7 February 1998

revised version, *Journal of Geophysical Research, in press*

Vertical velocity, vertical diffusion, and dilution by midlatitude air in the tropical lower stratosphere

Philip W. Mote,[1] Timothy J. Dunkerton,[1] Michael E. McIntyre,[2] Eric A. Ray,[3] Peter H. Haynes,[2] and James M. Russell III[4]

Abstract.

Air passing upward through the tropical tropopause is "marked" by an annually varying water vapor mixing ratio much as a tape recorder marks a magnetic tape; as the air ascends in the tropical stratosphere, these marks are effaced by a combination of vertical diffusion within the tropics and dilution of tropical air by sideways (isentropic) mixing-in of midlatitude air. We represent these processes using a one-dimensional advection-diffusion-dilution model, which we inverse-solve for the vertical profiles of three unknowns (vertical advection velocity, vertical diffusion coefficient, and dilution rate coefficient) after prescribing the vertical profiles of time mean methane [CH_4] and of amplitude and phase of the annually varying tape recorder signal in 2[CH_4]+[H_2O]. When tested on synthetic data generated by forward solving the same model, the method for inverse solution proved to be well conditioned and to give accurate results above 18 km. Applying the method to 5 years of smoothed data from the Halogen Occultation Experiment, we find a vertical advection velocity with a minimum of about 0.2 mm s^{-1} near 20 km, and both dilution rate coefficient and vertical diffusion coefficient with remarkably low minima near 22 km, 1/(6–7 year) and roughly 0.02 m^{2}s^{-1}, respectively. Our derived profile of vertical advection velocity agrees well, between 18 and 24 km, with an independent, radiatively derived, mass-budget-constrained transformed Eulerian mean calculation. Despite the

relatively modest values of the diffusion coefficient, vertical diffusion plays a significant role in attenuating the tape recorder signal, according to our model. The minimum value of the dilution rate coefficient corresponds to a relaxation timescale of 6–7 years, much longer than the timescales found in other studies. The long relaxation timescale at 20–24 km is, however, consistent with (1) the minimum in vertical velocity, (2) a reduced attenuation rate in the tape recorder signal, and (3) a decrease, hitherto unremarked, in the tropical vertical gradient of [CH_4] there.

1. Introduction

Profiles of water vapor in the tropical lower stratosphere have a vertical structure and time dependence that are determined, to a first approximation, by the ascent of the seasonally varying tropopause-level water vapor mixing ratio [*Mote et al.,* 1995, 1996; *Weinstock et al.,* 1995]. The tropical stratosphere thus resembles a "tape recorder" in that the water vapor content of rising air is "marked" by the seasonally varying saturation mixing ratio that it encountered at the tropical tropopause. These marks are slowly effaced as the air rises but are still discernible at 10 hPa, about 18 months after the air passed through the tropopause. *Mote et al.* [1996] (hereinafter M96) used data primarily from the Upper Atmosphere Research Satellite (UARS) to deduce several relevant aspects of the circulation and thermodynamics of the tropical lower stratosphere: the mean ascent rate between 100 and 10 hPa, some effects of the quasi-biennial oscillation (QBO), some characteristics of tropical troposphere-to-stratosphere mass transfer, and, for the 100 to 46 hPa and 46 to 10 hPa layers, upper bounds on the order of magnitude of vertical diffusion and on that of dilution of tropical air by sideways (isentropic) mixing-in of midlatitude air.

In this article we refine the work of M96 in three ways. First, we calculate extended empirical orthogonal functions (EEOFs) of the tape recorder signal in $2[CH_4]+[H_2O]$; the first two EEOFs are a conjugate pair describing a vertically propagating annual cycle. These two EEOFs and their associated time series are sufficient to reconstruct a smooth time-altitude plot of the tape recorder signal that is a good likeness of the time-altitude plot made directly from HALOE data (Plate 1). Second, by finding the slopes of features (extrema and zero crossings) on the smooth time-altitude plot (Plate 1b), we derive a vertical profile of the ascent rate w_{tr} of marks on the tape. Third, we obtain improved estimates of vertical diffusion and sideways dilution by fitting the profile of w_{tr} and other information in a one-dimensional (1-D) advection-diffusion-dilution model of the tropics. More precisely, the model fitting uses not only w_{tr} but also the amplitude of the tape recorder signal together with the vertical profile of time mean $[CH_4]$. We thus obtain consistent vertical profiles of the vertical advection velocity w (not necessarily the same as w_{tr}), the vertical diffusion coefficient K, and the rate or relaxation coefficient α for dilution by midlatitude air. The results indicate that in the altitude range from about 18 to 24 km, w_{tr} is close to the vertical advection velocity w, implying that for those altitudes the tape recorder signal can be taken at face value as a good indicator of mean vertical motion. The extraction of all three vertical profiles $w(z)$, $K(z)$, and $\alpha(z)$ is a well-conditioned, hence credible, calculation only because water vapor, with its high information content concerning the annual cycle, is used simultaneously with a long-lived trace gas, in our case $[CH_4]$.

Several other efforts have recently been made to characterize the exchange of air between tropics and middle latitudes. Some focused on mixing out from the tropics to middle latitudes by advecting passive tracers using observed winds [*Chen et al.,* 1994; *Waugh* 1996], while others considered 1-D budgets of observed long-lived trace gases, thus focusing on dilution of tropical air by mixing in from middle latitudes [*Avallone and Prather*, 1996; *Minschwaner et al.,* 1996; *Volk et al.,* 1996]. *Remsberg and Bhatt* [1996] used the zonal variance of nitric acid to infer the (qualitative) altitude dependence of dilution. In a different approach, *Schoeberl et al.* [1997] related dilution to the phase lag between water vapor mixing ratios and QBO winds. *Hall and Waugh* [1997b] independently derived average lower stratospheric values for w, K, and $1/\alpha$ using the tape recorder signal in $2[CH_4]+[H_2O]$ and a long-lived trace gas with a secularly increasing trend (sulfur hexafluoride). Our study differs from those just cited in that it simultaneously derives vertical velocity, vertical diffusion, and dilution rate coefficient as functions of altitude between 100 and 10 hPa, leading to a clear picture of the vertical structure of each of these three quantities.

2. Data

The constituent (CH_4 and H_2O) data used in this paper come from the Halogen Occultation Experiment (HALOE) aboard UARS. M96 used version 17 data ending July 1995; here we use version 18 data between September 1992 and August 1997. Time series are formed from monthly means, each comprising all tropical soundings for the month. We define the tropics as 14°S to 14°N, broad enough to include a greater number of profiles while still within the well-mixed region bounded by the subtropical mixing zones. Following M96, we use the quantity $\hat{H}=2[CH_4]+[H_2O]$, which has the advantage of being nearly conserved and homogeneous in the extratropical stratosphere, because photochemical breakdown of 1 mol of $[CH_4]$ produces approximately 2 mol of water vapor.

The vertical resolution of the data is the same as in M96 and is finer than that usually used (e.g., *Randel et al.* [1998]). HALOE level 2 data (where "level" refers to the level of processing) are available on a 0.3-km grid and are here interpolated to pressure levels 100.0, 82.5, 68.1,... hPa with a fractional spacing of $10^{1/12}$ in pressure. These pressure levels were converted to geopotential height using data from the U.K. Meteorological Office (UKMO) analyses [*Swinbank and O'Neill*, 1994], which form part of the UARS database. The spacing between HALOE pressure levels ranges from about 1.1 to 1.3 km.

3. Methods

3.1. Extended EOF Analysis

Plate 1a updates Plate 1b of M96 with about 2 more years of data. While the general sense of rising moist and dry anomalies is apparent, we wish to isolate the phenomenon of rising anomalies from other phenomena; to do this it will prove useful to apply a statistical technique known as extended EOF analysis [*Weare and Nasstrom,* 1982, *Wang et al.,* 1995].

Empirical orthogonal functions can identify coherent variations in noisy data by finding eigenvectors of the covariance matrix. In the case of tropical HALOE data, measurement noise is exacerbated by the sparse sampling that is characteristic of occultation techniques; EOF analysis can be performed using each month's profile to yield a clearer picture of vertically coherent variations. If the covariance matrix is formed by calculating covariance not of a single profile at each time but of several profiles at different lag times (in this case between -6 and $+6$ months), the resulting "extended EOF analysis" (EEOF analysis) yields a clearer picture of both vertically and temporally coherent variations.

We perform EEOF analysis on the HALOE $\hat{\mathrm{H}}$ data shown in Plate 1a, yielding a 4-year time series of coefficients since 6 months have been lost at each end. The first two EEOFs have a strong annual cycle and together they account for 68% of the variance. The next two EEOFs vary interannually and explain 17% of the variance. The coefficients of the first two EEOFs, when plotted against each other, trace nearly perfect circles with a period of one year, indicating that they are a conjugate pair. The coefficients of the next two EEOFs trace somewhat less than two orbits in 4 years, similar to the singular value decomposition results shown by *Randel et al.* [1998] for [CH_4].

With the coefficients of the first two EEOFs, multiplied at each time by the zero-lag profile of the EEOFs (and the missing 6 months at each end filled in by the EEOFs at lag $\pm 1, \pm 2, ... \pm 6$), we construct an altitude-time plot of $\hat{\mathrm{H}}$ anomalies (Plate 1b). When compared with the altitude-time plot of HALOE $\hat{\mathrm{H}}$ data (Plate 1a), the first two EEOFs produce almost a pure annual cycle. In the resulting picture, ascending minima and maxima of $\hat{\mathrm{H}}$, which are formed at the tropopause, retain their identity from the tropopause to 10 hPa (about 31 km).

From the tape recorder signal (for example, as represented smoothly in Plate 1b), two quantities can be deduced directly, both as functions of altitude: w_{tr}, the ascent rate of marks on the tape, and A, the amplitude of the signal. Both quantities depend on altitude. In the next subsection we consider A. The quantity w_{tr} (Figure 1) was obtained from the data in Plate 1b by following each ascending minimum, maximum, and zero from level to level on successive time series plots. Also shown in Figure 1 for later comparison is a transformed Eulerian mean vertical velocity $\bar{w}^*$ obtained by K. Rosenlof (personal communication, 1996), updating that of *Rosenlof* [1995]. (In section 5.2 we discuss whether we expect w_{tr} and $\bar{w}^*$ to be equal. Table 1 summarizes the definitions of these quantities.)

Rosenlof has recalculated $\bar{w}^*$ to make use of more UARS data in the radiative calculation, and the agreement between her new $\bar{w}^*$ (the curve marked by asterisks in Figure 1) and our w_{tr} is excellent, at least between 19 and 24 km. The standard deviations are calculated for 12 monthly profiles of $\bar{w}^*$, hence reflecting mainly seasonal variation, and are calculated for 20 separate features (extrema or zero-crossings of $\hat{\mathrm{H}}$) for w_{tr}, hence reflecting both uncertainty in our method of calculating w_{tr} and possibly seasonal and interannual variation.

3.2. Definition of Amplitude

Underlying the model-fitting concept to be used here is the tentative assumption that the pattern seen in Plate 1 is "typical" in the sense that, if previous or subsequent years were visible, one would see a very similar pattern, i.e., something like the periodic extension of Plate 1 in both time directions. To the extent that this is justified, one can talk about "the annual cycle," and by implication its amplitude and phase, with less ambiguity than is strictly the case with a record only 5 years long. With this point in mind, we seek to define the amplitude and, in section 3.3, the phase.

As was already mentioned, we use $\hat{\mathrm{H}}$ as the basis for estimating the amplitude A of the tape recorder signal. Variations in $\hat{\mathrm{H}}$ are due almost solely to the tape recorder signal, while $[H_2O]$ is more strongly affected by other kinds of variability such as the QBO or the semiannual oscillation (SAO). Indeed, we find that the variance of $[H_2O]$ exceeds that of $\hat{\mathrm{H}}$ above 25 km. Other sources of variability that might affect our estimation of A include the HALOE sampling pattern, instrumental or retrieval errors, or incursion of the subtropical barrier into the tropical region (14°S to 14°N). Such variability can be reduced by the use of EEOFs, since the EEOF technique identifies coherent variations.

Because of the variation in vertical velocity caused by the QBO, the temporal variations in $\hat{\mathrm{H}}$ become slightly distorted, with extrema shifted as much as 2 months away from a perfect 12-month period [M96]. Consequently, taking the mean annual cycle or applying Fourier analysis at each level independently (as was done by *Randel et al.* [1998]) may underestimate the tropopause-related variance above 20 km or so. Similarly, EEOF analysis links variations at different altitudes so that the stretching and compressing of the tape recorder signal by the QBO [M96; *Cordero et al.,* 1997] would be smeared out a bit, possibly leading to an underestimate of the variance associated with the tape recorder signal. In contrast to the mean annual cycle, however, the coefficients of the first two EEOFs have a small degree of interannual variability, allowing for some QBO effects on the tape recorder signal. We therefore expect that an amplitude definition based on EEOF analysis will be more successful than other definitions in distinguishing between variations originating at the tropical tropopause and variations arising from other sources.

With those considerations in mind, we present plots (Figure 2) of the logarithm of amplitude defined in several ways, all but one using $\hat{\mathrm{H}}$. Since only derivatives of $\ln A$ will be needed, the curves have been normalized by their amplitude at 100 hPa in order to emphasize their differences. Curve (a) shows the standard deviation, at each level, of the time series of $\hat{\mathrm{H}}$ (Plate 1b, the EEOF reconstruction). Curve (b) is the annual harmonic from fast Fourier transform analysis of monthly mean HALOE data (Plate 1a). Curve (c) is based not on HALOE $\hat{\mathrm{H}}$ data but on time series of tropical water vapor measured by the Microwave Limb Sounder (MLS) aboard UARS (see M96 for details). Curve (d) is obtained by first removing a 12-month running mean from the data in Plate 1a, then following individual maxima and minima upward, and taking the difference between the average maximum and the average minimum at each level. Curve (e) is the standard deviation of the time series at each level in Plate 1a ($\sigma(\hat{\mathrm{H}})$). Finally, the solid curve with plus marks is the peak-to-peak amplitude of the seasonal cycle, composed of the annual and semiannual harmonics of $\hat{\mathrm{H}}$, as derived by *Randel et al.* [1998] for the latitude range 4°S to 4°N on a subset of our levels.

By most definitions, the amplitude decreases monotonically with altitude, but all show a region of near-constant A (a "cliff") somewhere between 20 and 25 km. The total attenuation from 100 to 10 hPa is greatest for curve (a) and least for curve (e), with curves (b)–(d) clustered near the mean of these two extremes.

Our amplitude curves differ from that of *Randel et al.* [1998], probably because the narrow latitude range they chose (4°S – 4°N) results in far fewer good data points (especially at 100 hPa, where noise levels are high) and a less coherent annual signal.

For calculating first and second derivatives, it will be useful to have a smooth approximation $\ln f$ for $\ln A$. As was outlined above, we expect the curve based on EEOFs, curve (a), to provide the best definition of amplitude, and we fit smooth curves f to it and also to curve (d) in order to bracket the reasonable range of definitions of f. The data suggest that vertical attenuation rates are relatively fast at upper and lower levels of the domain, and relatively slow in the middle (near 20–23 km). It is not clear, however, how steep the cliff in the middle should be. One possibility is that the cliff is spurious. Another possibility is that the cliff is real and reflects a sharp decrease of attenuation rate in the middle layer, as might be expected in the lower part of the QBO for dynamical reasons [see *Dunkerton,* 1997].

We smooth the amplitude curves by first linearly interpolating them to high (approximately 70 m) vertical resolution, then applying a 1-2-1 filter to the second derivative 200, 500, or 1000 times, and finally inverting (integrating twice) to yield a set of curves f (Figure 3) with a range of behavior near the "cliff." Smoothed 1000 times, the cliff almost disappears in both cases. The curve labeled M is curve (a) smoothed 500 times, while curve M^* was obtained by constraining the attenuation rate to a constant above the cliff (for reasons described in section 3.4). The various curve fits provide a range of profiles spanning the uncertainty of the original estimates of A, thereby generating a range of values for w, α, and K.

3.3. One-Dimensional Model for Analysis

In order to derive vertical profiles of α and K, we assume that to within reasonable noise levels, the smoothed and meridionally averaged tropical data can be fitted to solutions of the simple one-dimensional advection-diffusion-dilution model

$$\chi_t + w\chi_z = \frac{1}{\rho}(\rho K \chi_z)_z - \alpha(\chi - \chi_{\mathrm{ML}}) + S \qquad (1)$$

where χ is a tracer mixing ratio, w is the vertical velocity, K is the vertical diffusion coefficient, $\rho = \rho_0 \exp(-z/H)$ is the mean density profile with $H = 7$ km, α is the dilution rate, χ_{ML} is the midlatitude value of χ, and S is a chemical source-sink term (of the form $-\gamma\chi$ for [CH_4], 0 for $\hat{\mathrm{H}}$). Partial derivatives with respect to time t and height z are indicated by subscripts.

Implicit in this model are several assumptions. (1) Tropical air is horizontally homogeneous within well-defined latitude limits (here, 14°N to 14°S) and is distinct, though not completely isolated, from midlatitude air. (2) The same values of w, K, and α apply to both [CH_4] and $\hat{\mathrm{H}}$ (see *Thuburn and McIntyre* [1997], for a discussion of this assumption). (3) As is commonly done [e.g., M96; *Volk et al.,* 1996], we assume dilution by midlatitude air on a given level to be a linear relaxation process in which tropical air tends toward the midlatitude mixing ratio χ_{ML} with relaxation rate α and with zero phase lag. In fact, the seasonal cycle of water vapor (hence $\hat{\mathrm{H}}$) in the lower stratosphere (below 20 km) at middle latitudes is strongly influenced by the tropical seasonal cycle, presumably by mixing out, with a phase lag of a month or two [e.g., *Mastenbrook and Oltmans,* 1983; *Rosenlof et al.,* 1997], This means that especially below 20 km the assumption of uniform extratropical $\hat{\mathrm{H}}$ fails, and dilution is a much more complex process than is assumed here.

With these caveats in mind, we first rewrite (1) as

$$\chi_t + \tilde{w}\chi_z = K\chi_{zz} - \alpha(\chi - \chi_{\mathrm{ML}}) + S \tag{2}$$

where

$$\tilde{w} \equiv w + (K/H) - K_z \tag{3}$$

arises when we expand the diffusion term in (1). Note that $\tilde{w}$ does not have a physical meaning but is introduced for mathematical convenience. Some of the advective effects of diffusion are represented in $\tilde{w}$; later we compare the variants of w listed in Table 1.

One usually solves differential equations like (2) for the dependent variable, in this case χ, but we want to find the three unknowns $\tilde{w}$, K, and α, through a process that could be called "inverse solving." There are a few possible methods for inverse solving (2) for the three unknowns. After testing these methods on synthetic data that were constructed (as is explained in the next section) from known profiles of $\tilde{w}$, K, and α, we chose one method that is both accurate and concise. (Another method, involving WKBJ or slow modulation theory, is outlined in Appendix A and will prove useful in sections 3.4 and 3.5.) Our method makes use of information about $[\mathrm{CH_4}]$ and $\hat{\mathrm{H}}$; specifically, we represent $\hat{\mathrm{H}}$ anomalies by the functional form $\chi = \mathrm{Re} f(z)\exp[i(\phi(z) - \omega t)]$, where $f(z)$ and $\phi(z)$ are real and are determined from $\hat{\mathrm{H}}$ data and where the midlatitude value of $\hat{\mathrm{H}}$ is assumed to be spatially and temporally constant at the mean value of the tropical $\hat{\mathrm{H}}$. Applying (2) to time mean $[\mathrm{CH_4}]$ and to the functional form of $\hat{\mathrm{H}}$ given above yields a set of three equations in three unknowns:

$$\begin{aligned} \alpha\,([\mathrm{CH_4}] - [\mathrm{CH_4}]_{ML}) \quad &+\tilde{w}\,[\mathrm{CH_4}]_z \\ - K\,[\mathrm{CH_4}]_{zz} \quad &= -\gamma[\mathrm{CH_4}] \end{aligned} \tag{4a}$$

$$\begin{aligned} \alpha &\quad +\tilde{w}\,(\ln f)_z \\ +K\,(m^2 - f_{zz}/f) &\quad = 0 \qquad \text{(4b)} \\ &\quad \tilde{w}\,m \\ +K\,(-2m(\ln f)_z - m_z) &\quad = \omega \qquad \text{(4c)} \end{aligned}$$

where $m \equiv \phi_z$. At each vertical grid point we solve (4a)–(4c) simultaneously for α, $\tilde{w}$, and K. This method is well-conditioned if the determinant is non singular, but well-conditionedness is no guarantee that the three quantities will be positive everywhere or have smooth profiles.

The input data needed are time mean profiles of $[CH_4]$ in the tropics and in middle latitudes, and profiles of γ, f, and m. The $[CH_4]$ profiles are taken from the climatology of *Randel et al.* [1998] using equivalent latitude limits of 30°–54° in both hemispheres for $[CH_4]_{ML}$; these limits were chosen to match the "surf zone" (see their Figure 7), where midlatitude values of trace constituents have only weak meridional gradients. The chemical reaction rates γ are taken from the $2\frac{1}{2}$-dimensional model of *Kinnersley* [1995]. The previous subsection described how we arrived at the smooth profiles f.

We have not yet defined ϕ, the vertical phase function from which m will be defined. It is possible, in principle, to derive ϕ directly from the HALOE data in Plate 1 and then differentiate to give m. We chose instead to derive m from the velocity w_{tr} already plotted in Figure 1; as is indicated in equation (B2), features of χ such as zero crossings rise at a rate equal to w_{tr}. In this context, attenuation does not affect the phase ϕ, so it is appropriate to consider the expression $\chi(z,t) = \cos(\phi(z) - \omega t)$, where ω is the annual frequency and ϕ is the phase. Differentiating χ first with respect to z and then with respect to t and applying (B2) yields the relation

$$m w_{\mathrm{tr}} = \omega \tag{5}$$

so that m is defined in terms of w_{tr}. The profile of m (not shown) has the shape one would expect for the reciprocal of w_{tr}, with a maximum at about 20 km.

3.4. Some Approximate Answers

In the discussion that follows, it will be useful to have rough estimates of the quantities to be calculated, as well as plausible ranges of values. It is possible to determine α directly from $[CH_4]$ data by taking the time mean of (2) and neglecting the diffusion term:

$$\alpha = \alpha_{\mathrm{m}} \equiv \frac{\tilde{w}\chi_z + \gamma\chi}{\chi_{\mathrm{ML}} - \chi} \tag{6}$$

with χ=$[CH_4]$. The profile of α_{m} is the solid curve in Figure 4a.

Some idea of extreme values of α and K can be estimated using the WKBJ approach set out in Appendix A. More accurate calculations based on (4a)–(4c) will check these estimates. The zeroth-order WKBJ equation is similar to (5), but with $\tilde{w}$ instead of w_{tr}, while the first-order equation (A6) is

$$-\tilde{w}(\ln f)_z = K m^2 + \alpha \tag{7}$$

which illustrates the simple relationship between the two observables, $\tilde{w}$ and f, and the two agents of attenuation, K and α.

From this equation alone, K and α cannot be uniquely determined, but we can estimate limiting values. We assume that $\tilde{w} \approx w_{tr}$ and alternately set $K = 0$ and $\alpha = 0$ to derive, respectively, the largest α (α_{max}) and the largest K (K_{max}) that are consistent with the observed attenuation of the smooth tape recorder amplitude f under the WKBJ assumption. Also, with $\alpha = \alpha_{\mathrm{m}}$ in (7), a profile of

K (K_{m}) can be derived that is consistent with α_{m}. That is,

$$-\tilde{w}(\ln f)_z = K_{\max} m^2 = \alpha_{\max} = K_{\mathrm{m}} m^2 + \alpha_{\mathrm{m}} \quad (8)$$

Figure 4 shows profiles of $\alpha_{\max}$, α_{m}, $K_{\max}$, and K_{m}. (For $\ln f$ we have used the curves M and M^* in Figure 3, and solutions using $f = M^*$ are indicated by plus symbols.) The profiles in Figure 4a of α_{m} (solid curve) and $\alpha_{\max}$ (dashed curves) both have a minimum of about 0.01–0.02 $\times\, 10^{-6}\mathrm{s}^{-1}$ at 22 km, with values increasing toward the tropopause. $K_{\max}$ (Figure 4b, dashed curves) has a minimum of about 0.02 $\mathrm{m}^2\mathrm{s}^{-1}$ also at 22 km and reaches much larger values near the tropopause. Where α_{m} exceeds $\alpha_{\max}$, K_{m} (solid curves) is negative; this occurs below 17 km and above 29 km (for M).

Solving (8) for the seven profiles of f shown in Figure 3, we find (results not shown) that $\alpha_{\max}$ above 18 km always lies between 0.01 and $0.07 \times 10^{-6}\mathrm{s}^{-1}$ and is sometimes smaller than α_{m} (implying that $K_{\mathrm{m}} < 0$) above 27 km. Both $\alpha_{\max}$ and $K_{\max}$ have a minimum near 22 km, since (from (7)) they vary as the product of the fractional attenuation rate and a power (1 for $\alpha_{\max}$, 3 for $K_{\max}$) of $\tilde{w}$. Curves fitted to curve (d) in Figure 2 tended to have smaller values of $\alpha_{\max}$ and $K_{\max}$ above 24 km than did those fitted to curve (a). The profiles of α_{m} and K_{m} for M^* were the most plausible, with α_{m} remaining fairly constant and K_{m} positive everywhere above 18 km. This is a consequence of fixing the attenuation rate to a constant above the cliff.

Most other researchers investigating the dilution of tropical air by midlatitude air [e.g., *Volk et al.,* 1996] have discussed, instead of the dilution rate α, the dilution timescale $\tau = 1/\alpha$. This quantity is shown in Figure 4c for the three profiles of α in Figure 4a. The minimum timescale (dashed curves, corresponding to maximum α) is about 5–15 months above 24 km and has a sharp peak of nearly 2 years at 22 km.

We emphasize that the curves shown in Figure 4 are not our best estimates of these quantities. The dashed curves represent extreme cases, under the WKBJ assumption, where only one of the two processes (dilution and diffusion) operates. In reality, $0 < K < K_{\mathrm{max}}$ and $0 < \alpha < \alpha_{\mathrm{max}}$. At each altitude there is (in the full solution) a unique combination of K and α consistent with the observed attenuation. Before solving (4a)–(4c) (the full, non-WKBJ set) to determine that unique combination, we first test the method using synthetic data.

3.5. Generation of Synthetic Data

We generate synthetic $[CH_4]$ and $\hat{\mathrm{H}}$ data using various profiles of w, K, and α to solve (2). We can then inverse solve the synthetic data to derive profiles of w, K, and α, and compare these with the input profiles. This provides a complete, end-to-end check on the forward and inverse methods and codes, as well as a check that the inverse method, to be used on smoothed real data, is well conditioned and robust.

The simple model used to generate the synthetic data, then, is a discretized form of (2) and is run five times, using the five combinations of K and α given in Table 2. The profile of w used in each case is simply w_{tr} from Figure 1, but because equation (2) uses $\tilde{w}$ instead of w, and $\tilde{w}$ depends on K (see equation (3)), $\tilde{w}$ is slightly different in each case. Advection is accomplished using essentially a semi-Lagrangian method, with the other terms time-split and semi-implicit; vertical resolution is quite fine (50 m), as is temporal resolution (0.5 day). To prevent error growth, the second derivative is evaluated on a reduced grid whose vertical resolution approximates that of the original HALOE data; then the result is interpolated to the 50 m grid. The chemical loss rate γ for $[CH_4]$ is again taken from the model of *Kinnersley* [1995].

The time mean $[CH_4]$ profile is calculated by integrating the model to steady state. The lower boundary value and midlatitude profiles are time mean values from *Randel et al.* [1998]. For $\hat{\mathrm{H}}$ the lower boundary varies sinusoidally with a minimum in mid-February and an amplitude of 1. Both S (chemical source-sink) and χ_{ML} are zero. The model is integrated for 4 years, and only the last year of data is used in the subsequent calculation.

Figure 5 shows profiles of both time mean observed and steady state synthetic [CH_4]. The synthetic profiles were both obtained using $\alpha = \alpha_{\mathrm{m}}$, one with $K = 0$ and one with $K = K_{\mathrm{max}}$ (see Figure 4b). Over most of the domain, the synthetic profile with $K = 0$ falls within 0.01 ppmv of the observed profile. For $K = K_{\mathrm{max}}$, however, differences are larger owing to the advective effect of $-K_z$ (equation (3)), which for K_{max} is quite large in the lower stratosphere up to 21 km.

We also calculated time-varying $\hat{\mathrm{H}}$ for the five combinations of K and α shown in Table 2. With each of the combinations of K and α, the resulting $\hat{\mathrm{H}}$ (Plate 2) qualitatively resembles Plate 1b. What distinguishes them is the rate of attenuation, which we will examine shortly, and the apparent ascent rate. In our 1-D model we can unambiguously identify the model's advective velocity ($\tilde{w}$, which includes diffusive effects) with the phase speed; in the real atmosphere the two quantities are not generally equal, though the tropical stratosphere is the only region where they are close [*Hall and Waugh,* 1997a]. The transit time from 100 hPa to 10 hPa of a (temporal) maximum or minimum decreases with increasing K, from 18.1 months for $K = 0$ to 17.1 months for $K = K_{\mathrm{m}}$ to 16.0 months for $K = K_{\mathrm{max}}$. The transit time estimated from HALOE data is 18.6 months. Note that in our 1-D model we have used $\tilde{w} = w_{\mathrm{tr}} + (K/H) - K_z$, so the faster ascent for nonzero K is to be expected. The transit time also depends weakly on α, differing by 14 days between combinations (iv) ($\alpha = 0$) and (v) ($\alpha = \alpha_{\mathrm{m}}$) but only by 1.5 days between combinations (i) ($\alpha = \alpha_{\mathrm{m}}$) and (ii)

($\alpha = \alpha_{\max}$).

The input profiles of K and α determine not just the ascent rate but also the amplitude of $\hat{\mathrm{H}}$ anomalies (Figure 6). Combination (ii) reproduces the observed profile M^* fairly well. Combinations (ii) and (iv) represent (from (7)) the extreme cases of no diffusion and no dilution, respectively, and since from (8) (ii), (iii), and (iv) all have the same profile of $Km^2 + \alpha$, they would produce the same attenuation were it not for the fact that $\tilde{w}$ is different for each one. For those combinations with $K \neq 0$ ((iii), (iv), and (v)), the ascent rate is quite fast in the lowest few kilometers of the stratosphere owing to the influence of diffusion; consequently, the tropical air has less time to be modified by diffusion and dilution, so there is less attenuation of the tape recorder signal at a given altitude.

4. Results

4.1. Deducing Known Profiles of w, K, and α From Synthetic Data

In order to test our method of inverse solving for K and α on the synthetic data, we must first determine m. This is done by calculating an ascent rate w_{synth} for each of the synthetic $\hat{\mathrm{H}}$ time series, in a manner similar to that used to derive w_{tr}. Using the profile of w_{synth} for each combination, we calculate a corresponding profile of m for inverse solving equations (4a)–(4c).

For each of the five combinations of K and α in Table 2, we sample the synthetic data at vertical resolutions ranging from 1.2 km (similar to HALOE data) to the 50 m grid of the synthetic data, interpolate back to the 50 m grid, then inverse-solve (4a)–(4c) at 50 m resolution. As is shown in Figure 7 for combination (iii) sampled at 1.2 km resolution, and for the other combinations and sampling resolutions (not shown), our method is fairly successful at recovering the input profiles of w ($= \tilde{w} - (K/H) + K_z$, from equation (3)), α, and K, at least between 18 and 28 km. Below 18 km the solution often differs substantially from the input values, a problem that is exacerbated at higher resolution. The problems below 18 km occur because the determinant of the matrix arising from (4a)–(4c) is small. In the high-resolution calculation, RMS errors (over the five combinations) between 18 and 28 km are small: 0.01–0.03 mm s^{-1} for w, $1 - 7 \times 10^{-9}$ s^{-1} for α, and 0.003–0.03 m^2s^{-1} for K. This method is clearly well suited for our purposes.

4.2. Deducing K and α From HALOE Data.

With some confidence that our method will give good results above 18 km, we apply it to each of the seven smooth amplitude curves shown in Figure 3. We use the observed [CH_4] profile in Figure 5 and other input data as described in section 3.3. As with synthetic data, the HALOE data are interpolated to a 50 m vertical grid for calculations, then subsampled on a 1.2-km vertical grid for plotting. The results are shown in Figure 8. In each panel the mean is shown as the solid curve, bounded by $\pm 1\sigma$, where σ is calculated over the six smooth curves (not including M^*). The derived profile of w is very similar to w_{tr} (long-dashed curve) but is somewhat lower above 24 km.

All seven definitions of f yield results for α that are substantially similar to each other and to α_{m} (Figure 8b), for reasons that will be discussed shortly. The minimum values are less than $1.0 \times 10^{-8}\mathrm{s}^{-1}$, and $\alpha < 4.0 \times 10^{-8}\mathrm{s}^{-1}$ everywhere above 18 km. The reciprocal of α, τ (Figure 8c), indicates far longer timescales than in previous research, for reasons that we discuss in section 5.4. At this point it is sufficient to note that the range of values above 18 km is 1.5–7 years, much longer than the advective timescale.

Profiles of K are reasonably consistent for all seven curves between 18 and 25 km, but differ markedly from each other above and below this altitude range.

For most definitions of f, K decreases again above 24 km and for many, K becomes negative at some altitude. The profiles with $f = M^*$ and with curve a smoothed 1000 times are the only ones for which $K > 0$ everywhere above 18 km. Below 18 km, most profiles are negative. The tentative extension of the solution to 100 hPa will be explained in section 5.2.

It is worth noting that a conceptually simpler approach, the WKBJ method outlined in Appendix A, yields very similar results. In essence, the WKBJ method says that $w = w_{\mathrm{tr}}$, $\alpha = \alpha_{\mathrm{m}}$ (from equation (6)), and $K = K_{\mathrm{m}}$. That is, $[CH_4]$ tells us about dilution, and the role of diffusion can be deduced from the excess attenuation of the $\hat{\mathrm{H}}$ anomalies above that expected from dilution alone.

5. Discussion and Conclusions

The ability of the first pair of extended EEOFs to reproduce the essential features of the tape recorder signal (Plate 1) underscores the simplicity of the phenomenon; marks left in the air by annual variations of tropopause temperature are the dominant variation in tropical water vapor below 10 hPa. The somewhat ragged HALOE data are smoothed nicely by EEOFs, allowing us to estimate directly the ascent rate of the tape recorder signal, w_{tr}. Knowing K allowed us to extract a profile of w (Figure 8a) from $\tilde{w}$.

Having simultaneously derived profiles of the three quantities w, α, and K, we now consider in more detail the assumptions and weaknesses in the calculation. We also discuss the implications of our findings and compare them to those of other researchers.

5.1. Comments on the Method of Deriving K and α

Several aspects of the calculation bear consideration: the reliability of the data, our definition of tape recorder amplitude, the curve-fitting techniques used, the appropriateness of equation (1), and our numerical approach to finding K and α (equations (4a)–(4c)).

The papers validating HALOE H_2O [*Harries et al.*, 1996] and CH_4 [*Park et al.*, 1996] provide estimates of the systematic and random errors of these two quantities (their Tables 1). The errors grow with pressure in the lower stratosphere; for example, the root-sum-square of all sources of error for CH_4 is 11% at 10 hPa and 19% at 100 hPa. While the larger errors at 100 hPa may affect our results, we believe this effect is minimal for several reasons. First, the error tables refer to individual profiles, but we have used monthly means; for large numbers of profiles (usually 50 or more) the random errors average to zero, leaving systematic errors at 100 hPa of 15% for CH_4 and 24% for H_2O. Second, $\hat{\mathrm{H}}$ data enter the solution mainly through $(\ln f)_z$. The only systematic errors that could have a significant effect on $(\ln f)_z$ are those with a seasonal, altitude-dependent bias; a uniform positive or negative bias at any level would have no effect on f at that level, nor would a seasonal bias that was independent of altitude. While it is possible that a seasonal, altitude-dependent bias exists, it would probably not be the largest component of the systematic errors. A known bias exists in the lower stratosphere between profiles taken during sunrise and those taken during sunset, but over the 5 years there is no seasonal bias between sunrise and sunset profiles. Finally, one can estimate the seasonal variations in $\hat{\mathrm{H}}$ at 100 hPa using observed 100 hPa temperatures to calculate saturation mixing ratios [M96]; these place an upper bound of about 3 ppmv on the amplitude, compared to 1.1–1.5 ppmv using the various definitions employed here. If we have underestimated the amplitude, that would merely serve to increase α and K below 18 km, where they are already large and perhaps unreliable.

A potential weakness of our calculation is the actual definition of the amplitude of the tape recorder signal. As is discussed in section 3.2, this definition should distinguish variations whose source is the tropical tropopause from all other variations. Furthermore, since α and K must be positive, and since w_{tr} is observed to be positive, by (7) the smoothed amplitude profile f must be monotonically decreasing with altitude. In our opinion, the definition using EEOFs (curve a in Figure 2) is the best one because the variations seen in Plate 1b are clearly linked to the tropopause. By some other definitions the amplitude decreases less quickly with altitude above 24 km, but these other definitions are more susceptible to variations attributable to other causes as discussed in section 3.2. Only two curves, one of them M^*, lead to positive K above 28 km, and it seems reasonable to conclude that the other curves are not decreasing fast enough above 28 km.

The next point of concern is the curve fitting. For a model like (1) to succeed, smoothing must be performed on the profiles of w_{tr} and especially the amplitude A. We have attempted to bracket the range of plausible curves by using two definitions of amplitude and three degrees of smoothing.

One might well ask whether the 1-D model represented in equation (1) accurately depicts the dynamics of the tropical stratosphere. Perhaps the greatest weakness in the model is the treatment of dilution as a linear relaxation process, especially below 20 km, where the tropics and middle latitudes are influenced by the swirl of monsoonal and other three-dimensional circulations protruding from the troposphere [*Dunkerton,* 1995]. Vertical diffusion too has more complex characteristics below 20 km. For these reasons and because of the poorer performance of the inverse solving method below 18 km even with "perfect" data (Figure 7), we have little confidence in our results below 18 km, especially for K.

Finally, we discuss our approach to inverse solving (2). Our method tests well on synthetic data except below about 18 km. It gives a definite result at every altitude, but with HALOE data at upper and lower levels it tends to give small or negative K except for the best profile M^*. Note that by assuming that temporally constant profiles of w_{tr}, α, and K apply to both [CH_4] and $\hat{\mathrm{H}}$, we ignore covariance terms like $w'\chi'_z$, leaving just $[w][\chi_z]$ (where $[w]$ represents a time mean and w' is a departure from that mean) and similar terms. Seasonal variations in w_{tr} [see *Rosenlof,* 1995], α [see *Bowman and Hu,* 1997], and K appearing in covariance terms would be very interesting, but their calculation is beyond the scope of this work.

5.2. Vertical Velocity

A preliminary version of our Figure 1 was presented by *Dunkerton* [1997], who compared a profile of w_{tr} derived from the (shorter, version 17) data used by M96 with the radiatively derived transformed Eulerian mean vertical velocity $\bar{w}^*$ of *Rosenlof* [1995] (which constrains the global vertical mass flux to zero); both showed a minimum ascent rate of about 0.2 mm s^{-1} at about 21 km. *Eluszkiewicz et al.* [1997] have also calculated $\bar{w}^*$, and a profile formed by averaging five of their profiles calculated using different inputs (not shown) generally falls within the error bars of the two curves in Figure 1 but has less vertical variation.

We now discuss the significance of similarities and differences among the quantities w, $\bar{w}^*$, w_{tr}, and $\tilde{w}$ (Table 1). In principle, $\bar{w}^*$ should be the same as the vertical advection velocity w. We consider two questions: (1) whether w_{tr} calculated mathematically with (B3) using the results of our inverse-solving method ($\tilde{w}$ and K) agrees with w_{tr} found using Plate 1b and (2) the implications of the differences between w_{tr} and $\bar{w}^*$ above 24 km and below 18 km.

To address the first question, we calculate the terms in the definitions of $\tilde{w}$ (equation (3)) and w_{tr} (equation B4) and plot the results in Figure 9. For K we have used the profile indicated by diamonds in Figure 8d, extended to 0.1

m^2s^{-1} at 100 hPa as indicated, for reasons that will shortly become clear. The vertical advection velocity w is shown as the solid curve; of all the curves, it has the smallest variation, being generally near 0.3 mm s^{-1}, but like the others it has a pronounced minimum at 20 km. Figure 9 implies that the true advection velocity is slower than the ascent rate of marks on the tape ($w_{\rm tr}$) and that this difference is due to diffusion. The K/H term (dotted curve) amounts to no more than 10% of $\tilde{w}$. The profile of $w_{\rm tr}$ derived mathematically from (B4) (dash-dot curve) is virtually identical to $w_{\rm tr}$ itself (diamonds), which is an encouraging demonstration of the consistency of $\tilde{w}$, K, and the derivatives used in calculating the approximation $1/h$ in (B5).

To address the second question, we consider first the differences above 24 km, where $\bar{w}^*$ is significantly larger than $w_{\rm tr}$, which in turn is larger than w. Assuming for the moment that our estimate of K is inaccurate and also that $\bar{w}^* = w$, (B4) could explain the difference only if K were large and negative, or K_z were large and positive, because the K terms compete with the $-K_z$ term. Because of this competition, only absurdly large values of K could explain the difference between $\bar{w}^*$ and $w_{\rm tr}$. It therefore seems that $\bar{w}^* \neq w$; either our estimate of $w_{\rm tr}$ is too small, or Rosenlof's $\bar{w}^*$ is too large. We note that Eluskiewicz's estimates of $\bar{w}^*$ are also smaller than Rosenlof's, generally 0.3–0.4 mm s^{-1}, and that a significantly larger w, and hence $w_{\rm tr}$, would mean that (for plausible values of K) $\hat{\rm H}$ anomalies would arrive at 10 hPa much faster than is observed.

Below 18 km our results generally suggest an increase of $\tilde{w}$, α, and K approaching the tropopause, but the uncertainty of K is relatively large, and more than one interpretation is possible. On the one hand, if $\bar{w}^*$ is qualitatively correct in this region, then a rapidly decreasing K might be responsible for the difference of about 0.2 mm s^{-1} between $\bar{w}^*$ and $w_{\rm tr}$. Such a profile of K

would be expected if vertical mixing were enhanced in a shallow layer owing to overshooting deep convection, as well as the local breaking of slow Kelvin and gravity waves launched by convective systems. Assuming a linear profile of K between 100 hPa and 83 hPa (as indicated by the dotted line in Figure 8d), it turns out that $K = 0.1\mathrm{m}^2\mathrm{s}^{-1}$ at 100 hPa will produce the difference of 0.2 mm s^{-1} between w and w_{tr}.

On the other hand, if $w_{\mathrm{tr}} \simeq w$, i.e., increasing rapidly approaching the tropopause (as in *Dunkerton*'s [1997] model of the QBO, for example) then K at 100 hPa must be much smaller and something must be missing from Rosenlof's radiative model, such that the actual diabatic heating rates are significantly larger at, and just above, the tropopause. One possibility is that heating associated with subvisible cirrus [*Wang et al.,* 1996], not included in Rosenlof's calculation, contributes to enhanced ascent in this layer.

Another explanation of the observed tape signal near the tropopause is that the "recording head," in reality, has a finite depth due to spatial and temporal variations of tropopause altitude, so that the source of the $\hat{\mathrm{H}}$ signal is somewhat fuzzy, giving the appearance of enhanced apparent ascent. To be sure, the observed vertical displacements of tropopause altitude occupy a significant fraction of the layer in which enhanced ascent is observed, and might therefore account, to some extent, for the observed signal. Nevertheless, none of the physical mechanisms described in the previous paragraph can be readily dismissed, and they merit further study. For example, knowledge of the population of overshooting convection, along the lines discussed by *Zhang* [1993], could (in principle) be used to model $K(z)$ for comparison to derived estimates. Similar models of diffusivity might be obtained from examination of breaking waves in long records of aircraft and rawinsonde data. A more thorough examination of these mechanisms will improve our understanding of stratosphere-troposphere exchange, dehydration, and radiative balance.

5.3. Diffusion

The derived profile of K has a minimum of about 0.02 m^2s^{-1} near 21 km; between 18 and 24 km, the values of K are similar for all definitions of f. Below 18 km the values of K are unreliable but probably increase downward. Above 24 km our results suggest a value of K around 0.1 m^2s^{-1}. These values are somewhat greater than the value of 0.01 m^2s^{-1} *Hall and Waugh* [1997b] derived for the layer between the 100- and 31.6-hPa HALOE levels, but they used *Randel et al.'s* [1998] definition of amplitude (plus symbols in Figure 2), which has smaller vertical derivative. Our values of K are somewhat less than the 0.07 m^2s^{-1} derived by *Remsberg* [1980] for 19°N, and are considerably less than the 1–7 m^2s^{-1} derived by *Patra and Lal* [1997] for 18°N, but the latter authors neglected vertical advection. The profiles of K that they showed, from their work and from several papers in the 1970s and 1980s, were otherwise qualitatively similar to ours, with a minimum around 19–21 km, sharp increases toward the tropopause, and slow increases up to 50 km.

Above 24 km the uncertainty in our derived K is very large and depends sensitively on the definition of the amplitude curve f. For most curves f in Figure 3, K becomes negative at some altitude. Only for $f = M^*$ and a highly-smoothed curve a does K remain positive.

The diffusion coefficient included in (1) represents the effects of a number of processes. These include small-scale processes such as turbulence due to gravity-wave breaking, but also larger-scale processes such as the diabatic dispersion discussed by *Sparling et al.* [1997]. In the tropical context the latter is likely to arise from the fact that as air parcels move longitudinally they experience different histories of radiative heating, due to the temporal and spatial variation of the heating field. Such variations may occur on the largest scales, e.g., in association with the warm pool region, or on mesoscales, e.g., in association with radiative effects of cirrus decks associated with individual convective systems.

5.4. Dilution

Unlike K, α is not very sensitive to the definition of f. All curves have a minimum of about 5×10^{-9} s^{-1}, somewhat smaller than for α_m (Figure 4b). Above 24 km, α is close to 2×10^{-8} and values are much larger below 18 km. The small values of α between 18 and 24 km approach the zero-dilution limit of *Plumb*'s [1996] "tropical pipe" model.

The reciprocal of α (τ) is the timescale for dilution by midlatitude air, and has a maximum at about 22 km (Figures 4c and 8c). Our maximum value of τ (80 months) is, however, extremely large compared to the values reported by M96 (15–18 months between 46 and 22 hPa), *Minschwaner et al.* [1996] (maximum 12–16 months), *Volk et al.* [1996] (13.5 months between 16 and 21 km), *Schoeberl et al.* [1997] (18 months between 20 and 28 km), *Hall and Waugh* [1997b] (16 months betwen 100 and 32 hPa), and *Randel et al.* [1998] (15 months between 68 and 32 hPa). While some previous studies [*Hitchman et al.,* 1994; *Minschwaner et al.,* 1996; M96; *Remsberg and Bhatt,* 1996] have suggested that there is an altitude range where air is more isolated from middle latitudes, our results indicate a much greater degree of isolation than other studies.

To understand why this is so, we mention a few relevant observations. First, the profiles of α and τ give different impressions; α contrasts the rapidity of dilution below about 18 km with the slowness above 18 km, while τ emphasizes the very long timescale over the 20–24 km altitude range. Coarser resolution of τ would diminish the maximum; in fact, over the 20–28 km altitude range of

Schoeberl et al. [1997], the reciprocal of our average α is 33 months. Second, since most other estimates of the dilution rate have neglected the role of diffusion, their estimates of τ should be viewed as lower limits, and indeed, when we neglect diffusion (Figure 4c), our profile of τ resembles that of other studies. Third, other studies have generally assumed a vertical velocity profile that is constant with altitude, but when the vertical velocity profile has a minimum (as does w_{tr}), the timescale for dilution necessarily increases at the altitude of the velocity minimum.

Fourth, other studies have not shown the cliff, that is, the sharp reduction in the vertical gradient of $[CH_4]$ and $\hat{\mathrm{H}}$ from HALOE and of $[H_2O]$ from MLS between about 20 and 23 km (Figures 2 and 5). We raised the possibility earlier that the cliff is spurious. However, it appears consistently not just in the quantities mentioned but also in HALOE HF [e.g., *Cordero et al.,* 1997] and CLAES CH_4 and N_2O [*Roche et al.,* 1996] data. The cliff occurs just at the flight ceiling of the ER-2 research aircraft (about 21 km).

This cliff, together with the minimum in vertical velocity, gives rise to much smaller values of α, hence larger values of τ. The actual values of τ are so much longer than the timescale for vertical advection as to be practically infinite (i.e., the transport barrier between tropics and middle latitudes is almost perfect) at these levels.

5.5. Final remarks

Using the WKBJ approach (equation (7)) and considering only the attenuation rate of the tape recorder signal in $\hat{\mathrm{H}}$, we derived profiles of the maximum dilution rate $\alpha_{\max}$and of the maximum vertical diffusion $K_{\max}$for the extreme scenarios where (respectively) $K = 0$ and $\alpha = 0$. We can use $\alpha_{\max}$ and $K_{\max}$, along with our derived profiles of K and α, to estimate the fractional attenuation due to dilution, $\alpha/\alpha_{\max}$, and the fractional attenuation owing to vertical diffusion, $K/K_{\max}$. In Figure 10 we show these quantities for $f = M$ and $f = M^*$. For virtually all the curve fits, the results resemble those for $f = M$: dilution dominates below 18 km, diffusion dominates between 18 and 28 km, and dilution again dominates above 28 km. For our "best" curve fit $f = M^*$, however, diffusion dominates over nearly the entire stratosphere.

Our analysis therefore shows that despite the fairly small values of K (as low as 0.02 m^2s^{-1}) in the altitude range 18–28 km, vertical diffusion plays a significant role in attenuating the tape signal. From Figure 6, even modest values of K can make a significant difference in the net attenuation at 10 hPa (compare curves (i) and (iii)). While these values of K will have little impact on the budget of a monotonically varying constituent like [CH_4], it is clear that one cannot neglect diffusion when considering a constituent whose second derivative is large, like water vapor or perhaps (at some altitudes) ozone. Improvements to 1-D models like those of *Avallone and Prather* [1996] or *Patra and Lal* [1997] would include all of the processes represented in equation (1).

The results presented here suggest that the tropical air column can be divided into three regions. Below about 18 km, α and w_{tr} are comparatively large, and diffusion is probably large there as well. *Holton et al.* [1995] and *Rosenlof et al.* [1997] distinguished this region from the stratospheric "overworld," and it falls under the direct influence of tropospheric circulations, as was noted in section 5.2. When the air rises to the second region, about 19 km, it finds itself extremely isolated from middle latitudes up to about 23 km, and this isolated air rises very slowly (it takes 6 months to travel those 4 km). In this region, dilution is so weak that the amplitude of $\hat{\mathrm{H}}$ anomalies is approximately constant with altitude (Figure 2), and [CH_4] decreases at a markedly reduced rate (Figure 5); what loss of [CH_4] occurs there can be attributed mostly to photochemistry.

Also, although K too has a minimum there, it appears to be largely responsible for the weak attenuation that does occur (Figure 10). This extreme isolation is consistent with the remarkably long duration of QBO westerlies below 23 km. When the air reaches the third region, above 23 km, α, K, and w_{tr} are again larger, and the air again feels the influence of middle latitudes. In contrast to the region below 18 km, where the swirl of tropospheric circulations is responsible for dilution, the agents of dilution above 23 km are probably Rossby waves propagating from middle latitudes [*O'Sullivan and Dunkerton,* 1997].

Acknowledgments. We thank the HALOE science team for the high-quality HALOE data set, Jonathan Kinnersley for the chemical reaction rates, Bill Randel for the constituent climatology, and Tim Hall and Don Delisi for fruitful conversations and helpful comments on the manuscript. This work was supported by National Aeronautics and Space Administration contracts NAS1-96071 and NAS5-32862.

References

Avallone, L.M., and M.J. Prather, Photochemical evolution of ozone in the lower tropical stratosphere, *J. Geophys. Res.,* 101, 1457–1461, 1996.

Bowman, K.P., and Y. Hu, Tropical mixing barriers in the lower stratosphere in the Geophysical Fluid Dynamics Laboratory SKYHI model, *J. Geophys. Res.,* 102, 21,465–21,478, 1997.

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.

Cordero, E.C., S.R. Kawa, and M.R. Schoeberl, An analysis of tropical transport: influence of the quasi-biennial oscillation, *J. Geophys. Res.,* 102, 16,453–16,461, 1997.

Dunkerton, T.J., Evidence of meridional motion in the summer lower stratosphere adjacent to monsoon regions, *J. Geophys. Res.,* 100, 16,675-16,688, 1995.

Dunkerton, T.J., The role of gravity waves in the quasi-biennial oscillation, *J. Geophys. Res.,* 102, 26,053–26,076, 1997.

Eluszkiewicz, J., D. Crisp, R. G. Grainger, A. Lambert, A. E. Roche, J. B. Kumer, and J. L. Mergenthaler, Sensitivity of the residual circulation diagnosed from the UARS data to the uncertainties in the input fields and to the inclusion of aerosols, *J. Atmos. Sci.,* 54, 1739–1757, 1997.

Hall, T.M., and D.W. Waugh, Timescales for the stratospheric circulation derived from tracers, *J. Geophys. Res.,* 102, 8991–9001, 1997a.

Hall, T.M., and D.W. Waugh, Tracer transport in the tropical stratosphere due to vertical diffusion and horizontal mixing, *Geophys. Res. Lett.,* 24, 1383–1386, 1997b.

Harries, J.E., J.M. Russell III, A.F. Tuck, L.L. Gordley, P. Purcell, K. Stone, R.M. Bevilacqua, M. Gunson, G. Nedoluha, and W.A. Traub, Validation of measurements of water vapor from the Halogen Occultation Experiment, *J.*

Geophys. Res., 10,205–10,216.

Hitchman, M.H., M. McKay, and C.R. Trepte, A climatology of stratospheric aerosol, *J. Geophys. Res.,* 20,689–20,700, 1994.

Holton, J.R., P.H. Haynes, M.E. McIntyre, A.R. Douglass, R.B. Rood, and L. Pfister, Stratosphere-troposphere exchange, *Rev. Geophys.,* 33, 405–439, 1995.

Kinnersley, J.S., A realistic three-component planetary wave model, with a wave-breaking parametrisation, *Quart.J.Royal Meteorol. Soc.,* 121, 853–881, 1995.

Mastenbrook, H.J., and S.J. Oltmans, Stratospheric water vapor variability for Washington DC/Boulder, CO: 1964–1982, *J. Atmos. Sci.,* 40, 2157–2165, 1983.

Minschwaner, K., A.E. Dessler, J.W. Elkins, C.M. Volk, D.W. Fahey, M. Loewenstein, J.R. Podolske, A.E. Roche, and K.R. Chan, The bulk properties of isentropic mixing into the tropics in the lower stratosphere, *J. Geophys. Res.,* 101, 9433–9439, 1996.

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

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.,* 101, 3989–4006, 1996.

O'Sullivan, D.J., and T.J. Dunkerton, The influence of the quasi-biennial oscillation on global constituent distributions, *J. Geophys. Res.,* 102, 21,731–21,743, 1997.

Park, J.H., J.M. Russell III, L.L. Gordley, S.R. Drayson, D.C. Benner, J.M. McInerney, M.R. Gunson, G.C. Toon, B. Sen, J.-F. Blavier, C.R. Webster, E.C. Zipf, P. Erdman, U. Schmidt, and C. Schiller, Validation of Halogen Occultation Experiment CH_4 measurements from the UARS, *J. Geophys. Res.,* 101, 10,183–10,204.

Patra, P.K., and S. Lal, Variability of eddy diffusivity in the stratosphere deduced from vertical distributions of N2O and CFC12, *J. Atmos. Terr. Phys.,* 1149–1157, 1997.

Plumb, R.A., A "tropical pipe" model of stratospheric transport, *J. Geophys. Res.,* 101, 3957–3972, 1996.

Randel, W.J., F. Wu, J.M. Russell III, A. Roche, and J.W. Waters, Seasonal cycles and QBO variations in stratospheric CH_4 and H_2O observed in UARS HALOE data. *J. Atmos. Sci.,* 55, 163–185, 1998.

Remsberg, E.E., Diffusion in the lower stratosphere as determined from lidar measurements of volcanic aerosol dispersion, *J. Atmos. Sci.,* 37, 2105–2112, 1980.

Remsberg, E.E., and P.P. Bhatt, Zonal variance of nitric acid vapor as an indicator of meridional mixing in the subtropical lower stratosphere *J. Geophys. Res.,* 101, 29,523–29,530, 1996.

Roche, A.E., J.B. Kumer, R.W. Nightingale, J.L. Mergenthaler, G.A. Ely, P.L. Bailey, S.T. Massie, J.C. Gille, D.P. Edwards, M.R. Gunson, M.C. Abrams, G.C. Toon, C.R. Webster, W.A. Traub, K.W. Kucks, D.G. Johnson, D.G. Murcray, F.H. Murcray, A. Goldman, and E.C. Zipf, Validation of CH_4 and N_2O measurements by the cryogenic limb array etalon spectrometer instrument on the Upper Atmosphere Research Satellite, *J. Geophys. Res.,* 101, 9679–9710.

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

Rosenlof, K.H., A.F. Tuck, K.K. Kelly, J.M. Russell III, and M.P. McCormick, Hemispheric asymmetries in water vapor and inferences about transport in the lower stratosphere, *J. Geophys. Res.,* 13,213–13,234, 1997.

Schoeberl, M.R., A.E. Roche, J.M. Russell III, D. Ortland, P.B. Hays, and J.W. Waters, An estimation of the dynamical isolation of the tropical lower stratosphere using UARS wind and trace gas observations of the quasi-biennial oscillation, *Geophys. Res. Lett.,* 24, 53–56, 1997.

Sparling, L.C., J.A. Kettleborough, P.H. Haynes, M.E. McIntyre, J.E. Rosenfield, M.R.

Schoeberl, and P.A. Newman, Diabatic cross-isentropic dispersion in the lower stratosphere, *J. Geophys. Res.,* 102, 25,817–25,829, 1997.

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

Thuburn, J., and M.E. McIntyre, Numerical advection schemes, cross-isentropic random walks, and correlations between chemical species, *J. Geophys. Res.,* 102, 6775–6798, 1997.

Volk, C.M., J.W. Elkins, D.W. Fahey, R.J. Salawitch, G.S. Dutton, J.M. Gilligan, M.H. Proffitt, M. Loewenstein, J.R. Podolske, K. Minschwaner, J.J. Margitan, and K.R. Chan, Quantifying transport between the tropical and mid-latitude lower stratosphere. *Science*, 272, 1763–1768, 1996.

Wang, R., K. Fraedrich, and S. Pawson, Phase-space characteristics of the tropical stratospheric quasi-biennial oscillation, *J. Atmos. Sci.,* 52, 4482–4500, 1995.

Wang, P.-H., P. Minnis, M.P. McCormick, G.S. Kent, and K.M. Skeens, A 6-year climatology of cloud occurrence frequency from Stratospheric Aerosol and Gas Experiment II observations (1985–1990), *J. Geophys. Res.,* , 29,407–29,429, 1996.

Waugh, D.W., Seasonal variation of isentropic transport out of the tropical stratosphere, *J. Geophys. Res.,* 101, 4007–4023, 1996.

Weare, B.C., and J.S. Nasstrom, Examples of extended empirical orthogonal function analysis, *Mon. Wea. Rev.,* 110, 481–485, 1982.

Weinstock, E.M., E.J. Hinsta, A.E. Dessler, and J.G. Anderson, Measurements of water vapor in the tropical lower stratosphere during the CEPEX campaign: Results and interpretation, *Geophys. Res. Lett.,* 3231–3234, 1995.

Zhang, C., On the annual cycle in highest, coldest clouds in the tropics, *J. Climate,* 1987–1990, 1993.

T.J. Dunkerton and P.W. Mote, Northwest Research Associates, PO Box 3027, Bellevue WA 98009; www.nwra.com.

P.H. Haynes and M.E. McIntyre, Department of Applied Mathematics and Theoretical Physics, Cambridge University, CB3 9EW, England. (http://www.atmos-dynamics.damtp.cam.ac.uk/)

E.A. Ray, NOAA/CMDL Mail Stop R/E/CG1, 325 Broadway, Boulder, CO 80303

J.M. Russell III, Dept of Physics, Hampton University, Hampton VA

Received ______________________

[1]Northwest Research Associates, Bellevue, Washington.

[2]Centre for Atmospheric Science at the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

[3]NOAA Climate Modeling and Diagnostics Laboratory, Boulder, Colorado

[4]Department of Physics, Hampton University, Virginia

Figure 1. Profiles of time mean ascent rate $w_{\rm tr}$ (plain curve) of the tape recorder signal, and the vertical velocity $\bar{w}^*$ calculated by Rosenlof (curve with asterisks), with $\pm 1\sigma$ bars (diamonds for $w_{\rm tr}$, plus symbols for $\bar{w}^*$).

Figure 2. Amplitude of the tape recorder signal defined various ways (curves a–e); see text for details.

Figure 3. Profiles of variously smoothed versions of curves a and d of Figure 2, the smoothing performed as described in the text. Curve M is the same as the dash-dotted version of curve a; see text for definition of M^*. Curves have been shifted left for clarity.

Figure 4. (a) Profiles of $\alpha_{\rm m}$ (solid) and $\alpha_{\rm max}$ (dashed) for curves M and M^* (the latter marked by plus symbols), (b) As in Figure 4a but for $K_{\rm m}$ and $K_{\rm max}$, (c) As in Figure 4a but for $\tau_{\rm m} = 1/\alpha_{\rm m}$ (solid) and $\tau_{\rm min} = 1/\alpha_{\rm max}$ (dashed).

Plate 1. (top) HALOE tropical $\hat{\rm H}$ and (bottom) its reconstruction from the first two EEOFs.

Figure 5. Time mean [CH_4] in the tropics as observed by HALOE (solid curve) and as simulated by 1-D model for two values of the diffusion coefficient K: $K = 0$ (dashed curve) and $K = K_{\max}$ (dash-dot curve).

Figure 6. Amplitude of $\hat{\mathrm{H}}$ anomalies simulated by 1-D model (broken curves). The solid curve is M^* from Figure 3, representing observations.

Figure 7. End-to-end consistency check by forward solving and then inverse solving equation (1). Dashed curves are the profiles used for forward solution and are, respectively, the same as (for w in Figure 7a) the plain solid curve in Figure 1, (for α in Figure 7b) the solid curve in Figure 4a, and (for K in Figure 7c) the solid curve with plus symbols in Figure 4b, modified as noted in Table 2. Solid curves in this figure are profiles derived from inverse solving. Differences give an idea of the effects of numerical truncation error.

Figure 8. Profiles of (a) w, (b) α, (c) $\tau = 1/\alpha$, and (d) K calculated by inverse solving the 1-D model with HALOE data, using equations (4). In each panel the mean of six solutions using the six broken curves in Figure 3 is shown as the solid curve, bounded by $\pm 1\sigma$. Also plotted are the results when $f = M^*$ (diamonds). The calculation is performed on a 50-m grid, but the results are plotted on HALOE levels. Long-dashed curves in Figures 8a, 8b, and 8c represent w_{tr}, α_{m}, and $1/\alpha_{\mathrm{m}}$, respectively. The solution for K (with $f = M^*$) has been extended to 100 hPa as is discussed in section 5.2.

Figure 9. Plots of $w = \tilde{w} - (K/H) + K_z$, $\tilde{w}$, $\tilde{w} - (K/H)$, w_{tr} as predicted by approximate formula (B4), and w_{tr} as originally estimated from Plate 1b. Solutions correspond to $f = M^*$, the case shown by diamonds in Figure 8. See text for details.

Figure 10. Fraction of attenuation of the observed $\hat{\mathrm{H}}$ signal due to diffusion (dashed curve) and dilution (solid curve) for $f = M$ and $f = M^*$ ('+' symbols). Curves do not quite add to 1 because results are shown as the ratio of the full solution to the maximum allowed with the WKBJ approximation.

Plate 2. $\hat{\mathrm{H}}$ anomalies simulated by 1-D model for the five combinations of K and α given in Table 2.

Table 1. Definitions of various vertical velocities.

term	definition
w	true vertical velocity[1]
$\bar{w}^*$	vertical velocity calculated by Rosenlof
w_{tr}	ascent rate of marks on the tape
$\tilde{w}$	$\equiv w + (K/H) - K_z$
w_{synth}	like w_{tr} but for synthetic $\hat{\mathrm{H}}$ data

[1]used in 1-D model

Table 2. Profiles of K and α used to generate synthetic data. See text for definitions, and Figure 5 for profiles, of K_{max}, K_{m}, α_{m}, and α_{max}.

	(i)	(ii)	(iii)	(iv)	(v)
K	0	0	K_{m}[1]	K_{max}	K_{max}
α	α_{m}	α_{max}	α_{m}	0	α_{m}

[1]The profile of K_{m} shown in Figure 4b was modified below about 17.5 km to increase monotonically down to a value of 0.1 $\mathrm{m^2 s^{-1}}$ at 100 hPa.