Convective Mixing near the Tropical Tropopause: Insights from Seasonal Variations

STEVEN C. SHERWOOD
Department of Geology and Geophysics, Yale University, New Haven, Connecticut

ANDREW E. DESSLER
Earth Systems Science Interdisciplinary Center, University of Maryland at College Park, College Park, Maryland

(Manuscript received 7 November 2002, in final form 26 April 2003)

ABSTRACT
It has been suggested that convection remains important in the budget of water vapor up to the tropical tropopause or even higher. But observed seasonal variations of CO$_2$ and water vapor in the lower stratosphere, and their timing, call the required convective transport into question. Here, these seasonal variations are modeled using several idealized representations of convection. First, a CO$_2$-like tracer is added to a previously published climate model in which convection explicitly transports air to a variety of heights ranging from 14 to 19 km, in a manner sufficient to dehydrate the stratosphere. It is found that these motions are consistent with the observed lags and amplitudes of seasonal variations near and above the tropopause, including a 2-month phase lag in CO$_2$ at 390 K relative to surface values and a similar lag in H$_2$O relative to tropopause temperatures. This result is explained in terms of the model’s mixing physics.

Next the ability of other models is considered, where convective outflows are confined below some ceiling at or below the cold-point tropopause, to account for the observed seasonal cycles. Behavior of such models is governed by the placement of the ceiling relative to a known stagnation surface in the radiatively balanced vertical velocity. It is found that convection must reach to within 1 km of the cold point in order for realistic seasonal cycles to exist above the tropopause in these simulations. Importantly, the properties of air entering the stratosphere must be determined by those of the planetary boundary layer rather than the upper troposphere. This work reinforces the view that convective mixing must evanesce gradually in importance through a tropical tropopause layer of substantial thickness, rather than stopping at any particular height.

1. Introduction
Since the work of Brewer (1949) more than 50 years ago, researchers have worked to understand the transfer of mass and water vapor into the stratosphere. Difficulty in explaining the dryness of the stratosphere led Johnston and Solomon (1979) and Danielsen (1982) to suggest that convective penetrations of the stratosphere were providing dry air. This provoked a debate, as yet unresolved, about the height to which tropical convection reaches and influences constituents. Stratospheric dryness can be explained without resort to direct convective effects, but the feasibility of this depends on unanswered questions about ice nucleation at cold temperatures and vertical motions near the tropopause (Holtton and Gettelman 2001; Jensen et al. 2001).

Recently, Sherwood and Dessler (2000, 2001, hereafter SD00 and SD01) proposed an explicit model of the interaction between overshooting convection and large-scale processes. They emphasized the importance of a tropical tropopause layer (TTL) between about 14 and 19 km, rather than a specific tropopause or other surface, as the true mediator of exchange between the troposphere and stratosphere. Air was carried up into the TTL by overshooting convective updrafts, was dehydrated immediately, and subsequently lofted out the top by radiative heating with no further dehydration. SD01 found that the observed water vapor and ozone profiles could be created by the effects of physically reasonable overshooting. In their model, some convective elements penetrated the cold-point tropopause though most did not.

Different observations have been cited as evidence for and against such convective penetrations. Those against include the paucity of radiometrically estimated cloud tops above 14 km (e.g., Highwood and Hoskins 1998), and the rapid change in vertical ozone gradient near 14 km—from a small value below, to a value roughly consistent with photochemical production above (Folkins et al. 1999). However, SD01 showed that neither of these observations rules out the required penetrations. Observations in favor of convective effects in the TTL include evidence of its effects on CO (Dessler
apparent control of stratospheric water by deep convective ice mechanisms (Sherwood 2002), and near-
constant isotopic depletion of water vapor throughout the TTL (Kuang et al. 2003). In a companion paper,
Dessler and Sherwood (2003, manuscript submitted to Atmos. Chem. Phys.) test the ability of the SD01 model
to account for the latter, affirming that it can, given a suitable assumption about the composition of lofted ice.

One other observed characteristic potentially at odds with significant TTL convection is the seasonal cycles
crossing into the TTL. Most of the above analyses were conducted as steady solutions to idealized equations. However,
the observed TTL experiences marked seasonal variations in temperature, water vapor, and CO₂, which
through the tropopause region with only modest attenuation (e.g., Andrews et al. 1999; Mote et al. 1996;
Weinstock et al. 1995). One important question is whether this propagation is compatible with the degree
of overshooting required, for example, to perform dehydration. Too much convection should produce sea-
sonal variations that advect faster than the background ascent rate, and damp rapidly, as air of a wide variety
of ages is mixed into the same potential temperature surface. Here we address this question by extending the
work in SD01 to include the prediction of a CO₂-like tracer on seasonal timescales. We also consider simple
alternative models of convection, in order to try to draw more general conclusions about what type of convective
transport is consistent with observed seasonal cycles.

2. Observed seasonal cycles

Seasonal variations of CO₂ occur in the lower troposphere due mainly to seasonal variations in plant res-
piration and decay. The seasonal variation is large in the Northern Hemisphere, vanishes near the southern
equator, and becomes small and of the opposite sign in the Southern Hemisphere extratropics. At 390 K, slightly above the tropical tropopause, the sea-
sonal variation remains similar to that at the surface near the equator, with little damping, but is lagged by 2 months (Boering et al. 1994). This curious result de-
mands explanation. The calculated damping depends on the amplitude of the near-surface cycle, which is im-
perfectly known because the observed cycle varies greatly with latitude and the latitudinal distribution of
air that reaches 390 K as a function of season is not well known.

Seasonal variations of water vapor, presumably related to seasonal variations in temperatures near the tro-
popause, also propagate into the stratosphere. The amplitude has been measured in the lower stratosphere by
aircraft and satellites (Mote et al. 1996; Weinstock et al. 1995); we report results taken from Randel et al.
(2001). As with CO₂, the main uncertainty in calculating damping from observations lies in assessing the und-
damped amplitude. This problem is more serious for H₂O than for CO₂ since the stratospheric entry value of
water vapor is determined by poorly understood sinks rather than conservative transport from the boundary
layer. This causes larger uncertainty in the observed damping of the H₂O cycle. Seasonal cycles are quanti-
fied further in section 3c.

3. Test of overshooting transport

We first test the proposal that overshooting cumulus clouds, sufficient to perform dehydration, would pro-
duce too much vertical mixing to be consistent with the observed seasonal cycles. We utilize the model of SD01.
The usefulness of this model is that it has already proven able to explain vertical profiles of key constituents; we
employ it here with minimal modifications, so that we can test the ability to explain the previous effects as
well as seasonal variations simultaneously.

a. The SD01 model

Here we briefly summarize aspects of the SD01 model most important to vertical mixing. That “1½-dimension-
al” model consisted of two one-dimensional columns or regions, representing 1) a region of convective over-
shooting and 2) a “capped” region with no explicit convection. Temperatures of each region were specified
from observations. The two regions were coupled by a large-scale diabatic circulation prescribed from obser-
vations, and by parameterized isentropic mixing.

The top and base of the TTL were defined by deep convection. The base (at 14–15 km) was near the typical
level of neutral buoyancy of undilute air parcels and was near or slightly below a stagnation surface (see sec-
tion 4) in the vertical velocity profile away from con-

vction. Thus entry into the TTL and crossing of the stagnation surface was entirely within convective up-
drafts. The maximum height of convective overshoots was determined by the environmental conductive avail-
able potential energy (CAPE), which was allowed to vary in a Gaussian distribution about a mean value of
3000 J kg⁻¹ in the standard model configuration. This meant that many convective elements penetrated the tro-

popause. In the SD01 model, the final potential tem-
perature of detraining air masses was governed mainly
by mixing and depended little on the variations of
CAPE.

The effects of overshooting convection in the model
were computed according to a buoyancy sorting scheme,
with a prescribed rate of overshooting into the TTL.
Mixing below the peak overshoot height was either
evenly distributed in pressure (our default case) or de-
creasing linearly to zero at the lowest pressure (an alter-
native we refer to here as “weak overshooting,” a
relative term since this still represents substantial excursions into the lower stratosphere). Mixtures were
computed according to the scheme of Emanuel (1991),
and each mixture sank to a new level of neutral buoy-
cy where it detrained into the environment. The re-
Fig. 1. Detrainment profile of the convection scheme in the SD01 model; units (day$^{-1}$) correspond to the ratio of mass flux detrained into a given layer to the total mass contained in that layer (in both model regions). Solid line denotes standard case; dashed line denotes the weaker overshooting model setting.

resulting detrainment profiles were concentrated mainly in the lower TTL (Fig. 1). Importantly, the resulting detrainment always occurs below the mixing (usually well below), and the environment is affected only by detrainment and not by mixing (except that some air is removed from the mixing level). Detrainment decreases strongly with height due to the limited buoyancy of mixtures and rapidly increasing environmental $u$, but detrainment is also proportional to static stability, which produces some vertical structure and shifts the detrainment profile upward.

The SD01 model contained a number of tunable parameters. The amount of overshooting convection was set as required by TTL energy balance; the amount of ice remaining to re-evaporate in overshoot mixtures (governed by the important parameter $c$), and the horizontal mixing rate, were tuned to produce the correct moistures above and below the TTL, respectively. The model was able to predict the approximate heights of the minimum in water vapor, rapid fall-off of cloud water, and first-order discontinuity of ozone—as well as the horizontal trends in water vapor and ozone—found in TTL observations. These predictions were robust to changes in tuned parameters (although the stratospheric moisture itself was quite sensitive to $c$ and somewhat sensitive to other parameters). The model was run only in DJF conditions in that study.

b. Model runs

Here we run the model in a transient mode with seasonal variations. In the TTL, observed June–August (JJA) temperatures are several degrees warmer than in DJF in both regions; we employ JJA sounding data from the same locations as before (Indonesia for the overshoot region and the Atlantic area for the capped region) to establish the JJA temperatures. The transient runs oscillate sinusoidally (at 1 cpy) between the DJF and JJA extrema.

1) STANDARD AND WEAK-OVERSHTTING CASES

Observations do not indicate any substantial seasonal variation in the statistics of CAPE through the Tropics, so we do not consider such variations. Sensitivity studies showed that modest seasonal variations in CAPE would have little effect on our results anyway.

For our standard configuration, we also assume no seasonal cycle in vertical velocity, horizontal mixing, or frequency of convection, retaining the SD01 values year-round as with CAPE. To examine the sensitivity to the degree of overshooting, we also compute results with the model in weak-overshooting mode throughout the year.

2) VARIABLE DYNAMICS CASE

Observations do indicate that tropical mean diabatic uplift in the lower stratosphere (Rosenlof 1995) and horizontal variations in mass flux near the tropopause (Sherwood 2000) are each only about half as strong in JJA as in DJF. We consider the possible importance of these circulation changes via an alternative run in which a sinusoidal seasonal variation is imposed on each of these quantities making them half as great in JJA as in DJF. Further sensitivity tests indicate that it is the tropical mean ascent rate that matters most, so we may think of this simulation as being characterized by seasonally varying tropical upwelling.

3) DIFFUSIVE ICE RE-EVAPORATION CASE

The parameter $c$, the ratio of retained ice to detrained vapor, was particularly important owing to the great influence of re-evaporated ice on stratospheric humidity found by SD01. In our standard configuration, $c$ is constant throughout the year at the value ($c = 4$) obtained empirically by SD01. As yet we have no theory to justify this value of $c$. However, the diffusion equation dictates that, all other things (ice surface area per unit mass, turbulence, etc.) being equal, the ice evaporation rate per unit mass will be proportional to the difference between the saturation mixing ratio and the ambient water vapor mixing ratio. In a saturated environment, for example, no reevaporation of ice can occur. Based on this argument, we consider for comparison with the standard case a simulation in which

$$c(t) = D_s[q_s(t) - q(t)],$$

$^1$ The energy closure used to determine the amount of convection demands that overshooting convective amount be approximately proportional to vertical heat advection, so these quantities are varied together.
with $q$ and $q_s$ evaluated at 100 hPa and $D_y$ an effective diffusion constant. This constant is chosen such that in DJF, $c$ has the same value as in the other configurations.

4) MINOR NUMERICAL ISSUES

Finally, several numerical issues in the SD01 model required attention. First, that model was originally run to equilibrium by alternating several times between sequences with a long time step where the TTL and stratosphere were updated, and a short time step where the whole atmosphere was updated (required by the higher vertical velocities in the troposphere). We retained this procedure here, with 10-day sequence alternation. Though this produced poor transient behavior below the TTL, that behavior had essentially no impact on the TTL itself. The physical reason for this is explained in section 4; we confirmed it quantitatively by performing a test run with no transient behavior in the troposphere, which changed the TTL seasonal variability by less than 1%. Second, we computed the net tendency due to convection and advection prior to updating concentrations, rather than applying these separately as had been done in SD01; this reduced artificial diffusion, improving the transient runs, but had very little impact on the time-invariant solutions reported in SD01. Remaining numerical diffusion appears to produce damping effects comparable to those in the real atmosphere. Finally, an error was discovered in the isentropic mixing scheme of the original model and was corrected. This caused small differences (<10%) in some constituent concentrations in the TTL.

c. Results

We characterize the tracer seasonal cycles in the lower stratosphere by examining their phase and normalized amplitude at 390 and 420 K, with normalization being relative to the (tropospheric) source. Predicted and observed values are plotted in Fig. 2. Normalizing and phasing the H$_2$O cycle is not straightforward; see the appendix for a description of how this was done.

1) CO$_2$

Carbon dioxide is easier to interpret since it is conserved. The results show that the seasonal cycles predicted by the SD01 model are not far away from those...
indicated by observations. They are slightly too weak at 390 K, and slightly early at 420 K, but agree fairly well overall given the crude character of the model. The suggestion that overshooting convection must severely dampen or eliminate seasonal variations in the lower stratosphere can clearly be rejected. Importantly, the simulations approximately reproduce the 2-month lag in CO₂ at 390 K. Changing either to weak overshooting or seasonal dynamics produces slightly greater lags, which is expected in both cases due to the lower-altitude starting point of tropospheric air, and the slower annual mean ascent rate, respectively.

2) H₂O

In interpreting water vapor results, we must remember that they reflect not only the damping effects important for CO₂ but also the seasonal variation of local sources and sinks. The importance of the latter is shown by the different results here compared with CO₂. In particular, most of the runs show much too little seasonal variation, only about 0.6 ppmv for the standard run. The small cycles in these runs are due mostly to the fact that overshoot temperatures respond only partially to changes in TTL temperature, therefore varying too little with the seasons to explain the observed H₂O cycle.

The “diffusive ice” simulation (in which the diffusion coefficient, rather than ice reevaporation parameter, was taken to be constant over the year), on the other hand, shows a strong cycle, on the high side compared to observations. There is more ice evaporation in JJA due to the lower relative humidities and higher temperatures. There are reasons to believe this assumption is plausible, if unproven. Sherwood (2002) observed that relative humidities in the TTL were controlled by variations in the surface area to volume ratio of lofted ice; tight tracking in time between satellite observations of the two quantities implied that ambient water vapor mixing ratios equilibrated rapidly (within a month or less) with changing ice distribution and temperature. This would not be the case unless the vapor source/sink were quite sensitive to ambient relative humidity. Though this does not confirm the assumption made in our diffusive-ice model, it does confirm that lofted ice is of primary importance to the moisture budget, lending credence to the assumption.

A Hovmöller diagram showing the “tape recorder” simulated by the diffusive-ice version of the model appears in Fig. 3. This simulation resembles the seasonal variation shown in Mote et al. (1996), and illustrates that overshooting convection falls far short of homogenizing the TTL. Damping of the cycle due to vertical diffusion is also apparent through the stratosphere, although in the real atmosphere most of this damping is probably due to horizontal mixing with the extratropical stratosphere.

4. The stagnation surface (and how to cross it)

An interesting feature of the simulations noted briefly above was the complete insensitivity of conditions in the TTL to those below the TTL. This runs counter to the common picture, in which net upward motion in the ascending branch of the Brewer–Dobson circulation extends down into the tropical troposphere, in turn implying that tropospheric properties should be advected (however slowly) into the lower stratosphere. Here, we offer a stark alternative to this picture.

The tropical troposphere is characterized by net clear-sky cooling (Qₐ < 0) and sinking, since longwave cooling from CO₂ and H₂O exceeds solar heating; by contrast, lifting occurs in the stratosphere where solar absorption by ozone, decreasing influence of H₂O, and (near the tropopause) decreased emission due to cold temperatures, tip the balance to net heating (Qₐ > 0). Mathematically, a zero must exist in the time-mean Qₐ at some level in between, where air does not move vertically in theta coordinates except by nonradiative heat sources or by transient fluctuations about the mean radiative heating. We term the points defining this level as the stagnation surface. The stagnation surface is too high for latent heating effects to be significant as a source of heat, which leaves mechanically forced mixing and radiative heating fluctuations as the only ways for air to cross this surface. The potential importance of this surface for transport has been emphasized (Gettelman and de Forster, 2002; Sherwood and Dessler, 2001) though not thoroughly investigated.

Radiative transfer calculations are necessary to locate the surface. Sherwood (2000) and SD01 reported a height of 135 hPa (14.8 km) in DJF, using mean soundings of temperature, ozone, and water vapor. However, Gettelman and de Forster (2002) reported a mean height
of 16 km using a different radiative transfer model with different observations. Neither study found any noteworthy seasonal or horizontal variation in the height of the surface. The stagnation level potential temperatures depend on horizontal position and season, ranging from about 352 to 360 K at 135 hPa and from 361 to 372 K at 16 km, with lowest values occurring near the deepest convection. JJA values are 2–4 K higher than DJF values.

In the SD01 model, the only means for transport across the stagnation surface was overshooting convection. Further, only updrafts air parcels from near the surface can possess sufficient \( \theta \) to reach the required levels; air entrained along the way, from within the troposphere, would retard the updrafts and reduce their buoyancy too much. Thus properties of the TTL (specifically, the air above the stagnation surface) must be determined entirely by those of the lower troposphere! This explains the near total lack of influence of the free troposphere upon the TTL state in the SD01 model. Whether this is true in reality depends on how transport across the stagnation surface is partitioned between convection and diffusion. Isentropic movement of an air parcel from one horizontal location to another may cause fluctuations in \( Q_{\eta}(\theta) \) (e.g., air just below the stagnation surface can move to where the surface is lower; ascend to higher theta, then return to its original horizontal position located above the stagnation surface). Temporal fluctuations in \( Q_{\eta} \) at a given point are also likely. Each of these will lead to random, diffusive transport across the surface, which could dominate convective transport and allow upper-tropospheric properties (such as humidity) to affect the TTL.

5. Tests of diffusive transport above capped convection

We investigate these matters by considering transport models that do not include overshooting convection. In particular, we consider the simple alternative picture in which convection prevails up to some ceiling or “capping” level \( \theta_{cap} \) where all convective effects cease. Outside of convection, transport is diffusive and advective. Since advection is zero at the stagnation level, however, diffusion is the only available means for getting material across this surface unless it is below the capping level.

Several things must be specified in such a model. These include the treatment of convection, the background flow field (i.e., the mean velocity field), the nature of the diffusive transport within the background flow (i.e., the higher-order statistics of individual parcel velocities) outside of convection, and of course the position of the capping level. We begin with a highly reduced specification, in order to capture the basic problem in what we believe is its purest form, then consider a couple of alternatives. We will examine the results as a function of the position of the capping level relative to that of the stagnation level.

a. Convective treatment

Convection is a mixing process that tries to homogenize passive, conserved tracers. Thus, one simple (though extreme) way of representing convection is to assume that it completely homogenizes such tracers below the cap, while having no effect above the cap—in other words, establishing a zero-age boundary condition at the capping level. This assumes a collection of up- and downdrafts so intense that convective timescales are much shorter than those of sources/sinks or larger-scale advection. Any net mass flux is a small residual difference between the much greater upward and downward draft fluxes. We will idealize convection in this manner to start with, though we will consider an alternative paradigm in section 6.

b. Background flow

We consider a 1D model with potential temperature \( \theta \) as its (vertical) coordinate. The time-averaged heating rate \( \overline{Q}(\theta) \); which is also the vertical velocity, varies linearly with \( \theta \), crossing zero at the stagnation level \( \theta_{stag} \); in other words, the diabatic flow “divergence”

\[
\frac{d\overline{Q}(\theta)}{d\theta}
\]

is constant. This causes mean advection away from \( \theta_{stag} \) in both directions. The divergence must be balanced by a combination of net convective outflow and/or net mass transport from higher latitudes, neither of which is currently well known. At some level \( \theta_i, \theta_i > \theta_{stag} \), we suppose \( \overline{Q}(\theta) \) plateaus at a constant value (following the observed behavior) though we find that this feature is not crucial. The heating profile is shown in Fig. 4.

c. Dispersion of parcels

In the real atmosphere, heating rates are not constant on a theta surface. Instantaneous heating rates \( \dot{Q} \) experienced by individual parcels will differ from \( \overline{Q}(\theta) \), leading to vertical dispersion of parcels. We treat this by tracking a large number of independent parcels through time. At each time \( j \), we add a heating anomaly \( Q'_i \) to each parcel \( i \),

\[
Q_i = \overline{Q}(\theta_i) + Q'_i.
\]

The heating anomalies are taken as independent random red noise processes for each parcel, each process characterized by a standard deviation \( \sigma \) and decorrelation time \( \tau \) (independent of position). The time step \( \delta t \) is one day. Each parcel is traced from a final (observation) level backward via

\[
\theta_i = \theta_{i+1} - Q_i \delta t.
\]

\[\text{We generated red noise by convolving a randomly generated white noise with the appropriate width Gaussian to yield a time series whose autocorrelation decays to zero with time lag } \tau.\]
until the parcel’s vertical position is below $\theta_{\text{cap}}$, yielding the parcel’s “age” (time since leaving convection). The amplitude and phase of the seasonal cycle of a constituent at the final level can be obtained by convolving the distribution of parcel “ages” at that level with the constituent’s tropospheric seasonal cycle (real and imaginary parts), which assumes that all convective parcels have identical constituent concentrations at a given time and that these are conserved after the air leaves convection.

d. Model behavior and parameters

Near the stagnation level, parcels will execute a classical random walk with no preferred direction. They may linger in this “stagnation region” for very long periods of time. However, parcels that stray too far from $\theta_{\text{stag}}$ have no chance of returning, since the mean flow will become stronger than the maximum dispersion velocities; such parcels may be said to have “escaped” into the stratosphere (or troposphere).

If stratospheric escape is already guaranteed before a parcel reaches the heating plateau at $\theta_0$ (Fig. 4), then the plateau height will be irrelevant in regulating escape. Simulations indicate that the observed $\sim 25$ K gap between $\theta_0$ and $\theta_{\text{stag}}$ is more than sufficient for this to be so. Thus, the only flow parameter relevant to parcel escape is the divergence $dQ/d\theta$, approximately $(50 \text{ days})^{-1}$ if we fit the heating profile to that calculated by a radiative transfer model of Rosenfield et al. (1994). The resulting values are approximately $\tau = 4$ days and $\sigma = 0.5$ K day$^{-1}$. Since $\tau \ll (dQ/d\theta)^{-1}$, the random walk will in effect be governed by a single diffusivity $\sigma^2/\tau \sim 1$ K$^2$ day$^{-1}$.

Finally, in this model, absolute vertical position does not enter. Thus, the dependence on convection occurs only through the convective “headroom”

$$\Delta \theta = \theta_{\text{cap}} - \theta_{\text{stag}}$$

between the capping level and the stagnation level.

To summarize, our Lagrangian transport model—which includes the mean transport/heating profile, the dispersion of parcels, and the behavior of convection—is encapsulated by only three parameters: the diabatic divergence $dQ/d\theta$, the diffusivity $\sigma^2/\tau$, and the headroom $\Delta \theta$. The first two were quantified from observations or radiative transfer calculations. The impact of changes in headroom, which is poorly known, is of primary interest.

e. Simulated seasonal cycle characteristics

The simulated seasonal cycle lags and normalized amplitudes are shown in Fig. 5 as a function of the amount of convective headroom $\Delta \theta$. With sufficient headroom for convection to approach the cold-point tropopause ($\Delta \theta \sim 20-25$), lags become too short compared to observations; as the headroom goes to zero, seasonal cycle lags become very long and amplitudes drop off. Reasonable results occur for $\Delta \theta \sim 5-10$, implying homogenization to a level at least 1 km above the stagnation level if it is at 15 km (or at least 500 m above, if it is at 16 km). Figure 6 illustrates why amplitudes become so unrealistically small as $\Delta \theta$ decreases: some parcels get trapped near the stagnation level for very short times, yielding unrealistically small amplitudes.
long periods of time, stretching out the distribution of ages so that no strong seasonal cycle persists. Easing transit across the stagnation surface by increasing the heating diffusivity does not significantly change the situation (Fig. 5).

It may seem obvious that convection must loft air above the stagnation surface in order to supply the mass going into the stratosphere. In principle, though, this mass could come from horizontal convergence, since flow fields are not well known. The results here show that significant convective flow must occur across the stagnation surface, independent of mass balance considerations.

6. Alternative formulations of the capped convection model

a. Different convective representation

We treated convection as a perfect homogenizer of the layer beneath the convective cap. This paradigm probably applies reasonably well to the region commonly called the “mixed layer” (typically within 1 km of the surface), but is definitely an extreme view when applied to the whole troposphere. Note that this assumption produces the smallest possible age dispersion, given $\theta_{\text{cap}}$, since all ages are at that level identically zero. Here we test our conclusions with an alternative representation of convection.

1) DESCRIPTION

The opposite extreme would be to assume that penetrative convection in the free troposphere is composed purely of undilute updrafts. Without downdrafts, mass continuity links the updraft mass flux directly to the environmental flow, which is in turn imposed by radiative balance (e.g., Betts and Ridgway 1988) and/or mechanical forcing such as “downward control” (see discussion in Mote et al. 1996). Updrafts inject air with near-surface properties into their detrainment level. Though clearly inappropriate at lower levels, this “pure updraft” paradigm appears to predict temperatures in the upper few kilometers of the troposphere successfully (Folkins 2002). We contrast the pure updraft paradigm with our original one in Fig. 7. We expect immediately that there will be greater dispersion of ages with pure updraft, since ages on the $\theta_{\text{cap}}$ surface will not all be the same.

To keep things simple, we assume here that detrainment is evenly distributed between $\theta_{\text{stag}}$ and $\theta_{\text{cap}}$, at a value so that the total detrainment supplies the tropical, lower-stratospheric upwelling mass flux. If $\theta_{\text{cap}} = \theta_{\text{stag}}$, the detrainment profile becomes a delta function and this paradigm is no different from the original one (the case of $\theta_{\text{stag}} > \theta_{\text{cap}}$ cannot be formulated with this assumption, but has already been shown to produce prohibitively large dispersion even under the original, most favorable assumption).

Establishing parcel “age” is a bit trickier here than in section 5c. When $\theta_{\text{stag}} \leq \theta_{\text{cap}}$, there is no guarantee but instead some (small) probability $P_i \leq 1$ that parcel $i$ just happens to have detrained from an updraft at the current time step $j$. Since heating rates are independent between parcels, $P_i$ equals the fraction of mass at $\theta_{\text{stag}}$ that detrains per time step (i.e., the quantity displayed in Fig. 1 for the SD01 model). Mass conservation, together with our assumption of vertically uniform mass insertion, yields

$$P = \frac{\bar{Q}_0}{\theta_{\text{cap}} - \theta_{\text{stag}}} \delta t$$

independent of $i$ and $j$. For a given net ascent, this is the smallest possible $P$, since any downdrafts must be balanced by more updrafts and therefore greater $P$. This $P$ may be contrasted with $P = 1$ with homogenizing convection; otherwise, the models are identical.

2) RESULTS

Figure 8 compares results from the pure updraft and homogenization cases. Two changes are observed. First, lags become longer with pure updrafts, as expected since

![Fig. 7. Cartoon illustrating two extreme paradigms for representing “capped” convection (see text).](http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(2003)060<2674:CMNTTT>2.0.CO;2)
Fig. 8. Lag vs amplitude of the passive, CO$_2$-like tracer cycle at 390 K, comparing standard version of random walk model and versions with minimal updraft convection. One updraft version is with the standard value of vertical background flow divergence, the other with this divergence doubled and with $\theta_0$ lowered accordingly to 367.5 K so as to preserve the same net mass flux into the stratosphere. Symbols denote $\Delta \theta = \theta_{up} - \theta_{stag}$, the height above the stagnation surface reached by convection.

Fig. 9. Lag vs amplitude of the passive, CO$_2$-like tracer cycle at 390 K, comparing the Lagrangian and Eulerian representation of capped, homogenizing convection. (nonnegative) parcel ages at $\theta_{up}$ are not all zero as they were before, so the mean age is greater. The lag difference increases with increasing headroom, so that large amounts of headroom do not produce such short lags.

The second change is that amplitudes, which previously increased monotonically with $\Delta \theta$, begin decreasing when $\Delta \theta > 10$. This is because not only mean ages but also dispersion of ages is increased, eventually damping the cycle amplitude. The damping becomes significant when $\Delta \theta > 20$, where the lofting time has become a significant fraction of a seasonal cycle.

Aside from these differences, the failure of the simulations for small values of $\Delta \theta$ is about the same as before. It appears that we must have $\Delta \theta \approx 5-10$ K with either paradigm, corresponding to a clearance by convection of $\approx 500$ m above the stagnation surface as placed by Gettelman and de Forster (2002), or $\approx 1000$ m above the surface if it is only at 15 km. This certainly places the top of convection within 1 km of the cold-point tropopause.

b. Different flow model

To examine the capped-convection model in a different setting, we incorporated the original (homogenizing) convective representation back into the SD01 model. To do this we simply deactivated the overshoot scheme, prescribed a capping level in each model region, and set the passive tracer and water vapor to a seasonally determined tropospheric value everywhere below the capping levels at each time step. This experiment allowed us, first, to test whether the above results were sensitive to the flow field and numerical representation; second, to consider spatially heterogeneous capping height; and third, to simulate seasonal cycles of water vapor easily.

In treating water, values below the capping level were set to saturation. Large-scale supersaturation was not allowed anywhere, and water vapor was removed as necessary to enforce this. This ended up occurring only in the colder model region (as was the case with convective dehydration in SD01), since horizontal transport was strong enough, given the thermal gradient, to prevent saturation in the warmer region. Therefore, dehydration in this model was similar to that in the slow-ascent model described by Holton and Gettelman (2001) though much more crudely resolved.

RESULTS

Passive tracer results were similar to those of the Lagrangian model, and are shown in Fig. 9. Slightly more deterioration of the seasonal cycle occurred, especially for small $\Delta \theta$. But the broad features of the result, and the $\Delta \theta$ that best fits observations, were similar and therefore do not appear to be particularly sensitive to details of the flow field or numerics. Independent variations of capping height in the two model regions (not shown) demonstrate that homogenization in either region alone is able to produce most of the effect observed with homogenization in both regions, which we would expect since the horizontal transport timescale is small compared to that of the seasonal variation.

The simulated, normalized water vapor seasonal cycle amplitude was 1.1, roughly double the observed estimate. This was essentially independent of $\Delta \theta$, not surprising since final dehydration always occurred above the convective capping level. The large amplitudes were associated with similarly too-high mean moisture. So if
some process (e.g., Kelvin waves) were added to further dehydrate the TTL, and if this process removed a similar proportion of the water vapor year-round, the resulting cycle amplitude would be fairly close to observations. Thus, given a suitable process for accomplishing mean dehydration, either the SD01 or gradual-dehydration models also appear capable of reproducing acceptable seasonal water vapor variations (subject to the caveats in section 3c). Further understanding of convective and gradual dehydration mechanisms is an active area of research at this time.

7. Discussion

It may seem surprising that convective overshooting to 70 hPa did not significantly damp the seasonal cycle or reduce the age of air above the tropopause (section 3). The keys to this puzzle are two. First, overshoots disturb the environment only where they detrain (cf. Fig. 1), not where they reach their maximum excursion or entrain environmental air. Detrainment occurs mostly below the tropopause where convective insertion of air is not a hindrance, but in fact, a necessity for a robust seasonal cycle (cf. section 5). Second, detrained material is a mixture of boundary layer and lower stratospheric air, rendering its properties close to those of the environment at the same level, as long as the mixing did not extend too far out of the linear range of vertical constituent variation with respect to $\theta$.

A back-of-the-envelope calculation is helpful here. Water vapor simulated by SD01 decreased above the cold-point tropopause by roughly an additional 10% due to dehydration above the tropopause (observations suggest that this also occurs in nature). Given that detrainment at such levels requires extremely cold overshoots, the detrained air masses have practically no water vapor, so 10% mass dilution is sufficient. This would reduce the seasonal cycle amplitude by 10% if the inserted air had random age. At lower levels within the TTL, there is more detrainment, but the constituent profiles are nearly linear with respect to $\theta$ so the impact on seasonal variations is small.

If undilute air is inserted at a wide range of altitudes (section 6a), the outcome is different. In this case a wide spread of ages must occur within the insertion layer bounded by $\theta_{\text{stg}}$ and $\theta_{\text{cap}}$, if the advective transit time of the layer stretches out to several months, cycles will begin to be seriously affected. This places an upper limit on the possible range of undilute insertion, although we do not find that this limit is highly restrictive and significant insertion all the way to the cold point (though not much higher) seems to be consistent with observed seasonal cycles.

The capped convection models show that convection in some form must get most of the mass up to at least 16 km in order to account for seasonality of lower-stratospheric CO$_2$. Meanwhile, the original and capped versions of SD01 show that H$_2$O seasonality depends less on the transport mechanism and more on the dehydration mechanism. Thus, though Holton and Gettelman (2001) were able to obtain realistic seasonal cycles of water vapor in a model with no convective effects above 14 km, we conclude that such a model would not also produce correct seasonal cycles of conserved quantities without additional convective transport.

8. Conclusions

Previous studies of transport across the tropical tropopause have been motivated largely by a desire to explain stratospheric water vapor. Here, we considered the ability of various idealized transport models to account for seasonal cycles, particularly that observed for CO$_2$. The models included one based on overshooting convection (Sherwood and Dessler 2001, SD01) shown previously to be consistent with observed water vapor profiles as well as other important features of the TTL (tropical tropopause layer).

We find that the level of zero net radiative heating, or “stagnation surface” in the diabatic velocity field, is a critical obstacle for penetration of seasonal variations into the stratosphere. Transport across this surface must be either convective or diffusive, with diffusive transport occurring primarily through random variations in heating rate experienced by a parcel.

Testing several simple models, we find consistently that convective transport, rather than diffusion, must remain dominant to $\theta$ surfaces at least 5–10 K above the stagnation level. Given various previous estimates of the altitude of the stagnation surface, these results imply dominant convective transport to at least 16–16.5 km. Otherwise, dispersion of parcels as they attempt to cross and/or escape from the stagnation surface produces seasonal cycles in the lower stratosphere that are unrealistically delayed and weak. An important corollary of this result is that the lower stratosphere is effectively insulated from the upper troposphere, receiving air mainly from near the surface. This conclusion rests on the validity of treating parcel heating variations outside convection as red noise, though we also tested a version of the SD01 model with some systematic horizontal variation in heating rates and obtained similar results. It is conceivable that a fuller (e.g., 3D) treatment of transport could yield dynamical (but nonconvective) mechanisms of parcel transport that differ sufficiently from the red-noise model to perform more like convection, though we cannot imagine any at present.

Convection cannot possibly be a thorough mixer up to the inferred height, as evidenced by ozone concentrations that increase sharply above 14 km. The inescapable conclusion is that convective outflows and effects must decrease gradually with height starting well below the classical tropopause, as inferred from tracer analyses (Dessler 2002). It seems unlikely that convective effects have completely disappeared prior to reaching the cold point.
While significant penetrations of updrafts beyond 16–17 km might be expected to destroy seasonal cycles by scrambling parcel “ages,” we find that this depends on the details of the convective process. Insertion of undilute, near-surface air at a wide range of heights (from the stagnation surface up to some maximum reachable level) does indeed lead to damping of seasonal cycles if the maximum reachable level passes beyond, roughly, the cold-point level. This process would require initial air parcels with quite a wide range of buoyancy ($\theta_*$) values.

Observations, however, indicate that the primary process affecting the fate of updrafts is not their initial $\theta_*$ but rather the amount and type of mixing they experience. To the extent that variations in insertion height are due to variations in turbulent overshoot mixing in the lower stratosphere, seasonal cycles will be relatively undisturbed since conservation laws ensure that mixtures tend to detraining at levels where their mean properties (e.g., $CO_2$) are not far from those of other air at that level. Consequently, the model of SD01 simulated quite realistic amplitudes and seasonal cycles, including the observed $CO_2$ phase lags with respect to the surface of 2 months on the 390-K surface and greater values higher up, despite having overshoots reaching 20 km and sufficient overshooting activity overall to fully dehydrate to observed stratospheric values. We conclude that observed seasonal cycles of $CO_2$ cannot be used to rule out a dominant role for overshooting convection in TTL budgets of water vapor, ozone, or energy.

Our examination of $H_2O$ seasonal cycles was inconclusive and points to the need for better understanding of the small-scale (e.g., microphysical or gravity wave) mechanisms that actually perform dehydration, and more precise knowledge of the temperatures at which dehydration occurs. We argue that net convective moistening is likely to respond to ambient relative humidity near the cold point, and if so, that convective dehydration would be expected to produce seasonal water vapor variations similar to those expected from in situ removal.

Acknowledgments. This work was supported by NASA EOS/IDS Grant NAG-59632 and Grants NAG-11849 and NAG-11897 to the University of Maryland and Yale University, and by a New Investigator Program in Earth Science Grant to the University of Maryland. The study was inspired partly by useful discussions with A. Andrews, E. Moyer, and C. Boering.

APPENDIX

Normalizing the $H_2O$ Seasonal Cycle

The “undamped” seasonal cycle of $CO_2$, against which the actual cycle may be normalized, is just that in the lower troposphere. Finding an equivalent for $H_2O$, however, is difficult.

The expected seasonal cycle of water vapor $q$ in the TTL is determined by those of the saturation mixing ratio $q_s$ and relative humidity $RH_{ac}$ at locations where dehydration occurs. In the SD01 model (including the version with capped convection), dehydration occurs in the overshoot region only. For a useful model evaluation, therefore, the seasonal cycle of $q_s$ in that model region must match that in the real atmosphere when and where dehydration occurs.

Unfortunately our ignorance of exactly where dehydration occurs, coupled with the steep slope of the Clausius-Clapeyron equation, leads to large uncertainties in the relevant, real-world means and seasonal variations of $q$, and $RH_{ac}$. Recall that we used soundings from the Indonesia area to set temperatures in the overshoot region. However, dehydration is not likely to be confined only to Indonesia in either season, especially JJA when the Asian monsoon develops.

For example, Sherwood (2002) used satellite and radiosonde data throughout the Tropics to estimate $q$, at 100 hPa, averaged specifically over sites where convective clouds were observed in the TTL, which should be a good indicator of dehydration (especially since other cloud types occur in basically the same locations). That analysis yielded DJF and JJA $q$, of 4.6 and 8.1 ppmv, respectively, as compared to 3.0 and 7.7 ppmv at the model 100-hPa level from the Indonesia data.

SD01 tuned the parameter $c$ to yield the observed DJF entry $q$ of roughly 2.5 ppmv; this was equivalent to diagnosing an effective $RH_{ac}$ of 83% (2.5/3.0) from the observed temperature and $q$ for that season. But if the temperatures of Sherwood (2002) had been used instead of the Indonesia ones, the diagnosed RH would have been 54% (2.5/4.6). If we supposed (for simplicity) $RH_{ac}$ to be constant throughout the year, we would anticipate a seasonal variation of $q$ of (2.5/3.0) $\times$ (7.7 $-$ 3.0) = 3.9 ppmv peak-to-peak based on Indonesia temperatures, but if we had used the Sherwood (2002) temperatures and followed the same procedure we would instead have expected (2.5/4.6) $\times$ (8.1 $-$ 4.6) = 1.9 ppmv.

Thus, the theoretical, “undamped” seasonal cycle of water vapor is quite sensitive to exactly where dehydration takes place at each time of year and what the temperature is there, even if we ignore (or expect the model to predict) any seasonal variation of relative humidity. For this paper, we take it to be the average of the two estimates above, or 2.9 $\pm$ 1.0 ppmv. We compute water vapor phase lag relative to the timing of the seasonal minima and maxima in 100-hPa temperature averaged over the Tropics. Actual damping and phase lags of the $H_2O$ cycle, then, are determined by comparing HALOE observations of $H_2O$ with the above estimates.

---

3 For example, in well-observed thunderstorms, only the strongest few updrafts come close to their theoretical undilute, adiabatic vertical velocities (e.g., Yuter and Houze 1995).
REFERENCES


