993] have utilized the downward control approximation to estimate the contribution of each extratropical hemisphere to global-scale STE in the solstice seasons. They used a 10-year record of daily stratospheric geopotential height analyses from the United Kingdom Meteorological Office (UKMO) as a data source and the 100-hPa pressure surface (about 16-km altitude) as a control surface for the STE calculation, since this surface is approximately at the tropical tropopause (Figure 1).

The results of *Rosenlof and Holton* [1993] for the rate of mass transfer across the 100-hPa surface are shown in Table 1. They found that upwelling is limited to equatorward of ±15° latitude, except in the southern hemisphere winter, when it extends to 30°S. The downward flow across the 100-hPa surface in the extratropics is dominated by the strong meridional circulation of the northern hemisphere winter season,

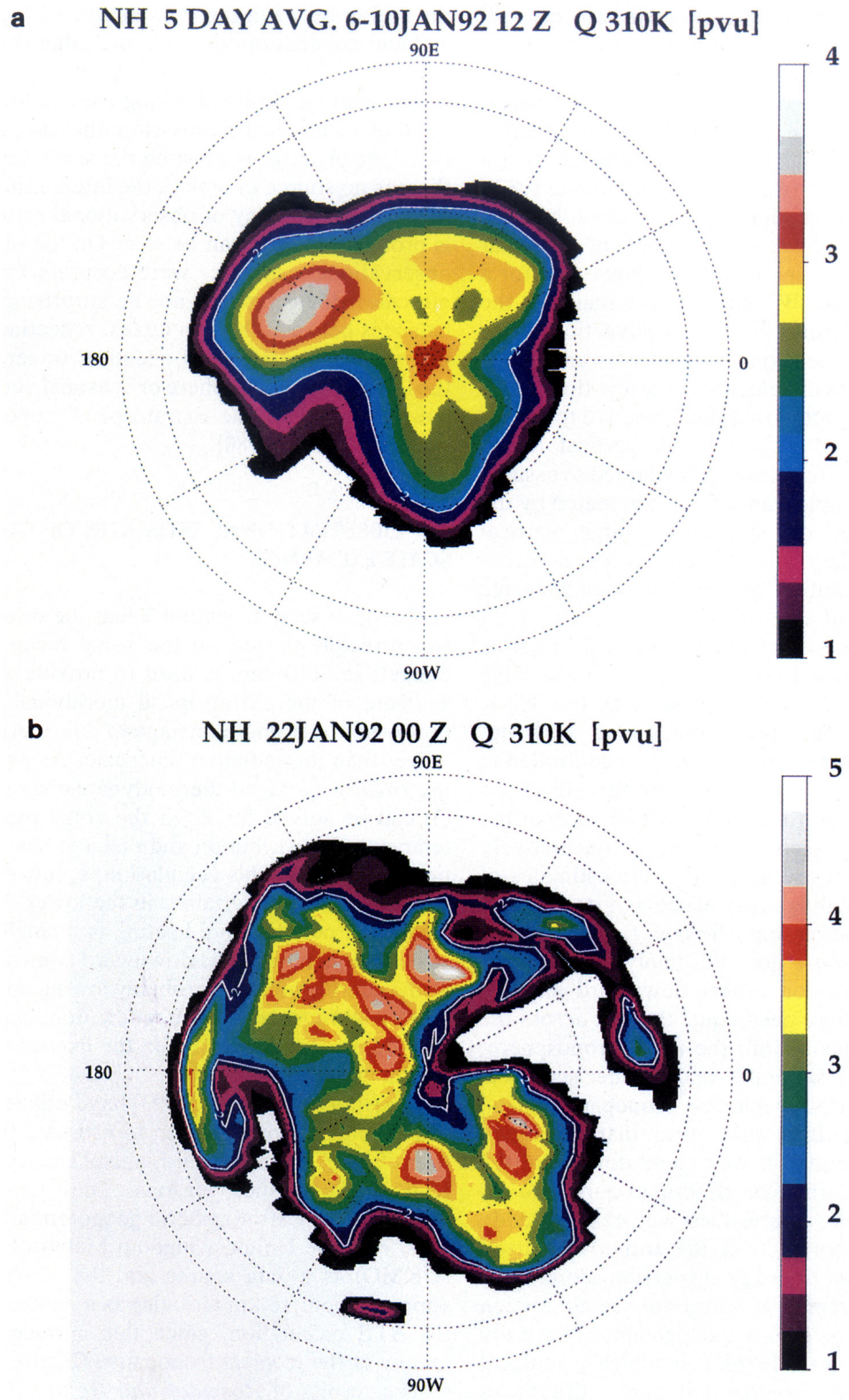

Plate 2. Potential vorticity distribution on the 310-K isentropic surface. The 2-PVU contour (shown in white) marks the approximate tropopause. (a) Five-day average of January 6–10, 1992, which shows the modulation of the tropopause latitude by the low-frequency Rossby waves. (b) The same field on January 22, 1992, showing strong meridional displacements of the tropopause associated with cutoff cyclone formation near the Greenwich meridian.

TABLE 1. Mass Flux Across the 100-hPa Surface From *Rosenlof and Holton* [1993]

	DJF	*JJA*	*Annual Mean*
NH extratropics	−81	−26	−53
Tropics	114	56	85
SH extratropics	−33	−30	−32

Units are 10^8 kg s^{-1}. Negative sign means downward flux. Abbreviations are DJF, December, January, and February; JJA, June, July, and August; NH, northern hemisphere; and SH, southern hemisphere.

which causes the compensating tropical upwelling to be about twice as large during the northern hemisphere winter as during the southern hemisphere winter.

If the tropical mass flux is assumed to be uniformly distributed between ±15° latitude, the vertical velocity at the 100-hPa level in northern hemisphere winter should be about 3×10^{-4} m s^{-1} greater than during southern hemisphere winter. Since upward motion causes adiabatic cooling in a statically stable environment, the annual cycle in the extratropical pumping should produce an annual cycle in tropical temperature near the 100-hPa level, with the lowest temperatures in the northern winter and the highest in the northern summer. Such an annual temperature cycle was found in conventional radiosonde data by *Reed and Vlcek* [1969] and has been confirmed in global satellite data by *Yulaeva et al.* [1994] using 12 years of data from channel 4 of the microwave sounding unit (MSU-4) flown on NOAA operational weather satellites. MSU-4 measures the mean temperature in a layer extending from about the 150- to the 40-hPa level, with a maximum response at about the 80-hPa level. Figure 6 shows the area-weighted annual march of temperature in the lower stratosphere from monthly mean MSU-4 data. In the tropics the minimum is in northern hemisphere winter and the maximum is in northern hemisphere summer, while in the extratropics the cycle is of similar amplitude but opposite phase. Because of the strong compensation between tropics and extratropics, the global mean temperature in this layer is nearly constant. Strong compensation between the tropical and extratropical signals is, of course, just what would be expected for a temperature signal produced by the adiabatic heating and cooling of an annually varying extratropically pumped meridional mass circulation, and provides strong support for the annual cycle in STE shown in Table 1 and for the global-scale circulation suggested in Figure 3.

The annual-mean tropical upwelling computed as an average of the two solstice seasons is equivalent to an upward mass flow across the 100-hPa surface of 2.7×10^{17} kg yr^{-1}. Since the total mass above the 100-hPa surface is 5.2×10^{17} kg, a notional turnover time of the order of 2 years is implied for the atmosphere above the 100-hPa level. Such a number must be viewed as only a rough approximation, however, since the atmosphere above the 100-hPa level cannot be regarded as a well-mixed layer.

The annual-mean global-scale mass exchange can also be estimated from the budget of a long-lived trace constituent such as the chlorofluorocarbon $CFCl_3$, which is inert in the troposphere but photodissociated in the stratosphere. Air parcels passing upward through the tropical tropopause will have $CFCl_3$ mixing ratios characteristic of the well-mixed troposphere; air parcels passing downward from the extratropical overworld into the lowermost stratosphere will have smaller mixing ratios owing to photodissociation during their passage through the overworld. The net flux of $CFCl_3$ into the overworld, F_c, must approximately equal the difference between the upward flux into the overworld in the tropics and the downward flux out of the overworld in the extratropics:

$$F_c = (F_{au}\chi_T - F_{ad}\chi_S)(m_c/m_a) \qquad (11)$$

Here F_{au} and F_{ad} are the upward and downward air mass fluxes, respectively, χ_T and χ_S are the mean volume mixing ratios of $CFCl_3$ transported upward and downward across the control surface defining the lower boundary of the overworld, and m_c/m_a is the ratio of the molecular mass of $CFCl_3$ to that of air. Conservation of mass requires that on average, $F_{au} = F_{ad}$, and the subtropical eddy-transport barriers in the overworld tend to keep the values χ_T and χ_S well apart.

Follows [1992] used observed values of χ_T and χ_S and an estimate of F_c deduced from data on the rate of release of $CFCl_3$ into the atmosphere and its rate of increase in the troposphere to solve (11) for the air mass flux F_{au}. To compare his results with those of *Rosenlof and Holton* [1993], we let the control surface

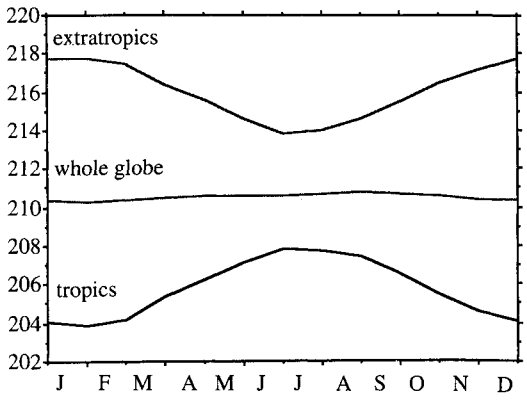

Figure 6. Mean annual march of lower stratospheric temperature based on MSU 4 data for the period 1979–1991, averaged over the tropics (30°S to 30°N), the extratropics (poleward of 30°S and 30°N), and the entire globe, showing evidence of control by extratropical pumping. (After *Yulaeva et al.* [1994]; copyright American Meteorological Society.)

in (11) be the 100-hPa level. From the vertical profile data of *Goldan et al.* [1980], $\chi_S \approx 0.6\chi_T$ at the 100-hPa level in the extratropics. If we use the estimate of *Rosenlof and Holton* [1993] that $F_{au} = 2.7 \times 10^{17}$ kg yr^{-1}, then (11) gives $F_c = 90 \times 10^6$ kg yr^{-1} as the net flux of $CFCl_3$ into the overworld. This agrees well with Follows's estimate of $F_c = 83\ (\pm 28) \times 10^6$ kg yr^{-1} and thus provides independent support for the downward control calculation.

The importance of carefully choosing an appropriate control surface when using conventional meteorological data to estimate global mass exchange is illustrated by the work of *Hoerling et al.* [1993]. Following *Wei* [1987], *Hoerling et al.* [1993] used an isentropic coordinate formulation in which the control surface was specified to coincide with the tropopause (defined approximately as in Figure 1), and the mass transport across this control surface was directly estimated from wind fields interpolated to isentropic coordinates. Although their estimated upward mass flux into the stratosphere in the tropics in January was surprisingly close to the December, January, and February (DJF) estimate of *Rosenlof and Holton* [1993], *Hoerling et al.* [1993] inferred that substantial transport into the stratosphere occurred across the tropopause at high latitudes. Such transport, if substantial, must be confined to a layer very close to the tropopause; otherwise it would not be consistent with the observed low water vapor mixing ratios in the lower stratosphere at high latitudes. This reinforces our theme that for most purposes the tropopause itself is not the most suitable control surface for study of mass exchange.

6. SMALLER-SCALE PROCESSES NEAR THE TROPOPAUSE: THE EXTRATROPICS AND SUBTROPICS

At the lower edge of the lowermost stratosphere, where the tropopause cuts across isentropic surfaces, STE can occur by isentropic transport, leading to displacement of the tropopause from its equilibrium position, as illustrated in Figure 2 and Plate 1, followed by nonconservative processes such as diabatic heating or cooling, layerwise two-dimensional turbulent mixing, and small-scale turbulent mixing. The boundary between stratospheric and tropospheric air along isentropes that span the tropopause is often marked by an undulating band of strong PV gradients (see Plate 2a) that coincides with the axis of an upper tropospheric jet stream. The existence of this band of strong PV gradients and indeed similarly strong gradients in mixing ratios of species such as ozone and water vapor, itself suggests that there must be rather strong dynamical resilience to transport along the isentropes, since otherwise vigorous mixing of stratospheric and tropospheric air would destroy the band of strong gradients.

The resilience to transport along the isentropes is provided by the Rossby-wave restoring mechanism, which causes the band of strong PV gradients to behave somewhat like an elastic band. Horizontal shear also plays an important role in limiting smaller-scale transports. In this sense the tropopause is much like the edge of the stratospheric wintertime polar vortex [e.g., *Juckes and McIntyre*, 1987; *McIntyre* [1992]. However, there are a number of important ways in which the tropopause is rather different from the stratospheric vortex edge. For example, displacements of the edge of the stratospheric polar vortex occur primarily through the upward propagation of Rossby waves from the troposphere. Large displacements of the tropopause, on the other hand, are generally associated with growing upper tropospheric cyclones. The distribution of horizontal shear is different, and the proportional jump across the tropopause is larger, of the order of unity, which is far larger than the jump in PV across the edge of the polar vortex, suggesting that quasi-geostrophic theory is highly limited in its applicability and that ageostrophic circulations have a significant role.

Exchange occurs when the tropopause is strongly distorted by large latitudinal displacements. Such distortions are characterized by tongues of anomalously high PV air, formed through the action of the large-scale deformation field, extending equatorward. Under some circumstances, some parts of these tongues may be stretched out into long filaments, and under others they may roll up to form isolated coherent structures containing high-PV air, generally referred to as "cutoff" cyclones (Figure 2 and Plates 1 and 2b). The ubiquity at midlatitudes of filamentary structures near the tropopause characterized by high-PV anomalies and likely to be associated with exchange processes is nicely shown by Meteosat water vapor images. An example was shown in Figure 2b (see also *Appenzeller and Davies* [1992]).

Cutoff cyclones have been argued to be important mechanisms of STE [e.g., *Bamber et al.* 1984]. Exchange in cutoff cyclones can occur by convective or radiative erosion of the anomalously low tropopause that is characteristic of cutoff cyclones, by turbulent mixing near the jet stream associated with the cutoff system, or as a result of tropopause folding along the flank of the system [*Hoskins et al.*, 1985; *Price and Vaughan*, 1993; *Wirth*, 1995]. However, Price and Vaughan conclude that convective exchange caused by cutoff cyclones is small compared to exchange caused by tropopause folding associated with upper level cyclogenesis. Cutoff cyclones, which can be substantial in size, have been successfully simulated and will be discussed further in section 8. Their subtropical counterparts, called "upper air cold vortices" by tropical meteorologists, must be presumed also to contribute to STE as suggested by the wavy arrow third from the bottom in Figure 3.

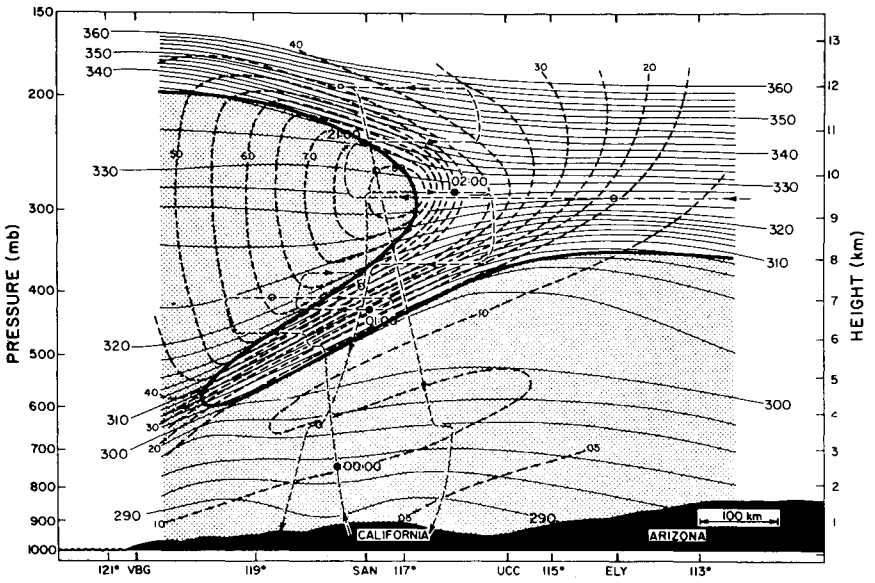

Figure 7. Cross section through strong three-dimensional tropopause-folding event of March 13, 1978, showing region of tropospheric air (stippled), potential temperature (thin solid contours), wind speed (in meters per second; dashed contours), research aircraft flight track (thin dashed lines), and potential vorticity tropopause (heavy solid line). (From *Shapiro* [1980]; copyright American Meteorological Society.)

The process of filamentation and the intrusion of stratospheric air deep into the troposphere, referred to as "tropopause folding," may also to some extent be regarded as the result of the systematic effect of the large-scale deformation field. There is some evidence to support this from contour-advection simulations of filamentary tropopause structure in reasonable agreement with the Meteosat water vapor images (Plate 1). In this sense again, the tropopause is similar to the edge of the stratospheric polar vortex, where vertical variations in the quasi-horizontal velocity field are expected to give filamentary, or sloping, sheet-like structures with small vertical and horizontal scales. However, for the midlatitude tropopause the effects of the large-scale deformation can be enhanced by ageostrophic frontogenetic circulations; indeed, in simple two-dimensional models of upper tropospheric frontogenesis [e.g., *Hoskins*, 1971], the ageostrophic circulation is entirely responsible for the appearance of small vertical scales, since the applied large-scale deformation flow is height independent. Stretching of stratospheric intrusions to ever finer scales leads to irreversible transport, often speeded up by turbulence resulting from shear instabilities. Much of the ozone transport from the lowermost stratosphere into the troposphere is believed to occur in connection with such tropopause fold events. There is a long history of case studies of midlatitude tropopause folding, which is reviewed by *WMO* [1986]. More recent examples are given by *Browell et al.* [1987] and *Kritz et al.* [1991]. These studies have confirmed that large episodic STE can occur in association with tropopause folding.

The structure and dynamics of the folding process associated with upper level frontogenesis was reviewed by *Keyser and Shapiro* [1986]. More recently, high-resolution mesoscale numerical models have been used to simulate tropopause folding associated with upper level baroclinic wave development [*Bush and Peltier*, 1994; *Rotunno et al.*, 1994; *Lamarque and Hess*, 1994]. The first two of these papers stressed the fact that the most extensive tropopause folding occurs in three-dimensional baroclinic developments associated with very strong jet streams. As the baroclinic wave development proceeds near the tropopause, strong subsidence develops in the region of northerly flow to the west of the upper level trough, together with downward tilt of the isentropes and intrusion of air of stratospheric origin into the troposphere as shown in Figure 7.

The quasi-isentropic transport of air out of the lowermost stratosphere initiated by tropopause folding could in principle occur in the absence of the global-scale diabatic circulation. But in that case, there would necessarily be on average an equal quasi-isentropic reverse transport of air from the troposphere into the lowermost stratosphere, in order to maintain mass balance. If this were the case, careful observations would be required to monitor the mesoscale correlations between trace constituent mixing ratios and air motions in order to estimate the net exchange of a given constituent. Although some studies [e.g., *Reiter et al.*, 1969; *Ebel et al.*, 1991; *Hoerling et al.*, 1993] suggest that irreversible transfer of tropospheric air into the stratosphere can occur in association with tropopause folding (and indeed, some two-way exchange is suggested by visualizations like Plate 1 and other Lagrangian tracer studies such as that of *Yang and Pierrehumbert* [1993]), the extremely low water vapor mixing ratios observed throughout the stratosphere [e.g., *Oltmans and Hofmann*, 1995] indicate that the quasi-isentropic exchange in middle latitudes is mostly a one-way transport into the troposphere, except perhaps in the lowest kilometer or so of the stratosphere.

The global-scale diabatic circulation is, of course,

required to transport stratospheric constituents such as ozone downward from the overworld to the lowermost stratosphere. The average rate at which such a species can be transported into the troposphere is thus ultimately determined by the rate at which the dynamically controlled global-scale circulation transports mass from the overworld into the lowermost stratosphere. For this reason the details of mesoscale tropopause-folding events may not be important for determining the global flux of ozone and other tracers from the stratosphere, although they will certainly strongly influence the time and space distribution of such transport.

Recent developments suggest, however, that the time and space distributions may be important for some chemical purposes and that some caution is required; for instance, much of the relationship between global-scale dynamics and photochemistry was developed before it was known that heterogeneous chemical processing occurring on lower stratospheric aerosols was important. Within the old framework of gas phase chemistry, most photochemical activity was confined to the upper stratosphere, so that the global transport in the overworld provided a good average representation of transport and an adequate constraint on STE. However, stratospheric reactions on aerosol surfaces may occur in the lowermost stratosphere. Furthermore, new initiatives to evaluate the impact of aircraft on the environment require the consideration of longitudinally asymmetric emissions directly in the lowermost stratosphere. For these problems, it is necessary to at least evaluate the sensitivity to the more detailed synoptic-scale approach of any conclusions that are drawn in chemical assessment studies.

7. SMALLER-SCALE PROCESSES NEAR THE TROPOPAUSE: THE TROPICS

As discussed in sections 3–5, the theory of the wave-driven circulation, illustrated in Figure 3 and confirmed and calibrated by observational knowledge of the relevant eddy effects, tells us that the extratropical stratosphere and mesosphere act nonlocally on the tropics as a global-scale fluid-dynamical suction pump. The pumping action is slow but inexorable and causes large-scale upward transfer of mass from the tropical troposphere into the tropical stratosphere, at a rate that is largely independent of local conditions near the tropical tropopause. Indeed, on the contrary, as was pointed out by *Yulaeva et al.* [1994] and by *Rosenlof* [1995], the local conditions must respond to the pumping rate, insofar as stronger extratropical pumping must tend to pull tropical upper tropospheric and lower stratospheric temperatures below radiative equilibrium, encouraging deep cumulonimbus and producing higher and colder tropopauses. This helps to explain the observed seasonal-cycle behavior illustrated in Figure 6.

The existence of the global-scale circulation itself is confirmed by chemical tracer evidence, now including some new and very precise evidence from the Halogen Occultation Experiment (HALOE) to be described shortly. Older, well-known evidence begins with the fact that long-lived chemical species with mixing ratios nearly uniform throughout the troposphere, such as methane, nitrous oxide, and the CFCs, (1) are observed to have their largest stratospheric mixing ratios immediately above the tropical tropopause, matching tropospheric values, and (2) show upward extending plumes of relatively large mixing ratios throughout the depth of the tropical stratosphere, consistent with upward transport in this region and a degree of isolation from the extratropics. Such isolation is, in turn, understandable from the existence of the subtropical eddy transport barriers mentioned in section 4.

Also well known are the facts that water vapor shows a similar upward extending plume, but that in contrast with the aforementioned tracers, very low values of the water vapor mixing ratio extend upward from a minimum within a few kilometers of the tropical tropopause. That minimum mixing ratio is only a few parts per million by volume (ppmv) and far lower than typical tropospheric values. It is generally accepted, with good reason, that the basic mechanism responsible for the low values is freeze-drying [*Brewer*, 1949], a process in which air passing through the tropical tropopause has its water vapor mixing ratio reduced to the ice saturation value at or near the tropopause. Satellite data since 1979 have broadly confirmed this classical global-scale picture of air entering the tropical stratosphere from below, being dehydrated as it enters, and then being gradually moistened, by methane oxidation and perhaps by weak mixing from the extratropics, as it is drawn upward in the tropics before being pushed poleward and eventually downward in middle and high latitudes [e.g., *Remsberg et al.*, 1984; *McCormick and Veiga*, 1992; *Carr et al.*, 1995]. The satellite data show the lowest zonal-mean water vapor mixing ratios in the tropical lower stratosphere during northern hemisphere winter (about 2–3 ppmv), with values increasing with altitude and latitude to 4–6 ppmv. It should be noted that most satellite measurements in the lower tropical stratosphere are highly uncertain because of cloud particle and aerosol contamination, especially for the 2 or 3 years following major volcanic eruptions. Also, the sharpness of tropical temperature and water vapor minima are difficult to determine with satellite measurements that resolve only vertical scales greater than 2–3 km.

Now a likely implication of the freeze-drying hypothesis is that water vapor mixing ratios of air entering the tropical stratosphere should vary seasonally in phase with the annual cycle of tropical tropopause

temperature. Thus, just as a signal is recorded on a magnetic tape as it passes the head of a tape recorder, the minimum saturation mixing ratio near the tropical tropopause should be "recorded" on each layer of air moving upward in the large-scale tropical stratospheric circulation. Since, as was mentioned in section 4, horizontal eddy transport by isentropic mixing between the tropics and extratropics is observed to be relatively weak, layers of air passing upward through the tropical tropopause should retain their water vapor signal for some time, perhaps many months, as they are advected upward by the large-scale circulation. With sufficient vertical resolution, the imprint of the seasonally varying saturation mixing ratio of the tropical tropopause should then be observable in a time-height section as alternating layers of low and high water vapor mixing ratio anomalies that move upward at the speed of the large-scale circulation.

This "tape recorder effect" has now been strikingly confirmed by analysis of water vapor mixing ratio data from the microwave limb sounder (MLS) and the Halogen Occultation Experiment (HALOE), both on the Upper Atmosphere Research Satellite (UARS) [*Mote et al.*, 1995a, b]. The data provide new and powerful evidence on two points. The first is the confirmation of large-scale upwelling speeds and their seasonal variation, entirely independent of the three observational approaches to global exchange rates summarized in section 5. The second is the remarkably long persistence of the "tape signal," showing only weak isentropic mixing into the tropics during the period in which adequate data were available from HALOE, January 1993 to October 1994, and roughly in the altitude range 20–25 km. The annually varying mixing-ratio signal produced by the freeze-drying process remains noticeable for as long as 2 years over a still larger altitude range of 15 km, well into the middle stratosphere. The implied ascent rates just above the tropical tropopause are approximately consistent with the calculated mass fluxes of Table 1, varying from about 0.2 mm s^{-1} in northern summer to about 0.4 mm s^{-1} in northern winter.

The HALOE "total hydrogen" ($2[CH_4] + [H_2O]$) data are of special interest here. For well-known chemical reasons connected with methane-oxidation chemistry, total-hydrogen mixing ratios are to good approximation uniform in the extratropical overworld, with values close to, or weakly fluctuating about, the time-averaged tape signal. It follows that the total-hydrogen signal can be due only to the "recording head" near the tropopause and that, in the overworld, isentropic entrainment or mixing in of extratropical air can hardly do other than attenuate the signal. Thus the total-hydrogen results provide new and very direct confirmation of, and perhaps the first quantitative information about, the weakness of isentropic entrainment from the extratropics; *Mote et al.* [1995b] give more detail. Here it is important to note, first, that eddy transport barriers cannot be described quantitatively in terms of eddy diffusivities [e.g., *McIntyre*, 1992, p. 364] and, second, that they can behave asymmetrically. As is clearly demonstrated by the more familiar example of the polar vortex edge, rates of entrainment or mixing in need not be comparable to rates of erosion or mixing out. Further information on these important questions is now emerging from recent airborne measurement campaigns (K. A. Boering, R. A. Plumb, and S. C. Wofsy, personal communication, 1995).

The remaining questions about the tape recorder analogy concern the amplitude of the observed tape signal in water vapor mixing ratios. Here our understanding is far from complete, but *Mote et al.* [1995b] suggest that the subtropical barrier is relatively permeable near the base of the overworld (consistent with classical bomb-debris observations), is less permeable a few kilometers farther up, where mixing in is minimal, and is then more permeable again in the middle stratosphere, especially during strong westerly quasi-biennial oscillation (QBO) episodes, as in January–February 1993. Such a pattern is consistent with what we know of horizontal eddy motion extending into the base of the overworld from the subtropical troposphere [e.g., *Hoskins et al.*, 1985, equation (33d), Figures 2a, 10, 15]; in this regard, Figure 3 is oversimplified. The mixing near the base of the overworld greatly attenuates the signal relative to what a simple estimate from 100-hPa synoptic-scale temperature analyses would imply, and it is now especially clear that dehydration arguments based on such data must be treated with caution, for reasons given in the classic "stratospheric fountain" work of *Newell and Gould-Stewart* [1981] and for additional reasons now to be summarized.

The tape recorder analogy, with seasonally varying tape speed and with allowance for mixing, helps to make better sense, retrospectively, of many past observational studies, including the in situ ER-2 aircraft measurements from the field campaigns based at Panama, Central America and at Darwin, northern Australia. This important pair of campaigns, to be discussed in more detail below, took place during northern hemisphere summer and winter, respectively, and showed the striking differences in the average water vapor mixing ratio profiles summarized in Figure 8 [*Kelly et al.*, 1993]. Figure 8 shows that the average lower stratospheric water vapor minimum in the northern hemisphere winter series (2.5 ppmv) occurs near the tropopause (10 flights from Darwin). However, the average minimum in the northern hemisphere summer series (3.5 ppmv) occurs near 19 km, which is well above the tropopause (7 flights from Panama). This dry "hygropause" layer cannot be explained in terms of freeze-drying by transport through the local tropopause at the same time of year, northern

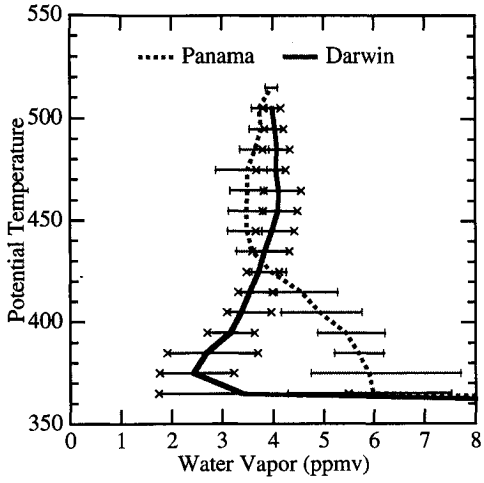

Figure 8. Panama (northern hemisphere summer, 1980) and Darwin (northern hemisphere winter, 1987) mean water vapor profiles measured by the NASA ER-2 research aircraft. The vertical axis is marked in units of potential temperature. Bars indicate the minimum and maximum observed values for individual flights. The northern summer minimum at Panama is about 75 K (~3 km) above the local tropopause, while the northern winter minimum at Darwin is coincident with the tropopause (From *Kelly et al.* [1993].)

summer [e.g., *Holton*, 1984]. However, the tape recorder analogy provides a basis for interpreting this elevated hygropause, to a first approximation, as a consequence of upward advection, during the intervening months, of low saturation mixing ratios that were near the tropopause during the northern winter, probably modified by mixing effects.

As explained in section 5, the extratropically controlled upward mass transport into the tropical stratosphere is believed to be about twice as large in northern hemisphere winter as in northern hemisphere summer, and regionally averaged tropopause temperatures over the northern Australia region during January do appear to be low enough to dehydrate air to prevailing minimum stratospheric values [*Selkirk*, 1993]. A uniform rising motion would, however, be expected to cause water vapor saturation and formation of a widespread thick cirrus cloud layer near the tropopause. *Robinson* [1980] and others suggested that such cloudiness is not observed and therefore that upward flow across the tropical tropopause must be associated entirely with the mass flow from tropical deep convection that penetrates into the lower stratosphere. How such mass flows could occur and lead to stratospheric dehydration was discussed by *Danielsen* [1982], in terms of small-scale turbulent entrainment into overshooting cumulonimbus turrets, followed by particle sedimentation within anvils, leaving sheets of dry air behind.

It is possible that such cumulonimbus penetrations could "hyperventilate" the lower tropical stratosphere, in the sense that there could be a greater convective mass flux than can ultimately be taken up by the extratropically controlled large-scale upwelling. *Mote et al.* [1995b] argue that such hyperventilation is most likely in the northern summer monsoon season, when surface conditions tend toward strengthening and deepening the cumulonimbus convection but the global-scale upwelling is slowest. Such hyperventilation would have to introduce a statistical bias toward colder, dryer dehydration conditions than any synoptic-scale temperature analysis would indicate. This is because in the tape recorder analogy the deepest penetrations act like recording heads that are located farthest along the direction of tape travel. The dehydration signals from such recording heads have the greatest chance of surviving, that is, of not being erased or displaced back downward or sideways by subsequent higher injections as the notional tape moves upward. Mote et al. point out that hyperventilation could therefore be a contributing factor in determining the amplitude of the observed tape signal in total hydrogen. It would reinforce the effects of isentropic entrainment or mixing in during northern summer and work against them during northern winter.

The Panama and Darwin campaigns both used the high-altitude ER-2 research aircraft and aimed at testing by in situ observation whether, on the scale of cumulonimbus anvils and their immediate surroundings, dehydration took place in the way suggested by *Danielsen* [1982]. The campaigns were joint efforts of the National Aeronautics and Space Administration and the National Oceanic and Atmospheric Administration. The Panama campaign, the Water Vapor Exchange Experiment (WVEE) of September 1980, was the first aircraft investigation of tropical convective exchange. The fundamental measurements in this experiment were temperature, pressure, water vapor (using a new Lyman-Alpha hygrometer capable of measuring very low mixing ratios of water vapor [*Kley et al.*, 1982], and ice particles [*Knollenberg et al.*, 1982]. The experimental plan involved sampling air both within and downstream of convectively produced anvils so that the nature and evolution of the convectively influenced air could be ascertained.

The WVEE observations demonstrated that convective systems do penetrate into the stratosphere and do mix tropospheric and stratospheric air together [*Danielsen*, 1982], so that the resulting mixture can have stratospheric values of potential temperature, $\theta > 380$ K, despite the maximum observed tropospheric equivalent potential temperature of 367 K [*Selkirk*, 1993]. Evidence for such penetration was shown by observations in a convectively produced anvil where high ice crystal amounts (several hundred ppmv of total water substance) were found in an imperfectly

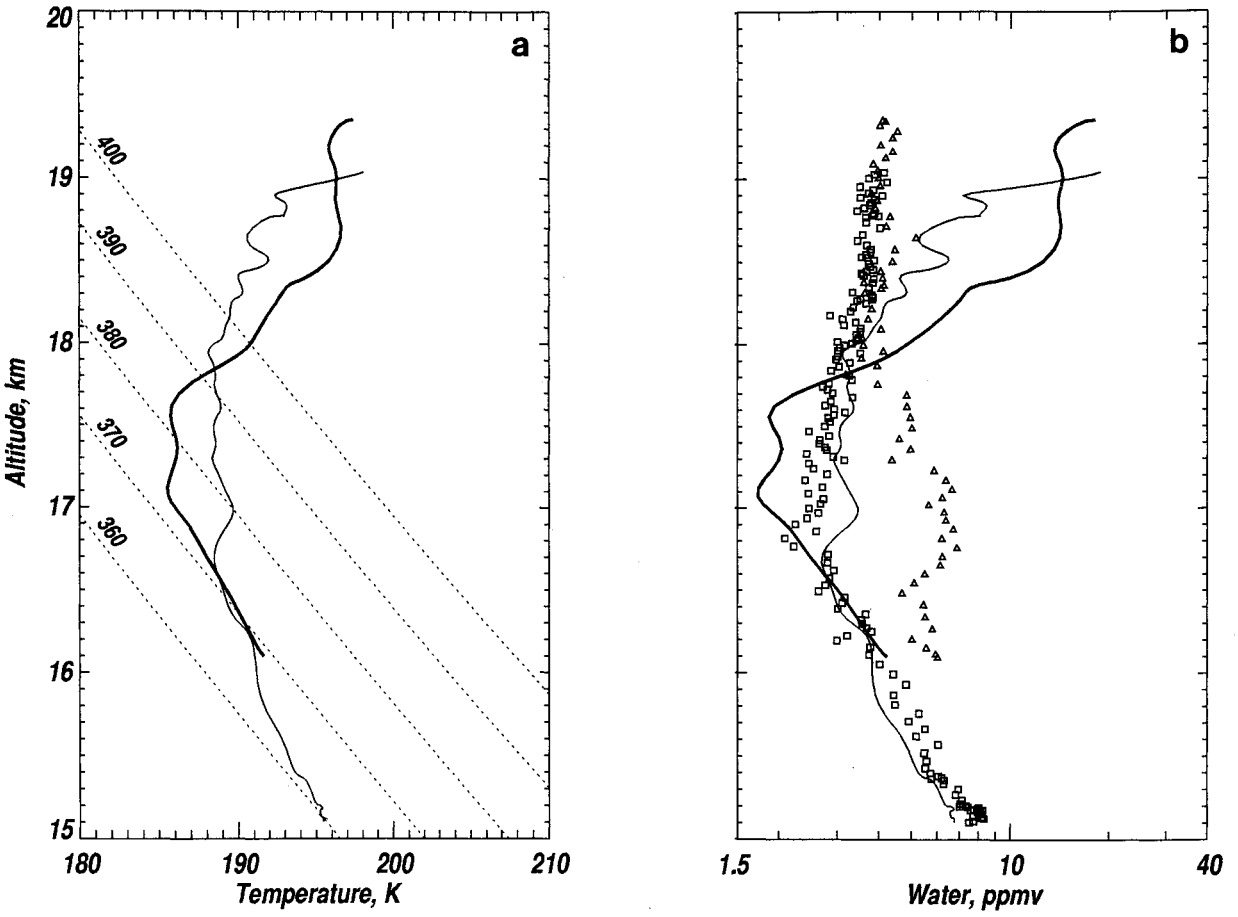

Figure 9. Vertical profiles from the flight of January 23, 1987, during STEP/Tropical. (a) Temperature at the edge of the anvil (heavy solid line) and outside the anvil (light solid line). (b) Total water and ice saturation mixing ratio at the edge of the anvil (triangles and heavy solid line, respectively) and outside the anvil (squares and light solid line, respectively).

mixed layer between 16 and 17 km. The air in this layer was clearly stratospheric, with a potential temperature of about 385 K. The observations also showed that the ice crystals in the mixed layer (16–17 km) appeared to fall out, leaving a mixture of stratospheric and tropospheric air with a water vapor content of about 6 ppmv, essentially the saturation mixing ratio for temperatures found at the top of the convective system (about 193 K at 17 km, or about 2.5° colder than the seasonal mean at Panama). Finally, the particle measurements [*Knollenberg et al.*, 1982] showed that the bulk of the ice in the sampled anvil was in particle sizes of 100 μm or more, implying that the ice would fall 1 km in less than 1 hour. Thus the evolution from the anvil center (with substantial excess ice) to the anvil edge (with no excess ice) is consistent with the particle data.

No cumulonimbus penetration events were observed, however, that reached above 450 K and produced the minimum Panama mixing ratios associated with the dotted curve in Figure 8. This prompted the second campaign, the tropical experiment of the Stratosphere-Troposphere Exchange Project (STEP/Tropical), based in Darwin, Australia, during January and February 1987. The aim was to deploy the ER-2 aircraft in the "stratospheric fountain" region of the Indonesian maritime continent [*Newell and Gould-Stewart*, 1981], where seasonally averaged temperatures at 100 hPa were estimated to be as low as 188 K, implying saturation mixing ratios of 2.4 ppmv. This is already at the low end of climatological-mean observed water vapor mixing ratios, 2.5–3.5 ppmv, for the tropical lower stratosphere. Furthermore, if the coldest cloud tops were 2.5° colder than the mean at 100 hPa, then moist convection could conceivably inject air with mixing ratios as low as 1.6 ppmv.

That extreme conditions of this order can and do occur in reality was verified by one of the flights from the Darwin campaign. Figure 9 shows the evolution of an anvil from a large continental thunderstorm near the south end of the Gulf of Carpentaria. The heavy solid lines and triangles show the anvil penetration. Here the presence of both measurable radon gas (not shown) [*Kritz et al.*, 1993] and total water exceeding the ice saturated mixing ratio [*Kelly et al.*, 1993] indicates that tropospheric air has mixed upward to 17.9

km with clearly stratospheric potential temperatures of 390 K. Atmospheric radon has its source at the Earth's surface, and its atmospheric lifetime is governed by its 3.8-day radioactive half-life. The lowest temperatures, with a potential for dehydration to 1.7 ppmv, were found between 17 and 17.6 km (potential temperatures of 372 K and 383 K). All of this air, including the air in the quasi-adiabatic mixed layer below 17 km, has stratospheric characteristics as well, as demonstrated by the elevated potential temperatures. It also has ozone mixing ratios of 100 parts per billion by volume (ppbv) or higher [*Danielsen*, 1993], most likely showing isentropic entrainment from the extratropics and suggesting why actual mixing ratios (squares in Figure 9b) are not quite as low as 1.7 ppmv.

Thus as in WVEE, tropospheric air was observed to have penetrated and mixed into the stratosphere. The major difference is that the cloud tops are colder and somewhat higher, with minimum saturated mixing ratios of 1.7 ppmv (light solid curve in Figure 9b) instead of around 6 ppmv. The light solid lines and squares in Figure 9 show a sounding outside the anvil, northwest and roughly (though not exactly) downstream of the heavy solid lines. Some warming in the 16- to 18-km layer is evident (mostly attributable to adiabatic descent [see *Kelly et al.*, 1993], but the excess ice has fallen out, leaving air with minimum water mixing ratios of 2.2 ppmv (squares in Figure 9b) at a potential temperature of 372 K, again, clearly stratospheric. Measurable radon is present at this location as well, clearly demonstrating the following: (1) Tropospheric air has penetrated into the stratosphere and been mixed irreversibly with stratospheric air, (2) excess ice has fallen out, and (3) the resulting air has water vapor mixing ratios similar to or below the climatological-mean values of 2.5–3.5 ppmv.

The STEP/Tropical experiment thus clearly showed that convective exchange in the tropics can transport tropospheric air upward and dehydrate it at the same time. However, it is not clear from the data gathered on the remaining flights whether penetration events like that of Figure 9 are frequent enough to dominate the tropical stratospheric mass and water vapor budget, or whether they are "too frequent" in the sense that hyperventilation might be taking place. The remainder of this section discusses some of the questions thus raised.

Before a budget calculation or hyperventilation estimate could be undertaken based on, for instance, global satellite coverage, one would need to know whether the mechanism outlined by *Danielsen* [1982] and illustrated in Figure 9 is representative of deep moist convection that reaches and overshoots the tropopause. In fact, the cloud top characteristics of the bulk of the moist convection observed during STEP/Tropical differed substantially from Figure 9. The storm corresponding to Figure 9 occurred over the Australian continent during a period of suppressed moist convection in the Australian region [*Russell et al.*, 1993]. By contrast, most of the moist convection during STEP/Tropical occurred near the northern coast during the active periods of the monsoon. Moreover, averaged over the month of January 1987, the lowest saturation mixing ratios from radiosonde measurements occurred not over the continent, but over the northern Australian coast [*Selkirk*, 1993]. An important feature of this coastal moist convection was its lesser penetration into the stratosphere, as evidenced by the much lower (10–12 K) tropopause potential temperatures.

Figures 10a to 10c show three additional examples of soundings at cloud tops taken during the STEP/Tropical experiment. Along with Figure 9, they show the diversity of moist convection during the Australian monsoon (see *Danielsen* [1993] for additional examples). Figure 10a is from a convective system over Darwin observed soon after monsoon onset, when tropopause and near-tropopause temperatures (as measured by radiosondes) were at their lowest points of the winter of 1987 [*Selkirk*, 1993]. The minimum saturated mixing ratios are comparable to or lower than those in Figure 9 (1.6 ppmv), but they occur in what is essentially pure tropospheric air (ozone mixing ratios of 10 ppbv and potential temperatures of 360 K). If these clouds are to contribute substantially to the stratospheric mass budget, the air in these anvils must attain stratospheric potential temperatures, either by turbulent diffusion at the anvil top [*Danielsen*, 1993] or by radiative heating and subsequent rise [*Danielsen*, 1982; *Ackerman et al.*, 1988]. Though there is substantial theoretical evidence for enhanced radiative heating of tropical anvils, the observational evidence is still quite inconclusive [*Russell et al.*, 1993]. (Of course, radiative heating is indeed what would occur if the stratosphere were "underventilated," with insufficient mass injection, in which case the extratropical pump would simply and inexorably pull air upward on a large scale. As was already pointed out, the absence of a large-scale cirrus veil can be taken as evidence against such underventilation.)

Figure 10b is from a flight north of Darwin about 2 weeks later, near the beginning of a second active period during the 1987 Australian monsoon. Tropopause radiosonde temperatures on the northern Australian coast are about 4 K higher than during monsoon onset, and the aircraft measurements are clearly consistent with this. The moist convection clearly penetrates into the stratosphere (as is shown by the excess ice), but only to potential temperatures of about 370 K. Moreover, this system is clearly hydrating the stratosphere below 16.3 km, with saturated mixing ratios of 3.2–4 ppmv.

Figure 10c is from the northern end of the Gulf of Carpentaria over a tropical cyclone. A notable feature is the deep region of negative (but clearly subadiabatic) vertical temperature gradients below 17.9 km,

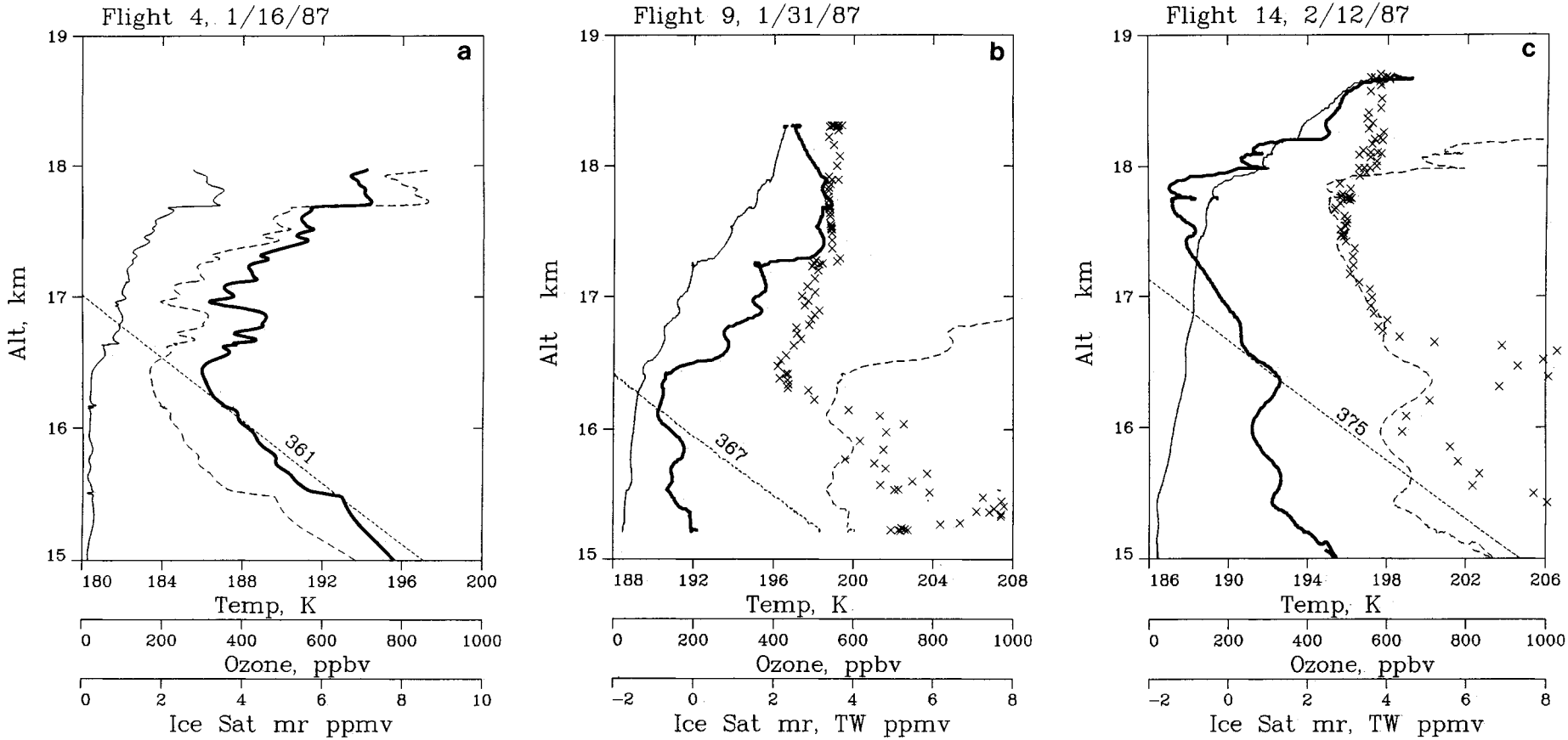

Figure 10. (a) Vertical profile through an Australian monsoon onset convective system from the flight of January 16, 1987. The heavy solid line denotes temperature, in kelvins, the light solid line denotes ozone, in parts per billion by volume; and the dashed line denotes ice saturation mixing ratio, in parts per million by volume. The dotted line is the 361-K isentrope. (b) Vertical profile through a midmonsoon convective system from the flight of January 31, 1987. Heavy solid line denotes temperature in kelvins; light solid line denotes ozone, in parts per billion by volume; dashed line denotes ice saturation mixing ratio, in parts per million by volume; crosses denote total water mixing ratio, in parts per million by volume. The dotted line is the 367-K isentrope. (c) As in Figure 10b, except for a tropical cyclone from the flight of February 12, 1987. The dotted line is now the 375-K isentrope.

which is capped by a strong inversion. The fact that the entire region is saturated with respect to ice is evidence for anvil penetrations to 17.9 km. In fact, ice was observed over the same tropical cyclone 4 days earlier at 18.3 km [*Kelly et al.*, 1993; *Pfister et al.*, 1993]. A succession of overshooting and detraining convective turrets have here produced a region of weak stability that is a mixture of stratospheric and tropospheric air. The remains of one such turret are apparent at 16.5 km, with marked ice crystal loading. These systems are clearly able to dehydrate the tropical stratosphere, since the minimum saturated mixing ratio is about 2.5 ppmv.

The most important implication of this variety of convective cloud tops for tropical exchange and dehydration is that it complicates attempts at global bulk calculations of water vapor and mass budgets. Though all the types of observed moist convection produce upward transfer of some sort, the continental systems are probably most effective on an individual basis. Currently available cloudiness statistics are hardly adequate for estimating the spatial and temporal distributions of near tropopause cloud mass fluxes and hence the degree to which convection hyperventilates the lower tropical stratosphere. An additional problem involves ice crystal fallout. No very large ice crystals (>100 μm in diameter) were observed during STEP/Tropical. In fact, within the regions of high ice crystal content (200 ppmv of ice or more), perhaps 2% of the mass was in sizes of less than 10 μm [*Knollenberg et al.*, 1993]. These crystals fall very slowly, taking over a day to fall 1 km. Thus given the possibility of significant anvil heating rates [*Ackerman et al.*, 1988], there is the real possibility of evaporation of these ice crystals. Though 2% is a small percentage, this could account for 4 ppmv or more of ice, certainly significant for the stratospheric water budget, particularly if there is substantial hyperventilation.

The difficulty of quantitatively accounting for observed tropical stratosphere dehydration based on the penetrative cumulus convection model led *Potter and Holton* [1995] to propose an alternative scheme in which convectively generated buoyancy waves provide an important role in dehydration. Vertical parcel displacements produced by such waves can produce local saturation and promote the formation of thin ice clouds in the lower tropical stratosphere (analogous to wave clouds in the lee of mountains). Such clouds would be expected to form in a quasi-random fashion in the vicinity of tropical moist convection. Even if typical ice particle sizes are very small so that particle sedimentation rates are slow, the net result averaged over many realizations of such clouds could be a significant downward flux of water, which would tend to dehydrate the lower stratosphere and thus contribute to the maintenance of the observed water vapor minimum.

Since particles that form as a result of cooling over several days rather than several hours tend to be somewhat larger [*Jensen and Toon*, 1994], variations in tropopause temperatures on longer timescales may be even more promising for dehydrating the tropopause region. One possibility is the tropical Kelvin wave having a timescale of about a week. (See *Andrews et al.* [1987] for a review of Kelvin wave properties.) *Tsuda et al.* [1994] have documented the passage of such a Kelvin wave over Indonesia and noted a temperature variation from 185 to 193 K, implying a saturation mixing ratio variation from 1.8 to 4.8 ppmv. If the particles that are formed are large enough to fall significantly in a few days, this is certainly a potential water removal mechanism near the tropical tropopause. Similar slow cooling mechanisms are responsible for dehydration in the Antarctic winter stratosphere [*Toon et al.*, 1989].

The WVEE and STEP/Tropical experiments showed that cumulonimbus penetration and dehydration events can take place not only during northern winter, but also even in northern summer when tropopause temperatures are nearly 5 K higher than in northern winter, with correspondingly higher saturated mixing ratios. This argues against the original suggestion of *Newell and Gould-Stewart* [1981] that such penetration is strictly confined to the maritime continent in northern winter. Though, as is discussed in section 5, upward motion in the tropics is only about half as strong in northern summer as in the northern winter because of weaker extratropical pumping from the southern winter hemisphere [*Rosenlof and Holton*, 1993], there is still considerable moistening at the tropopause, as is shown in Figure 8. As was discussed earlier, this moistening, combined with the upward lifting of tropical air dehydrated during the previous northern winter (the tape-recorder effect), can explain the observed lifted hygropause shown for Panama in Figure 8. The poleward spread of the dehydrated air is also the most likely explanation for the observed annual minimum in water vapor mixing ratios in the midlatitude southern hemisphere lower stratosphere during March–July [*McCormick et al.* 1993]. In fact, significant southward spreading of dry air from the Australian monsoon is evident from STEP/Tropical, where a 100-hPa anticyclone clearly transports air from just above the tropopause in the northern Australian monsoon region to the midlatitude jet over the Southern Ocean [*Danielsen*, 1993; *Kelly et al.*, 1993; *Kritz et al.*, 1993].

8. STRATOSPHERE-TROPOSPHERE EXCHANGE IN MODELS

In this review we have emphasized that STE must be regarded as part of the general problem of global transport and that nonlocal dynamical forcing (extratropical pumping) plays a primary role in driving global

transport. In section 5, observational evidence was presented in support of the extratropical-pumping model of transport and STE. Because much of the extratropical pumping is generated by Rossby waves, general circulation models (GCMs) that correctly simulate the generation, propagation, and dissipation of Rossby waves should correctly represent the global-scale aspects of transport and STE. Thus GCMs represent a valuable tool for testing the conceptual model of transport and STE presented above.

It is also reasonable to ask whether the two-dimensional (longitudinally averaged) models commonly used in chemical assessment studies properly represent global transport and STE. The nonlocally driven global-scale diabatic circulation (either externally specified or determined internally) is used as the basis for the advective transport in most two-dimensional models. Thus such models should adequately represent the global-scale mass exchange across the lower boundary of the overworld. Any other contribution to STE in such models, such as by isentropic mixing in the lowermost stratosphere, can only be included via the eddy-flux parameterization.

Three-dimensional models will include other contributions to STE through explicit simulation of planetary-scale and synoptic-scale eddies, but at present it is only those with relatively high resolution that can resolve features such as tropopause folds. For example, *Cox et al.* [1995] report a recent simulation of a fold in a global GCM, with horizontal resolution of about 100 km and vertical resolution of about 1.5 km near the tropopause. Coarser-resolution global models will represent exchange in such cases, largely through parameterized mixing. Mesoscale (limited area) models may not be so reliant on parameterization and may be used for detailed simulation of local processes (see section 6), but it is difficult to put the results from such models in a global context. This section will explore further how two- and three-dimensional models actually represent STE.

Global constituent transport in general circulation models has been studied for more than a decade, particularly at the Geophysical Fluid Dynamics Laboratory at Princeton University [e.g., *Mahlman et al.*, 1980, 1986; *Plumb and Mahlman*, 1987; *Strahan and Mahlman*, 1994a, b]. However, there have been only a few general circulation model studies that specifically focused on STE. An early example was the study of *Allam and Tuck* [1984a, b], who attempted to simulate the transport and exchange of water vapor. Their efforts were hampered both by the limited resolution of their model and by its unrealistically large rate of diffusion by subgrid-scale processes. More recently, *Mote et al.* [1994] have used the National Center for Atmospheric Research community climate model 2 (CCM2), a model with much better resolution in the region of the tropopause and better numerical treatment of tracer transport, in a detailed study of model-simulated STE.

The CCM2 qualitatively simulates the stratospheric transport driven by extratropical pumping. However, the simulated mass transport across the 100-hPa surface is too strong, while the transport at higher altitudes is too weak, consistent with the fact that the zonal forcing in the model is too strong in the lower stratosphere and too weak in the upper stratosphere and mesosphere. Nevertheless, the CCM2 captures the correct seasonal variation in exchange; there is a maximum in transport across the 100-hPa surface in northern hemisphere winter and a minimum in northern hemisphere summer.

The simulation of water vapor in the CCM2 is not as satisfactory as is the simulation of air mass exchange. Although there is some evidence of the tropical tape recorder effect in the model, latitudinal transport in the lower stratosphere is much too strong; the imprint of the seasonally varying tropopause saturation mixing ratio decays with height much more rapidly than is observed. Furthermore, although the tropical tropopause temperatures in the model are colder than those observed, the lower stratosphere in the model is actually moister than that observed. The model does not resolve convective-scale processes, and the parameterized representation of moist convection in the model is apparently unable to mimic the dehydrating effects of overshooting cumulonimbus turrets. On the other hand, the model is rather successful in simulating isentropic exchange with the extratropical lowermost stratosphere caused by wave breaking associated with synoptic-scale weather disturbances. In other words, the model has some representation of the lowermost wavy arrows in Figure 3, including some of the transport effects normally associated with tropopause folding, even though the folds themselves are not resolved realistically.

Two-dimensional chemistry and transport models have an important role to play, because it is not feasible, even with current computing power, to run extended multiyear simulations in three-dimensional GCMs with a full representation of chemistry. Such extended simulations are necessary to calculate, for example, the impact of increased fluorocarbons and aircraft exhaust pollutants on stratospheric ozone, including its role in the radiative balance of the troposphere. It is standard in the two-dimensional models to assume that transport may be approximated by an averaged circulation and a spatially varying, nonisotropic diffusion and also to make simplifying assumptions concerning the chemistry, for example, neglect or approximation of diurnal variation. Together with the small number of degrees of freedom, such approximations greatly reduce the computational requirements. Such models have been used extensively for major stratospheric ozone assessments [e.g., *WMO*, 1986, 1995], and comparison with observation shows

that the models can be useful for long-term simulations. However, the two-dimensional models have obvious limitations, and in particular, it seems likely that for species that are controlled by transport as well as by photochemistry, simulated temporal variability can be regarded with any confidence only for timescales of a season or longer.

Two-dimensional models are clearly not capable of representing the details of STE. The horizontal, vertical, and temporal resolutions are insufficient. However, since they are based on the global-scale circulation, they might be expected to represent the bulk magnitude of STE with some accuracy when the model exchange is integrated over a year. Also, since the two-dimensional model circulation is usually derived from observations in the overworld, the annual mass exchange between the stratosphere and troposphere calculated from a two-dimensional model [e.g., *Douglass et al.*, 1992] should be similar to that calculated in other ways [*Rosenlof and Holton*, 1993; *Follows*, 1992; *Robinson*, 1980].

The two-dimensional model transport from the troposphere to the stratosphere takes place mainly in the tropics. Transport from the stratosphere to the troposphere is due to advective transport at middle to high latitudes, horizontal transport where the two-dimensional model tropopause height is changed by a grid box or more, and vertical diffusion to the troposphere at all latitudes outside the tropics [e.g., *Douglass et al.*, 1993; *Shia et al.*, 1993]. Though observations indicate that most transport from the stratosphere to the troposphere takes place at middle latitudes and is associated with synoptic-scale storms (see section 6), in many two-dimensional models, stratosphere-to-troposphere transport tends to be placed poleward of the storm tracks at middle to high latitudes. Three-dimensional model simulations, on the other hand, are able to simulate the observed distribution of stratosphere-to-troposphere transport [e.g., *Douglass et al.*, 1993; *Rasch et al.*, 1994]. This difference in placement of the downward branch of STE is the most important difference between most two-dimensional and three-dimensional tracer simulations when comparisons are made for seasonal or longer timescales.

Introduction of strong eddy diffusion in two-dimensional models to represent quasi-isentropic transport across the tropopause may have unrealistic consequences. For instance, a pollutant that is introduced into the northern hemisphere lowermost stratosphere may rapidly enter the southern hemisphere lowermost stratosphere via rapid diffusion along isentropes, crossing the equator on the tropospheric portion of such isentropes. Such a case is discussed by *Shia et al.* [1993]. Since, however, these processes are basically specified in the two-dimensional formulation, the specification can be changed to make the STE morphologically correct. Specification of such parameters compromises the robustness of the models' sensitivity to change.

For applications in which the sources of stratospheric constituents have strong latitudinal, longitudinal, and temporal variability, the inadequacies of the current two-dimensional formulations will have serious consequences. However, two-dimensional models are thought to be an appropriate tool for calculating the impact of chlorofluorocarbon increase on stratospheric ozone. The chlorofluorocarbons $CFCl_3$ and CF_2Cl_2 are both photolyzed mainly above 25 km; these molecules are released in industrialized areas but are well mixed in the troposphere and thus represent a zonally symmetric source of inorganic chlorine in the upper stratosphere. The fraction of chlorine that participates directly in ozone destruction (Cl and ClO) and the fraction stored in reservoirs (HCl and $ClONO_2$) are determined photochemically. The only loss process for stratospheric chlorine is transport to the troposphere, where soluble species are removed by rainout. Although transport to the troposphere from the stratosphere is not necessarily zonally symmetric, on an annual basis the net loss from the stratosphere is approximated well by the diabatic circulation.

The hydrochlorofluorocarbons (HCFCs) that are increasingly being used as substitutes for $CFCl_3$ and CF_2Cl_2 are shorter-lived and have significant photochemical loss rates (and thus production rates of reactive chlorine species) in the lower stratosphere. Owing to the difficulties of accurately representing transport and mixing in the lower stratosphere in two-dimensional models, assessment of the relative impact of these species on ozone using such models is more problematic.

Efforts to evaluate the atmospheric effects of subsonic and supersonic aircraft are motivating more research into tracer transport and STE [e.g., *Sparling et al.*, 1995]. By contrast with the inorganic chlorine source, which the stratosphere sees as zonally symmetric, supersonic aircraft emissions will be restricted to oceanic corridors and are thus a highly nonzonal source of pollutants. Other aircraft emissions are also nonzonal, reflecting the geography of the main air traffic corridors. Pollutant transport in the lowermost stratosphere is very much a three-dimensional process, involving rapid advection and shearing along jet streams as well as eddy dispersion away from them in much the way illustrated in Figure 2b. All of these processes are generally much stronger in winter than in summer. These three-dimensional aspects are not dealt with adequately by parameterizations in current two-dimensional models and may be beyond such models altogether, even with improved parameterization.

Many of the issues raised by the aircraft emissions problem require evaluation of three-dimensional transport effects. This does not, however, mean that it is

necessary to use a full general circulation model. An alternative approach to three-dimensional modeling of constituent transport and exchange is to solve the continuity equations for the relevant trace constituents using wind data that may come either from a previous run of a general circulation model [*Rasch et al.*, 1994; *Chipperfield et al.*, 1994] or from real data analyzed with a data assimilation system, such as the STRATAN stratospheric analysis system [*Coy et al.*, 1994] used by *Douglass et al.* [1993]. For either of these approaches, the net mass exchange calculated diagnostically (e.g., by downward control and its generalizations) will be similar, provided that the general circulation model produces a reasonable climatology of winds and temperature. For both approaches, the meridional circulation (and the annual mass exchange between the troposphere and stratosphere) will be similar to that calculated using a two-dimensional model.

Global three-dimensional transport models address some of the shortcomings of the two-dimensional models, in that the upper tropospheric synoptic scales are explicitly represented in the model. For transport associated with synoptic scales, the storm tracks and the upper tropospheric synoptic-scale activity are defined realistically when assimilated fields are used for the transport. The results of simulations may be verified with constituent observations; however, such comparisons do not provide a strong test of the accuracy of the net stratosphere-to-troposphere mass flux calculation. Spatial and temporal resolution of such models is still too coarse to resolve the actual STE processes. Small differences in horizontal behavior can lead to large differences in vertical transport; comparisons of calculated and observed profile shapes are not a strong test because irreversible transport is not clearly identified. The models may adequately represent the effects of the STE processes; judgment on this subject requires careful analysis of three-dimensional model behavior, both for long (annual) timescales and shorter timescales associated with specific events.

Consideration of specific events in global transport models illustrates the heterogeneity of their behavior. The variability of cutoff cyclone systems is illustrated by two examples which are found in the STRATAN 500-hPa geopotential height fields on January 21 and 22, 1989 (Plates 3a and 3b). Relatively low values of geopotential height are found at 35°E and 315°E on January 21; by January 22, both of these areas are identified by closed contours. An ozone simulation, initialized on December 28, 1988, and total ozone mapping spectrometer (TOMS) observations for January 22, 1989, are given in Plates 4a and 4b. The areas of the cutoff cyclones are characterized by relative maxima in the TOMS data and in the simulation. In both the data and the simulation, the maxima at 315°E are larger than at 35°E. However, vertical cross sections of ozone (Plates 5a and 5b) show that the downward displacement of mixing ratio surfaces associated with the weaker cutoff cyclone (and the less intense ozone maximum) at 35°E is greater than that associated with the stronger cutoff cyclone at 315°E. This suggests that the exchange is associated with more derived (or sophisticated) parameters that define the characteristics of the cutoff cyclones.

Price and Vaughan [1993] have suggested that only a relatively small percentage of the STE is associated with cutoff cyclones. Instead, they argue that strong baroclinic folding events [e.g., *Danielsen*, 1968] are more important. The basic features of the transport are entirely consistent with the mechanisms proposed by *Danielsen* [1968]. Irreversibility is closely correlated with diabatic heating terms, both radiative and moist-convective, consistent with the mesoscale studies of *Lamarque and Hess* [1994]. Once again, however, zeroth-order parameters such as surface pressure do not provide good predictors of the STE associated with a particular event; quantities such as static stability, baroclinicity, and transience made irreversible by Rossby-wave breaking are important in determining the amount of stratospheric air which is carried to altitudes low enough to lead to effective STE. There is some modeling and observational experience to suggest that rapidly growing events [see *Ucellini et al.*, 1985] might play a particularly important role in midlatitude STE.

Diagnosis of the representation of tropical STE in three-dimensional models has received relatively little attention. As we have argued above, the overall rate of upward mass transport across the tropical tropopause is determined nonlocally as part of the overall general circulation and is not likely to be sensitive to the details of tropical moist convection. Tropical STE of water vapor and of short-lived tropospheric tracers is, however, likely to be strongly influenced by the manner in which tropical moist convection is parameterized in the model as discussed above for the CCM2 simulation of *Mote et al.* [1994].

9. SUMMARY AND CONCLUSIONS

In this review we have argued that to understand STE, it is essential to include it within a global-scale dynamical perspective and that it is useful then to distinguish between the part of the atmosphere whose isentropes lie entirely above the tropopause (the overworld, $\theta \gtrsim 380$ K) and the part of the atmosphere whose isentropes span the troposphere and the stratosphere (whose stratospheric part we have called the lowermost stratosphere). Further, we have argued that chemical transport between the lowermost stratosphere and the overworld is controlled largely by the same nonlocal dynamical processes that are responsible for the global-scale mean circulation in the middle atmosphere, with upward mass flux in the tropics and

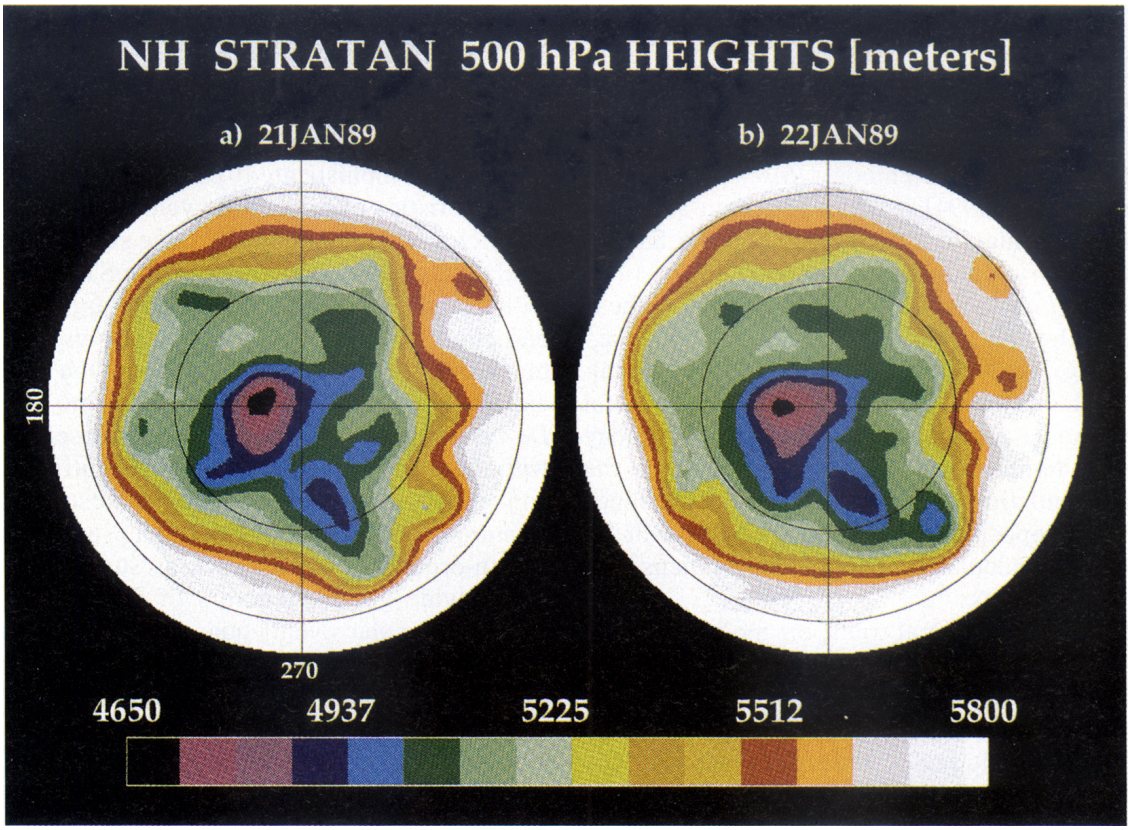

Plate 3. The 500-hPa geopotential height fields from the STRATAN data assimilation system for (a) January 21, 1989, and (b) January 22, 1989. Cutoff cyclones are found near 35°E and 315°E in Plate 3b.

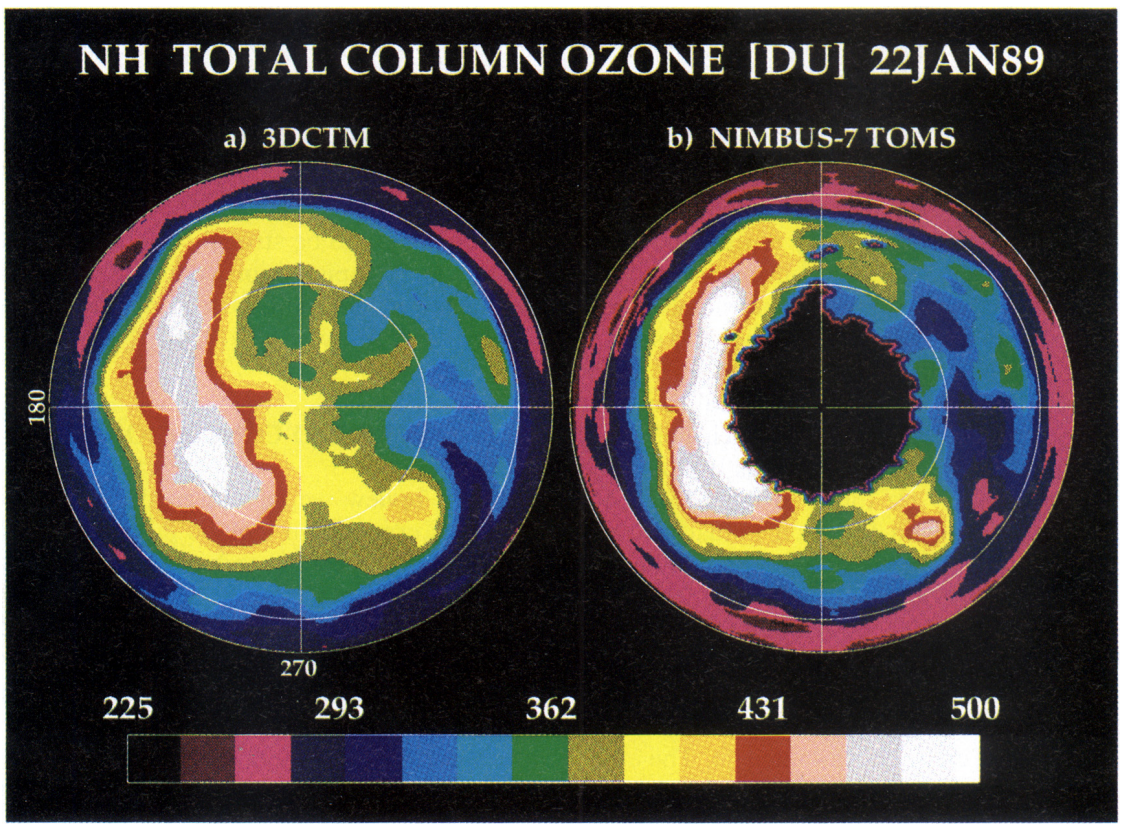

Plate 4. (a) Ozone simulation in a three-dimensional chemical transport model initialized on December 28, 1988, and (b) TOMS observations for January 22, 1989. Areas corresponding to the cutoff cyclones of Plate 3b show anomalously high total ozone in both data and simulation.

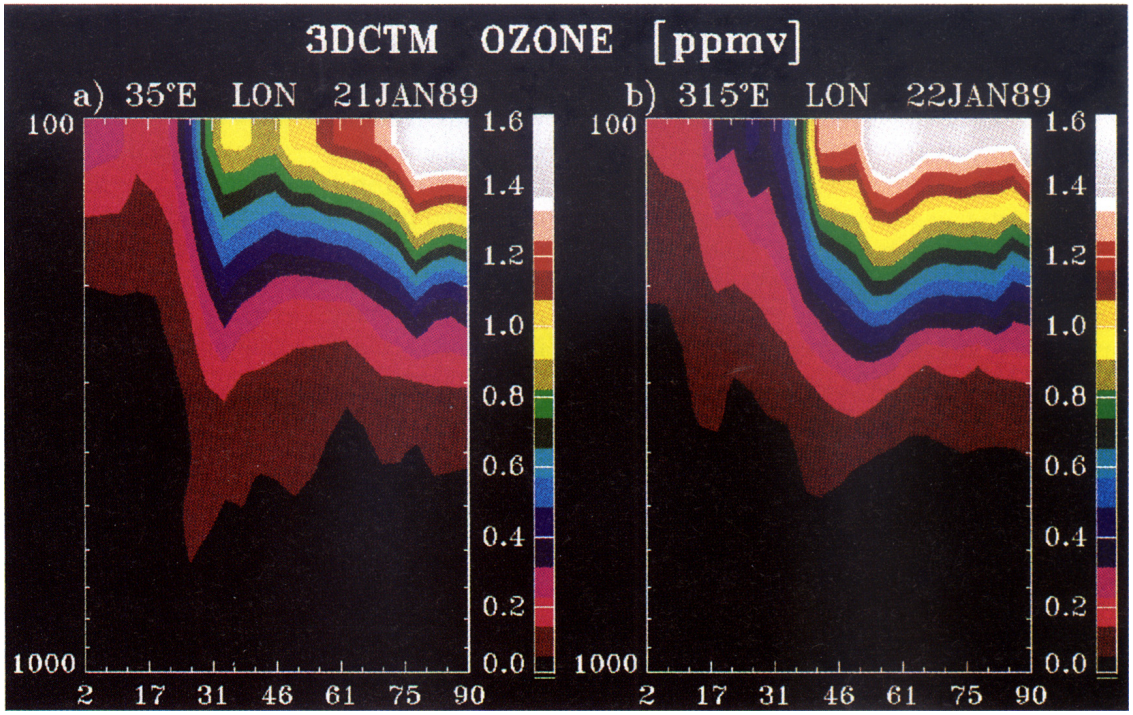

Plate 5. Vertical cross sections of ozone for the two cutoff cyclones near (a) 35°E and (b) 315°E, in Plate 3b.

downward mass flux in the extratropics. These upward and downward mass fluxes largely govern global-scale chemical transport, because of the weakness of eddy transport along isentropes between the tropics and the extratropics, that is, because of the effectiveness, to a sufficient extent for this purpose, of the subtropical eddy transport barriers in the overworld explained in section 4 (and now directly evidenced by the "tape recorder effect" [*Mote et al.*, 1995a, b].

These considerations explain the remarkable fact that STE, viewed from this global perspective and on seasonal and longer timescales, is controlled not by local events near the tropopause but by the generation of large-scale and small-scale waves mainly in the troposphere, followed by their propagation into the stratosphere and mesosphere and their dissipation to produce a predominantly and persistently westward wave-induced zonal force. In the extratropics, which are strongly influenced by the Earth's rotation, such a persistent westward force tends to come into balance with the Coriolis force associated with a mean poleward mass flux. In other words, air persistently pushed westward has a persistent tendency to move poleward, a kind of quasi-gyroscopic effect, producing the pumping action discussed in section 3. By mass continuity and through the tendency toward "downward control" exhibited by the extratropical response to the zonal force (the thought experiment of Figure 4b being especially relevant), most of the air pushed poleward must go downward. Mass continuity now requires a corresponding upward mass flux in lower latitudes, and as Figure 4b suggests, with a realistic force varying on the seasonal timescale, much of this mass flux is drawn up from the deep tropics. Therefore it is reasonable to regard the wave-induced force as exerting a "downward-sideways control" on the transport across the lower boundary of the overworld, even though the troposphere in turn appears to be the ultimate source of the waves, and hence the wave-induced force, in the stratosphere and mesosphere.

A complete picture would need to include the pumping effects of gravity-wave drag in the upper stratosphere and in the mesosphere, at present not well quantified. Together with the transient response to the seasonal cycle of solar heating, these probably affect the details of the tropical upwelling in the upper stratosphere. But, more importantly for the purposes of this review, none of these effects change the gross picture just summarized, particularly the fact that upwelling mass fluxes in the lower tropical stratosphere are extratropically controlled.

It can hardly be overemphasized that this downward-sideways control view of stratosphere-troposphere exchange is a drastic departure from previous conceptual models for exchange. In the extensive review of the literature on stratosphere-troposphere exchange given by *WMO* [1986], the focus is almost entirely on synoptic-scale and small-scale processes occurring near the tropopause. Although no attempt was made in that review to provide estimates of global mass exchange rates, it is clear from the context that the authors felt that it would be necessary to compile statistics on the climatology of exchange by individual tropopause-folding events and by cutoff cyclones, cor-

responding mostly to the lowermost wavy arrow in Figure 3, in order to derive seasonal estimates of exchange for each hemisphere.

There is also a sense conveyed in the discussion of tropical exchange by *WMO* [1986] that air is forced into the stratosphere by tropical cumulus convective turrets and that estimates of tropical mass exchange would thus require accurate measurement of the cumulus mass flux near the tropopause. We have argued, however, as was just summarized, that the net irreversible mass flux into the tropical stratosphere is primarily determined nonlocally by dynamical forcing in the form of wave-driven pumping from the extratropical stratosphere, which is the immediate cause of the slow diabatic ascent in the tropical stratosphere. Because transport across the isentropic surfaces near 380 K defining the lower boundary of the overworld is independent of the details of tropopause folding, of formation of cutoff cyclones, and of the distribution and intensity of tropical cumulus convection, it is possible, as is discussed in section 5, to obtain seasonal estimates of the total upward transport into the tropical stratosphere and the total downward transport into the lowermost stratosphere in each hemisphere by estimating the extratropical wave-induced forcing from general circulation statistics using the method of downward control. This has substantial advantages over the strategy of adding up the exchange from individual events. The latter is likely to be very difficult to carry out, given the indications from modeling studies (see section 8) that there is large variability in the STE associated with nominally similar synoptic events. Furthermore, there is no reason why all the exchange across the mean tropopause should take place in a conspicuous way; there could be significant global-scale subsidence or, for that matter, ascent, and significant contributions from the smaller-scale eddies illustrated in Figure 2b.

The recognition of global-scale nonlocal control, mainly in the downward-sideways sense explained above, also provides an explanation for the observed springtime maxima in the fluxes of ozone and radioactive tracers into the northern hemisphere troposphere. It has usually been assumed [e.g., *WMO*, 1986] that these seasonal maxima are caused by a springtime peak in upper level cyclogenesis and tropopause-folding events. *Mahlman* [1969], however, showed that although surface radioactivity and upper level cyclogenesis were strongly correlated on the timescale of individual cyclones, the seasonal cycle in observed surface radioactive tracer debris concentration in the northern hemisphere was not significantly correlated with the seasonal cycle in an index of upper tropospheric cyclone activity.

The findings of *Mahlman* [1969] are not consistent with the view that the seasonal cycle in the rate of mass transfer into the troposphere is controlled by the frequency and amplitude of tropopause folding events or cutoff cyclone formation. They are, however, consistent with the concept of nonlocal control. According to that concept, the occurrence of a maximum in wave dissipation in the northern hemisphere winter stratosphere leads to a maximum downward mass flux in the extratropics and hence to a maximum flux of ozone and other stratospheric tracers across the lower boundary of the overworld into the lowermost stratosphere during the northern winter. This winter season peak in the fluxes into the lowermost stratosphere inevitably leads to a buildup of tracers in the lower stratosphere until the tracer mixing ratios become sufficiently large that the fluxes from the overworld can be matched by increased fluxes into the troposphere. Thus a springtime maximum in flux of ozone and radioactive tracers into the troposphere would exist even if the rate of irreversible mass transport by tropopause folding and other small-scale processes remained constant all year round.

Notwithstanding the attractive simplicity of the nonlocal-control concept and its implications for the global circulation and chemical transport in the overworld, there are important temporal and spatial variations in the occurrences of tropopause folding and in the formation of cutoff cyclones. Problems such as aircraft emission impact assessments, removal of volcanic debris, and seasonal ozone depletion will almost certainly require the consideration of detailed mechanisms. Detailed modeling of the geographical and temporal dependence of transport across the tropopause requires careful consideration of the various processes that operate on the synoptic scale and on the mesoscale. Similarly, an understanding of the stratospheric water vapor budget and of the longitudinal variations in transfer of mass and trace constituents across the tropical tropopause requires careful consideration of the physics and dynamics of tropical convective systems and in particular of the roles of overshooting convective towers and equatorial gravity waves in the possibly hyperventilated layer just above the tropical tropopause. These problems, important as they are, should be clearly distinguished from the problem of global mass exchange, which as we have argued is determined as part of the general circulation of the global stratosphere.

GLOSSARY

Baroclinic waves and eddies: synoptic-scale disturbances (horizontal length scales usually of the order of 1000 km) that develop partly through a process known as baroclinic instability, in which potential energy stored in the background flow and associated with the horizontal temperature gradient is released. Baroclinic waves and eddies are most prevalent in the storm track regions coincident with tropospheric jet

streams and strong low-level temperature gradients associated for instance with land-sea contrasts. Such waves and eddies (e.g., Figure 2 and Plate 1) lead to deformation and sharpening of the tropopause and to irreversible exchange.

Contour advection: an adaptive computational technique for accurately following a quasi-material contour, for example, the isoline of a tracer, on an isentropic surface by computing the isentropic trajectories of a large number of parcels placed initially at short intervals along the contour. Resolution is maintained even in "surf-zone" conditions by systematically redistributing and adding parcels as necessary.

Diabatic circulation: the vertical circulation associated with cross-isentropic flow. In the global-scale stratospheric "overworld" (see Figure 3), the most persistent and chemically important contribution to the diabatic circulation is that caused by the extratropical pumping discussed in section 3. The resulting vertical motion, upward in the tropics and downward in the extratropics, pushes temperatures away from the radiatively determined values toward which they would otherwise relax, thus inducing the required diabatic heating or cooling.

Downward control: the extratropical wave-driven pumping discussed in section 3 in the long-time limiting case of Figure 4c. In Figure 4c, a steady state has been reached in which the pumping action of a steady force extends purely downward. That is, the nonlocal control of the pumped circulation is felt, after a long enough time, only directly below the extratropical locations where the force is nonzero. For a more realistic, seasonally varying global-scale force, the limiting configuration does not apply in the tropics. This is illustrated by the thought experiment of Figure 4b. The control is then mainly downward-sideways; that is, the extratropical pumping by such a seasonally varying force is felt mainly (as regards mass flux) below and to the side of the force, most powerfully to the tropical side. In the long-time mean the downward-sideways effect on the tropics depends on nonlinear aspects of the response to the pumping.

Elliptic partial differential equations: these often arise in models of physical systems with nonlocal effects approximated as acting instantaneously. Standard textbook examples include the Laplace and Poisson equations arising in electrostatics and in static models of elastic membranes. Equation (5) is also elliptic when $\sigma/\alpha \neq 0$. Changing conditions at one location generally affect the solution everywhere else, as when poking an elastic membrane.

Eulerian (zonal) mean: an instantaneous average taken around a latitude circle. In the conventional Eulerian mean the average is taken at constant altitude or constant pressure; the resultant mean meridional circulation does not provide a useful approximation for the advective transport of tracers. In isentropic coordinates the average is taken at constant potential temperature; the resultant diabatic circulation does provide a good approximation for the vertical advection of tracers. In the transformed Eulerian mean (TEM) the average is taken at constant pressure, but the resulting mean meridional circulation is a good approximation to the zonal-mean diabatic circulation for seasonal and longer timescales.

Frontogenesis: a meteorological term applied to certain cases of strain or deformation in the wind field sharpening the horizontal gradient of an advected quantity. Such gradient sharpening is commonplace and often produces streaky or filamentary features like those in Plate 1 or those produced by stirring cream on coffee. The classical cases called "frontogenesis" refer to the sharpening of large-scale horizontal gradients and are of two main kinds: (1) those in which the advected quantity is potential temperature near the Earth's surface, in extreme cases producing the sudden temperature drops that can accompany changes in surface weather, and (2) those in which the advected quantity is PV on isentropes near the tropopause, in extreme cases producing the highly distorted streaky structures known as "upper air fronts" or "tropopause folds" (see below), a section through one of which is shown in Figure 7.

Geostrophic flow: a hypothetical flow in which the horizontal component of the Coriolis force due to the rotation of the Earth is exactly balanced by the horizontal pressure gradient force. Geostrophic flow is a good approximation to the actual flow (with errors of the order of, say, 10% or better) for most large-scale extratropical wind fields, the main exceptions being strong cutoff cyclones and strongly curved jet streams, in which the relative centrifugal acceleration can be fairly important.

Log-pressure coordinate (z): a vertical coordinate, approximately equal to geopotential height, which is commonly used as an independent variable in dynamical meteorology. It is defined as $z = -H \ln (p/p_s)$, where H is a mean density scale height, p is pressure, and $p_s = 1000$ hPa.

Potential temperature: defined for a perfect gas as $\theta = T(p_0/p)^{R/c_p}$, where T is the temperature, p is pressure, $p_0 = 1000$ hPa, and R and c_p are the gas constant and specific heat for dry air, respectively. Thus θ is the temperature that an air parcel would attain if it were adiabatically compressed to pressure p_0. Potential temperature is a function of specific entropy alone and is thus conserved by fluid parcels when the motion is adiabatic. Since diabatic processes operate on the timescale of weeks in the stratosphere, on shorter timescales parcels move approximately along constant θ surfaces. Isentropic coordinates use potential temperature or specific entropy as the vertical coordinate and have the advantage that under adiabatic conditions there appears to be no vertical motion. Thus adiabatic motion appears to be horizontally

two dimensional (albeit compressible) when viewed in isentropic coordinates.

Potential vorticity (PV): somewhat analogous to spin angular momentum. For large-scale atmospheric motion, PV is conventionally defined as $P = (\zeta_\theta + f)(-g\partial\theta/\partial p)$, where ζ_θ is Rossby's "isentropic relative vorticity," a vorticity-like quantity approximately equal to the component of relative vorticity normal to an isentropic surface; f is the Coriolis parameter (twice the local vertical component of the Earth's angular velocity); and g is the gravitational acceleration. Here θ can be replaced by any monotonic function $F(\theta)$, but the conventional choice, $F(\theta) = \theta$, is empirically useful for defining the extratropical tropopause. Since $-\partial p/\partial\theta$ may be regarded as a local measure of the depth (in pressure units) of the layer between two potential temperature surfaces, an increase in $-\partial p/\partial\theta$ implies stretching of vortex tubes and an increase in the absolute vorticity, and vice versa, somewhat like the spin angular momentum of a ballerina or figure skater.

PV substance: the concept underlying the analogy between PV and chemical mass mixing ratios. PV substance means an imaginary quasi-chemical substance whose mixing ratio, or amount per unit mass of air, is equal, by definition, to the PV. The term "PV substance" does not mean the amount of anything per unit volume. *Haynes and McIntyre* [1990] show that the PV substance can be imagined to consist of massless, positively or negatively charged particles that, despite appearances in some cases, are unable to cross isentropic surfaces. So for inviscid, adiabatic motion the PV substance is simply advected like an inert chemical tracer, but it cannot, for instance, be mixed like a chemical.

Quasi-geostrophic approximation: in the zonally (longitudinally) averaged equations, an approximation assuming that the zonal wind component is in geostrophic balance with the latitudinal pressure gradient and neglecting the effect of the mean meridional circulation in the zonal momentum equation, except in the Coriolis term, and in the thermodynamic energy equation, except in the static stability term.

Rossby wave: a wave that owes its existence to the spatial variation of potential vorticity along isentropic surfaces, for example, in the latitudinal direction. Because of the variation of potential vorticity, a fluid parcel displaced along the potential vorticity gradient (and materially conserving its potential vorticity and potential temperature) will have potential vorticity different from that in the local environment and will induce a perturbation velocity disturbance. This will cause parcel displacements of the same sign to the west of the original displaced parcel and of the opposite sign to the east. The result is a tendency of the PV wave pattern in the potential vorticity field to undulate quasi-elastically, the undulations propagating westward relative to the mean flow.

Rossby-wave breaking: a concept basic to the theory of wave-induced forces (e.g., *McIntyre* [1993] and many other references), as well as to the formation of the "surf zones" conspicuous in stratospheric PV and tracer distributions. "Breaking" does not mean fracturing in any literal sense; rather, it refers to situations in which PV contours deform irreversibly, or fold sideways, on isentropic surfaces, instead of merely undulating as in a simple Rossby wave. These irreversible deformations may take place in a complicated, layerwise two-dimensional turbulent manner, as illustrated in Plate 1.

Tropopause folding: a process in which a thin band of stratospheric air intrudes more or less deeply into the troposphere along the strongly tilted isentropes associated with an upper tropospheric frontal zone. See Figure 7 for a vertical cross section through a large midlatitude fold. It is usually found that part of the stratospheric air in the fold returns reversibly to the stratosphere, and part is drawn irreversibly into the troposphere by advection round an anticyclone; this is a special case of "Rossby-wave breaking" (see above). The distinction between reversible and irreversible behavior is important for a correct estimation of the resulting eddy fluxes of constituents from the lowermost stratosphere to the troposphere.

ACKNOWLEDGMENTS. We wish to thank Alan Plumb and Jerry Mahlman for their perceptive reviews, Julia Slingo for kindly providing unpublished insight and information on tropical convection, Warwick Norton for furnishing Plate 1, and Christof Appenzeller for preparing Figure 1 and for his comments on the manuscript. This paper is based partly on discussions at the NATO Advanced Research Workshop, Stratosphere-Troposphere Exchange, held in Cambridge, England, in September 1993. We wish to thank the Scientific and Environmental Affairs Division of NATO for its support. Partial support was provided by the NASA Atmospheric Chemistry Modeling and Analysis Office under NASA grants NAGW-3794 and NAGW-662 and by the U. K. Natural Environment Research Council through the Universities Global Atmospheric Modelling Programme.

REFERENCES

Ackerman, T. P., K.-N. Liou, F. P. J. Valero, and L. Pfister, Heating rates in tropical anvils, *J. Atmos. Sci.*, *45*, 1606–1623, 1988.

Allam, R. J., and A. F. Tuck, Transport of water vapour in a stratosphere-troposphere general circulation model, I, Fluxes, *Q. J. R. Meteorol. Soc.*, *110*, 321–356, 1984a.

Allam, R. J., and A. F. Tuck, Transport of water vapour in a stratosphere-troposphere general circulation model, II, Trajectories, *Q. J. R. Meteorol. Soc.*, *110*, 357–392, 1984b.

Andrews, D. G., J. R. Holton, and C. B. Leovy, *Middle Atmospheric Dynamics*, 489 pp., Academic, San Diego, Calif., 1987.

Appenzeller, C., and H. C. Davies, Structure of strato-

spheric intrusions into the troposphere, *Nature*, *358*, 570–572, 1992.

Appenzeller, C., H. C. Davies, and W. A. Norton, Fragmentation of stratospheric intrusions, *J. Geophys. Res.*, in press, 1995.

Bamber, D. J., P. G. W. Healey, B. M. R. Jones, S. A. Penkett, A. F. Tuck, and G. Vaughan, Vertical profiles of tropospheric gases: Chemical consequences of stratospheric intrusions, *Atmos. Environ.*, *18*, 1759–1766, 1984.

Boering, K. A., et al., Measurements of stratospheric carbon dioxide and water vapor at northern midlatitudes: Implications for troposphere-to-stratosphere transport, *Geophys. Res. Lett.*, in press, 1995.

Brewer, A. M., Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, *Q. J. R. Meteorol. Soc.*, *75*, 351–363, 1949.

Browell, E. V., E. F. Danielsen, S. Ismail, G. L. Gregory, and S. M. Bleck, Tropopause fold structure determined from airborne lidar and in situ measurements, *J. Geophys. Res.*, *92*, 2112–2120, 1987.

Bush, A. B. G., and W. R. Peltier, Tropopause folds and synoptic-scale baroclinic wave life cycles, *J. Atmos. Sci.*, *51*, 1581–1604, 1994.

Carr, E. S., et al., Tropical stratospheric water vapor measured by the microwave limb sounder (MLS), *Geophys. Res. Lett.*, *22*, 691–694, 1995.

Chen, P., J. R. Holton, A. O'Neill, and R. Swinbank, Isentropic mass exchange between the tropics and extratropics in the stratosphere, *J. Atmos. Sci.*, *51*, 3006–3018, 1994.

Chipperfield, M. P., D. Cariolle, and P. Simon, A three-dimensional transport model study of polar stratospheric cloud processing during EASOE, *Geophys. Res. Lett.*, *21*, 1463–1466, 1994.

Cox, B. D., M. Bithell, and L. J. Gray, A general circulation model study of a tropopause folding event at middle latitudes, *Q. J. R. Meteorol. Soc.*, *121*, 883–910, 1995.

Coy, L., R. B. Rood, and P. A. Newman, A comparison of winds from the STRATAN data assimilation system to geostrophic and balanced wind estimates. *J. Atmos. Sci.*, *51*, 2309–2315, 1994.

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

Danielsen, E. F., A dehydration mechanism for the stratosphere, *Geophys. Res. Lett.*, *9*, 605–608, 1982.

Danielsen, E. F., In situ evidence of rapid, vertical, irreversible transport of lower tropospheric air into the lower tropical stratosphere by convective cloud turrets and by larger-scale upwelling in tropical cyclones, *J. Geophys. Res.*, *98*, 8665–8681, 1993.

Dewan, E. M., Turbulent vertical transport due to thin intermittent mixing layers in the stratosphere and other stable fluids, *Science*, *211*, 1041–1042, 1981.

Dickinson, R. E., On the excitation and propagation of zonal winds in an atmosphere with Newtonian cooling, *J. Atmos. Sci.*, *25*, 269–279, 1968.

Dickinson, R. E., Theory of planetary wave–zonal flow interaction, *J. Atmos. Sci.*, *26*, 73–81, 1969.

Dickinson, R. E., Analytic model for zonal winds in the tropics, I, Details of the model and simulation of gross features of the zonal mean troposphere, *Mon. Weather Rev.*, *99*, 501–510, 1971a.

Dickinson, R. E., Analytic model for zonal winds in the tropics, II, Variation of the tropospheric mean structure with season and differences between hemispheres, *Mon. Weather Rev.*, *99*, 511–523, 1971b.

Douglass, A. R., C. H. Jackman, R. B. Rood, A. C. Aikin, R. S. Stolarski, M. P. McCormick, and D. W. Fahey, Natural cycles, gases, in *The Atmospheric Effects of Stratospheric Aircraft: A First Program Report, NASA Ref. Publ.*, *1272*, 33–61, 1992.

Douglass, A. R., R. B. Rood, C. J. Weaver, and M. C. Cerniglia, Implications of three-dimensional tracer studies for two-dimensional assessments of the impact of supersonic aircraft on stratospheric ozone, *J. Geophys. Res.*, *98*, 8949–8963, 1993.

Dunkerton, T. J., Nonlinear Hadley circulation driven by asymmetric differential heating, *J. Atmos. Sci.*, *46*, 956–974, 1989.

Dunkerton, T. J., Nonlinear propagation of zonal winds in an atmosphere with Newtonian cooling and equatorial wavedriving, *J. Atmos. Sci.*, *48*, 236–263, 1991.

Dunkerton, T. J., C.-P. F. Hsu, and M. E. McIntyre, Some Eulerian and Lagrangian diagnostics for a model stratospheric warming, *J. Atmos. Sci.*, *38*, 819–843, 1981.

Ebel, A., H. Hass, H. J. Jakobs, M. Laure, M. Memmesheimer, A. Oberreuter, H. Geiss, and Y.-H. Kuo, Simulation of ozone intrusion caused by a tropopause fold and cut-off low, *Atmos. Environ., Part A*, *25*, 2131–2144, 1991.

Eliassen, A., Slow thermally or frictionally controlled meridional circulation in a circular vortex, *Astrophys. Norv.*, *5*(2), 19–60, 1951.

Emanuel, K. A., J. D. Neelin, and C. S. Bretherton, On large-scale circulations in convecting atmospheres, *Q. J. R. Meteorol. Soc.*, *120*, 1111–1143, 1994.

Fels, S. B., Radiative-dynamical interactions in the middle atmosphere, *Adv. Geophys.*, *28A*, 277–300, 1985.

Follows, M. J., On the cross-tropopause exchange of air, *J. Atmos. Sci.*, *49*, 879–882, 1992.

Garcia, R. R., On the mean meridional circulation of the middle atmosphere, *J. Atmos. Sci.*, *44*, 3599–3609, 1987.

Garcia, R. R., and B. A. Boville, "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere, *J. Atmos. Sci.*, *51*, 2238–2245, 1994.

Goldan, P. D., W. C. Kuster, D. L. Albritton, and A. L. Schmeltekopf, Stratospheric $CFCl_3$, CF_2Cl_2, and N_2O height profile measurements at several latitudes, *J. Geophys. Res.*, *85*, 413–423, 1980.

Haynes, P. H., and M. E. McIntyre, On the conservation and impermeability theorems for potential vorticity, *J. Atmos. Sci.*, *47*, 2021–2031, 1990.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces, *J. Atmos. Sci.*, *48*, 651–678, 1991.

Held, I. M., On the height of the tropopause and the static stability of the troposphere, *J. Atmos. Sci.*, *39*, 412–417, 1982.

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

Hoerling, M. P., T. K. Schaack, and A. J. Lenzen, A global analysis of stratospheric-tropospheric exchange during northern winter, *Mon. Weather Rev.*, *121*, 162–172, 1993.

Holton, J. R., Troposphere-stratosphere exchange of trace constituents: The water vapor puzzle, in *Dynamics of the Middle Atmosphere*, edited by J. R. Holton and T. Matsuno, pp. 369–385, Terra Sci., Tokyo, 1984.

Holton, J. R., Meridional distribution of stratospheric trace constituents, *J. Atmos. Sci.*, *43*, 1238–1242, 1986.

Hoskins, B. J., Atmospheric frontogenesis models: Some solutions, *Q. J. R. Meteorol. Soc.*, *97*, 139–153, 1971.

Hoskins, B. J., Towards a PV-theta view of the general circulation, *Tellus, Ser. A*, *43*, 27–35, 1991.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, On the use and significance of isentropic potential vorticity maps, *Q. J. R. Meteorol. Soc.*, *111*, 877–946, 1985. (Correction, *Q. J. R. Meteorol. Soc.*, *113*, 402–404, 1987.)

Jensen, E. J., and O. B. Toon, Ice nucleation in the upper troposphere: Sensitivity to aerosol number density, temperature, and cooling rate, *Geophys. Res. Lett.*, *21*, 2019–2022, 1994.

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

Kelly, K., M. H. Proffitt, K. R. Chan, M. Loewenstein, J. R. Podolske, S. E. Strahan, J. C. Wilson, and D. Kley, Water vapor and cloud water measurements over Darwin during the STEP 1987 tropical mission, *J. Geophys. Res.*, *98*, 8713–8723, 1993.

Keyser, D., and M. A. Shapiro, A review of the structure and dynamics of upper level frontal zones, *Mon. Weather Rev.*, *118*, 1914–1921, 1986.

Kley, D. A., A. L. Schmeltekopf, K. Kelly, R. H. Winkler, T. L. Thompson, and M. McFarland, Transport of water vapor through the tropical tropopause, *Geophys. Res. Lett.*, *9*, 617–620, 1982.

Knollenberg, R. G., A. J. Dascher, and D. Huffman, Measurements of the aerosol and ice crystal populations in tropical stratospheric cumulonimbus anvils, *Geophys. Res. Lett.*, *9*, 613–616, 1982.

Knollenberg, R. G., K. Kelly, and J. C. Wilson, Measurements of high number densities of ice crystals in the tops of tropical cumulonimbus, *J. Geophys. Res.*, *98*, 8639–8664, 1993.

Kritz, M., S. W. Rosner, E. F. Danielsen, and H. B. Selkirk, Air mass origins and troposphere-to-stratosphere exchange associated with midlatitude cyclogenesis and tropopause folding inferred from ^{7}Be measurements, *J. Geophys. Res.*, *96*, 17,405–17,414, 1991.

Kritz, M., S. W. Rosner, K. K. Kelly, M. Loewenstein, and K. R. Chan, Radon measurements in the lower tropical stratosphere: Evidence for rapid vertical transport and dehydration of tropospheric air, *J. Geophys. Res.*, *98*, 8725–8736, 1993.

Lamarque, J., and P. G. Hess, Cross-tropopause mass exchange and potential vorticity budget in a simulated tropopause folding, *J. Atmos. Sci.*, *51*, 2246–2269, 1994.

Levy, H., II, J. D. Mahlman, and W. J. Moxim, A stratospheric source of reactive nitrogen in the unpolluted troposphere, *Geophys. Res. Lett.*, *7*, 441–444, 1980.

Lindzen, R. S., The Eady problem for a basic state with zero PV gradient but $\beta \approx 0$, *J. Atmos. Sci.*, *51*, 3221–3226, 1994.

Mahlman, J. D., Long-term dependence of surface fallout fluctuations upon tropopause-level cyclogenesis, *Arch. Meteorol. Geophys. Bioklimatol., Ser. A*, *18*, 299–311, 1969.

Mahlman, J. D., H. Levy II, and W. J. Moxim, Three-dimensional tracer structure and behavior as simulated in two ozone precursor experiments, *J. Atmos. Sci.*, *37*, 655–685, 1980.

Mahlman, J. D., D. G. Andrews, D. L. Hartmann, T. Matsuno, and R. G. Murgatroyd, Transport of trace constituents in the stratosphere, in *Dynamics of the Middle Atmosphere*, edited by J. R. Holton and T. Matsuno, pp. 387–416, D. Reidel, Norwell, Mass., 1984.

Mahlman, J. D., H. Levy II, and W. J. Moxim, Three-dimensional simulations of stratospheric N_2O: Predictions for other trace constituents, *J. Geophys. Res.*, *91*, 2687–2707, 1986.

McCormick, M. P., and R. E. Veiga, SAGE II measurements of early Pinatubo aerosols, *Geophys. Res. Lett.*, *19*, 155–158, 1992.

McCormick, M. P., E. W. Chiou, L. R. McMaster, W. P. Chu, J. C. Larsen, D. Rind, and S. Oltmans, Annual variations of water vapor in the stratosphere and upper troposphere observed by the Stratospheric Aerosol and Gas Experiment II, *J. Geophys. Res.*, *98*, 4867–4874, 1993.

McIntyre, M. E., Atmospheric dynamics: Some fundamentals, with observational implications, in *The Use of EOS for Studies of Atmospheric Physics*, edited by J. C. Gille and G. Visconti, pp. 313–386, North-Holland, New York, 1992.

McIntyre, M. E., On the role of wave propagation and wave breaking in atmosphere-ocean dynamics, in *Theoretical and Applied Mechanics 1992*, edited by S. R. Bodner, J. Singer, A. Solan, and Z. Hashin, pp. 281–304, Elsevier Sci., New York, 1993.

McIntyre, M. E., and T. N. Palmer, The "surf zone" in the stratosphere, *J. Atmos. Terr. Phys.*, *46*, 825–849, 1984.

Mote, P. W., J. R. Holton, and B. A. Boville, Characteristics of stratosphere-troposphere exchange in a general circulation model, *J. Geophys. Res.*, *99*, 16,815–16,829, 1994.

Mote, P. W., K. H. Rosenlof, J. R. Holton, R. S. Harwood, and J. W. Waters, Seasonal variations of water vapor in the tropical lower stratosphere, *Geophys. Res. Lett.*, *22*, 1093–1096, 1995a.

Mote, P. W., K. H. Rosenlof, M. E. McIntyre, E. S. Carr, J. C. Gille, J. R. Holton, J. S. Kinnersley, H. C. Pumphrey, J. M. Russell III, and J. W. Waters, An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, *J. Geophys. Res.*, in press, 1995b.

Newell, R. E., and S. Gould-Stewart, A stratospheric fountain?, *J. Atmos. Sci.*, *38*, 2789–2796, 1981.

Norton, W. A., Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours, *J. Atmos. Sci.*, *51*, 654–673, 1994.

Oltmans, S. J., and D. J. Hofmann, Increase in lower-stratospheric water vapour at a mid-latitude northern hemisphere site from 1981 to 1994, *Nature*, *374*, 146–149, 1995.

Pfister, L., K. R. Chan, T. P. Bui, S. Bowen, M. Legg, B. Gary, K. Kelly, M. Proffitt, and W. Starr, Gravity waves generated by a tropical cyclone during the STEP tropical field program: A case study, *J. Geophys. Res.*, *98*, 8611–8638, 1993.

Plumb, R. A., Zonally symmetric Hough modes and meridional circulations in the middle atmosphere, *J. Atmos. Sci.*, *39*, 983–991, 1982.

Plumb, R. A., and J. D. Mahlman, The zonally averaged transport characteristics of the GFDL general circulation/transport model, *J. Atmos. Sci.*, *44*, 298–327, 1987.

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

Potter, B., and J. R. Holton, The role of monsoon convection in the dehydration of the lower tropical stratosphere, *J. Atmos. Sci.*, *52*, 1034–1050, 1995.

Price, J. D., and G. Vaughan, On the potential for stratosphere-troposphere exchange in cut-off low systems, *Q. J. R. Meteorol. Soc.*, *119*, 343–365, 1993.

Ramaswamy, V., M. D. Schwarzkopf, and K. P. Shine, Radiative forcing of climate from halocarbon-induced global stratospheric ozone loss, *Nature*, *355*, 810–812, 1992.

Randel, W. J., Global variations of zonal mean ozone during

stratospheric warming events, *J. Atmos. Sci.*, *50*, 3308–3321, 1993.

Randel, W. J., J. C. Gille, A. E. Roche, J. B. Kumer, J. L. Mergenthaler, J. W. Waters, E. F. Fishbein, and W. A. Lahoz, Stratospheric transport from the tropics to middle latitudes by planetary wave mixing, *Nature*, *365*, 533–535, 1993.

Rasch, P. J., X. Tie, B. A. Boville, and D. L. Williamson, A three-dimensional transport model for the middle atmosphere, *J. Geophys. Res.*, *99*, 999–1017, 1994.

Reed, R. J., and C. L. Vlcek, The annual temperature variation in the lower tropical stratosphere, *J. Atmos. Sci.*, *26*, 163–167, 1969.

Reid, G. C., and K. S. Gage, On the annual variation in height of the tropical tropopause, *J. Atmos. Sci.*, *38*, 1928–1938, 1981.

Reiter, E. R., M. E. Glasser, and J. D. Mahlman, The role of the tropopause in stratospheric-tropospheric exchange process, *Pure Appl. Geophys.*, *75*, 185–218, 1969.

Remsberg, E. E., J. M. Russell III, L. L. Gordley, P. L. Bailey, W. G. Planet, and J. E. Harries, Implications of the stratospheric water vapor distribution as determined from the Nimbus 7 LIMS experiment, *J. Atmos. Sci.*, *41*, 2934–2945, 1984.

Robinson, G. D., The transport of minor atmospheric constituents between troposphere and stratosphere, *Q. J. R. Meteorol. Soc.*, *106*, 227–253, 1980.

Rosenlof, K. H., The seasonal cycle of the residual mean meridional circulation in the stratosphere, *J. Geophys. Res.*, *100*, 5173–5191, 1995.

Rosenlof, K. H., and J. R. Holton, Estimates of the stratospheric residual circulation using the downward control principle, *J. Geophys. Res.*, *98*, 10,465–10,479, 1993.

Rotunno, R., W. C. Skamarock, and C. Snyder, An analysis of frontogenesis in numerical simulations of baroclinic waves, *J. Atmos. Sci.*, *51*, 3373–3398, 1994.

Russell, P. B., L. Pfister, and H. B. Selkirk, The tropical experiment of the stratosphere-troposphere exchange project (STEP): Science objectives, operations, and summary findings, *J. Geophys. Res.*, *98*, 8563–8589, 1993.

Schneider, E. K., Axially symmetric steady-state models of the basic state for instability and climate studies, II, Nonlinear calculations, *J. Atmos. Sci.*, *34*, 280–297, 1977.

Selkirk, H. B., The tropopause cold trap in the Australian monsoon during STEP/AMEX 1987, *J. Geophys. Res.*, *98*, 8591–8610, 1993.

Shapiro, M. A., Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and the troposphere, *J. Atmos. Sci.*, *37*, 994–1004, 1980.

Shia, R.-L., M. K. W. Ko, M. Zou, and V. R. Kotamarthi, Cross-tropopause transport of excess ^{14}C in a two-dimensional model, *J. Geophys. Res.*, *98*, 18,607–18,616, 1993.

Solomon, S., J. T. Kiehl, R. R. Garcia, and W. L. Grose, Tracer transport by the adiabatic circulation deduced from satellite observations, *J. Atmos. Sci.*, *43*, 1603–1617, 1986.

Sparling, L. C., M. R. Schoeberl, A. R. Douglass, C. J. Weaver, P. A. Newman, and L. R. Lait, Trajectory modeling of emissions from lower stratospheric aircraft, *J. Geophys. Res.*, *100*, 1427–1438, 1995.

Strahan, S. E., and J. D. Mahlman, Evaluation of the SKYHI general circulation model using aircraft NO_2 measurements, 1, Polar winter stratospheric meteorology and tracer morphology, *J. Geophys. Res.*, *99*, 10,305–10,318, 1994a.

Strahan, S. E., and J. D. Mahlman, Evaluation of the SKYHI general circulation model using aircraft NO_2 measurements, 2, Tracer variability and diabatic meridional circulation, *J. Geophys. Res.*, *99*, 10,319–10,332, 1994b.

Swinbank, R., and A. O'Neill, A stratosphere-troposphere data assimilation system, *Mon. Weather Rev.*, *122*, 686–702, 1994.

Toon, O. B., R. P. Turco, J. Jordan, J. Goodman, and G. Ferry, Physical processes in polar stratospheric ice clouds, *J. Geophys. Res.*, *94*, 11,359–11,380, 1989.

Toumi, R., S. Bekki, and K. S. Law, Indirect influence of ozone depletion on climate forcing by clouds, *Nature*, *372*, 348–351, 1994.

Trepte, C. R., and M. H. Hitchman, Tropical stratospheric circulation deduced from satellite aerosol data, *Nature*, *355*, 626–628, 1992.

Trepte, C. R., R. E. Veiga, and M. P. McCormick, The poleward dispersal of Mount Pinatubo volcanic aerosol, *J. Geophys. Res.*, *98*, 18,563–18,573, 1993.

Tsuda, T., Y. Murayama, H. Wiryosumarto, S. W. B. Harijono, and S. Kato, Radiosonde observations of equatorial atmosphere dynamics over Indonesia, 1, Equatorial waves and diurnal tides, *J. Geophys. Res.*, *99*, 10,491–10,506, 1994.

Tung, K. K., On the two-dimensional transport of stratospheric trace gases in isentropic coordinates, *J. Atmos. Sci.*, *39*, 2330–2355, 1982.

Ucellini, L. W., D. Keyser, K. F. Brill, and C. H. Wash, The Presidents' Day cyclone of 18–19 February 1979: Influence of upstream trough amplification and associated tropopause folding on rapid cyclogenesis, *Mon. Weather Rev.*, *113*, 962–988, 1985.

Waugh, D. W., and R. A. Plumb, Contour advection with surgery: A technique for investigating fine-scale structure in tracer transport, *J. Atmos. Sci.*, *51*, 530–540, 1994.

Wei, M.-Y., A new formulation of the exchange of mass and trace constituents between the stratosphere and the troposphere, *J. Atmos. Sci.*, *44*, 3079–3086, 1987.

Wirth, V., Diabatic heating in an axisymmetric cut-off cyclone and related stratosphere-troposphere exchange. *Q. J. R. Meteorol. Soc.*, *121*, 127–147, 1995.

World Meteorological Organization (WMO), Atmospheric ozone 1985, *WMO 16*, Geneva, Switzerland, 1986.

WMO, Scientific Assessment of Ozone Depletion: 1994, *WMO 37*, Geneva, Switzerland, 1995.

Yang, H., and R. T. Pierrehumbert, Production of dry air by isentropic mixing, *J. Atmos. Sci.*, *51*, 3437–3454, 1994.

Yulaeva, E., J. R. Holton, and J. M. Wallace, On the cause of the annual cycle in the tropical lower stratospheric temperature, *J. Atmos. Sci.*, *51*, 169–174, 1994.

A. R. Douglass and R. B. Rood, Laboratory for Atmospheres, Code 916, NASA Goddard Space Flight Center, Greenbelt, MD 20771. (e-mail: douglas@aurora.gsfc.nasa.gov; rood@oz.gsfc.nasa.gov)

P. H. Haynes and M. E. McIntyre, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, England. (e-mail: phh@amtp.cam.ac.uk; m.e.mcintyre@damtp.cam.ac.uk)

J. R. Holton, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640. (e-mail: holton@atmos.washington.edu)

L. Pfister, Space Science 245-3, NASA Ames Research Center, Moffett Field, CA 94035. (e-mail: pfister@margie.arc.nasa.gov)