Response to Reviewer Comments, revised version

Based on the additional reviewer comments, in particular those of reviewer 3, we have implemented several modifications to our manuscript (see detailed comments below). Reviewer 3 strongly encouraged to include HALOE results and this advice has been followed: Figures 3-6 now show HALOE results alongside MLS results, while versions of Figures 2 and 7 showing HALOE results are provided as supplement.

Reviewer 1:

Some comments regarding my reasons for rejecting this paper:
My previous review and comments on this paper have focused on the lack of including dehydration in the idealized 1D model calculations, which may lead to spurious derived vertical diffusion near the cold point tropopause. The authors have addressed this criticism by including calculations with idealized dehydration during winter, but the strongest dehydration is specified at 100 hpa, not near the cold point (the minimum in observed water vapor during boreal winter is near 83 hPa in both MLS and HALOE retrievals). In the end the authors discount the importance of explicit dehydration from these tests because it "…does not improve the overall simulation throughout the year…and increases the dry bias during the summer". However, the summer results (at 80 hPa) seem unphysical to begin with (in my opinion – see below), so this is a poor argument for not including dehydration (which is the fundamental cause of the tape recorder to begin with).

To the extent that dehydration (or any other source / sink) is prescribed, i.e. does not depend on the simulated water vapor mixing ratio, the details of how it is prescribed should not matter much qualitatively – they will result in simple (possibly seasonally depend) offsets of the simulated tape recorder signal. In the extreme case one could simulate the entire tape recorder purely by specifically designed sources and sinks, without any transport contributions – in strong conflict with what is well-known about lower stratospheric transport. In our view, it is preferable then to not include any prescribed source/sink term (as we have done, following Mote et al.) and rather diagnose missing tendencies after the fact. We agree that the low bias of our simulated water vapor during boreal winter and spring is consistent with a missing sink (e.g. due to dehydration), as discussed in the manuscript.

Note that the bulk of dehydration happens below 80 hPa, and therefore is included in the lower boundary condition of our simulations. So, the statement that we didn't include any dehydration is false. After all, even the Mote et al. control setting (small vertical mixing) produces a qualitatively realistic looking tape recorder signal at 80 hPa. As an aside, dehydration is only one fundamental cause of the tape recorder, the other being vertical transport (~the movement of the tape).

While intuitively I don't like the neglect of an important term in the idealized model, it is the detailed output of the model that is problematic in my opinion. Fig. 7 shows vertical mixing maximizing with identical timing to vertical advection, with extrema during Nov-Dec and Aug-Oct (which they term 'summer'). This timing seems unphysical to me; why should vertical mixing be tied to mean vertical advection? And what is the physical mechanism for these seasonal maxima in derived mixing?

First of all, we disagree that the timing between the vertical advection and vertical mixing terms in Fig. 7 is "identical". For example, the strongest negative tendency due to vertical mixing happens in late November, 1-2 months before that of the vertical advection. Vertical mixing goes through zero tendency in early February when vertical advection is at its strongest. Second, even if both terms did have identical timing we don't see how that alone indicates anything "unphysical": 1) the mathematical relation between the two (based on the vertical gradient and vertical curvature, respectively) certainly allows for a physical relation (e.g. think of an exponential); 2) a vast range of atmospheric processes that have absolutely no physical relation to each other still share similar seasonality (e.g. arctic sea ice extent and Colorado snow cover) – in general, common seasonality is a very poor indicator of direct physical relationships between two variables.

That said, we agree that our maximized tendency due to vertical mixing during August-September seems surprising at first. However, note that this doesn't indicate that more vertical mixing occurs during that time, but that it has a bigger effect on water vapor. The vertical mixing tendency is a function of the background vertical curvature of the water vapor profile. This curvature maximizes during August-September giving rise to a large vertical mixing tendency, even if the diffusivity is held constant throughout the year. In principle, the mixing tendency could be larger even if the amount of mixing is smaller (i.e. in our case smaller K during August-September), as long as the curvature contribution overcompensates.

Also, the diagnosed horizontal advection is relatively small and peaks during boreal spring (~May), very different from large boreal summer transport expected from the Asian summer monsoon (e.g. Ploeger et al 2012 explicitly calculate a 400 K maximum for water vapor horizontal mixing during August-October). Because of the fundamental problems in these important details, I am not convinced of the key result of large diagnosed vertical mixing from this analysis. As a note, it would have been helpful to have some statistical uncertainty estimates included with the results, although I did not recommend this in my previous review.

We think this is a misunderstanding on the reviewer's side. In fact, Fig. 1a in Ploeger et al. (2012) very clearly shows that the <u>tendency</u> due to horizontal mixing is largest during boreal spring: this tendency will be largest when the H_2O difference between in-mixed extratropical air and tropical air is largest and this happens during boreal spring (e.g. difference between full red and black lines in Fig. 1A of Ploeger et al). The strongest effect in tropical mean tracer mixing ratio will be found when the <u>accumulated</u> tendency is strongest: this happens during late boreal summer when the tendency crosses through zero. The timing of this zero crossing is within a month between our result and Ploeger et al., so there is no discrepancy.

This has been further clarified in the manuscript.

The authors argue that the lack of strong vertical mixing derived using isentropic coordinates is physically consistent with the pressure coordinate calculations. However, I am not familiar with the quoted statement that "vertical mixing due to breaking gravity waves may be assumed to take place quasi-adiabatically", and a reference is not provided. The derived horizontal mixing for the isentropic case (Fig. 12) still peaks in the wrong season compared to previous work. Overall, the isentropic-coordinate calculations do not convince me that the pressure-coordinate calculations are correct (the p-coordinate calculations are more straightforward, and I believe they are not correct based on my comments above).

Basic gravity theory assumes adiabatic processes (e.g. Gill's book). In the stratosphere, diabatic processes are predominantly due to radiation. Radiative time-scales are long in the lowermost stratosphere (~30 days or longer) – long enough so that it seems reasonable to assume that gravity wave breaking happens fast enough to not "feel" any radiative damping. A clarifying statement has been included as footnote in the introduction.

We emphasize again that our isentropic coordinate results (based on the Mote et al control settings, so the associated tendencies are consistent with Mote et al) confirm existing results in the literature and that we feel that this provides some confidence in our approach and assumptions.

Reviewer 2:

We again thank the reviewer for her/his help to improve our paper. The reviewer suggests to run (and report on) more sensitivity tests for the seasonality in transport parameters. We have performed such tests and have now included additional text in the paper (section 3.2). In short, our score drops somewhat for reductions in the vertical advection seasonal cycle amplitude, slightly improves for a moderate enhancement (from 50% to 67%), but very strongly drops for more significant enhancements (beyond 75%). The score drops somewhat for very strong seasonality in K, but is otherwise hardly sensitive to changes in seasonality of the mixing parameters. In order to keep things simple we have opted to keep the 50% seasonality in all parameters.

Reviewer 3 (Tim Dunkerton):

We thank Tim Dunkerton for the overall encouraging remarks and the constructive criticism, which we address below. We have incorporated several corresponding changes in the manuscript.

This paper extends earlier analysis of the tape recorder signal in water vapor by including the seasonal cycle and focusing primarily on the lowermost part of the signal near 80 hPa. If I were interested in further research on this topic I would certainly want to consult this paper to build on the authors' results. I regard their findings as a partial improvement, perhaps half-way to where we need to go to understand fully the implications of water vapor behavior for mean vertical motion, vertical diffusion, and lateral in-mixing. The MLS data are inferior to HALOE for this purpose, the model is, well, just a model, and ERA-I is excessively diffusive, perhaps worthless of this purpose. I would probably rely on extending the HALOE analysis (Dunkerton, 2001 JAS) with the precision (for lowermost levels) illustrated by the authors.

We have now included several results based on HALOE data in the manuscript. Figures 3-6 now all include additional panels or lines based on HALOE. Versions of Fig. 2 and 7 based on HALOE are provided as supplement. The HALOE data set we used already came as zonal monthly means, which prevented us from running our analyses in isentropic coordinates (the first author now works at NCAR on a different project, so with the time constraint of a major revision we weren't able to run the entire isentropic coordinate analysis from instantaneous HALOE data, which is not as straightforward to process). Overall, results agree quite well between MLS and HALOE; if anything, HALOE suggests even greater vertical mixing strength. It is encouraging to see that the Mote et al. control setting performs much better for HALOE data (Fig. 6), on which it was based.

Two issues are overlooked that otherwise limit my appreciation for this work. First, it is now well established that the moist part of the tape recorder is driven by vertical, then lateral, transport of water vapor injected by overshooting convection in the Asian summer monsoon (Gettelman et al.,, 2004 JGR, et seq.). It is surprising if, in fact, no one has repeated Mote et al. with time-varying side boundary conditions to mimic this effect.

We agree that lateral transport, e.g. associated with the monsoons, may potentially be important, although Ploeger et al. (2012) conclude that it has a negligible effect on tropical water vapor just above the tropopause. The issue is that even though the monsoons may provide a great deal of mass transport into the TTL, the effect on tracers also depends on their background gradient. In the case of water vapor this background gradient is quite small during the relevant seasons.

Note that our 1-d model setup already included the effect of time-varying side boundary conditions: see text under Eq. 1 (section 3.2). This is now emphasized more in the text and the Gettelman et al. reference is included.

Second, although the authors seem to casually ascribe vertical diffusion to breaking gravity waves, it is well-known that such waves (if undergoing local convective instability in their phase of overturning) are not effective in mixing heat and constituents vertically (Coy et al., 1988 JAS). Inertia-gravity waves may undergo shear instability at large amplitude, altering this result possibly in a significant way (Dunkerton, 1985 JAS, et seq.).

We agree that our discussion of what may cause the vertical mixing is vague and hand-wavy. If gravity wave breaking contributes significantly then this has to happen in a non-classical way. The strong wave guide given by the large stratification jump and the tropopause inversion layer may perhaps give rise to non-linear effects that haven't been considered much in the past. Another possibility is that cloud-radiative effects modify diabatic damping in a way that gives rise to vertical dispersion, although this would then also be felt in isentropic coordinates. We admit that we don't have a clear conceptual picture at this point how this may work, but note that we are currently doing more work on this using cloud modeling.

Meanwhile, what is the role of overshooting convection in penetrating local theta surfaces? The main conclusion emerging from our work and subsequent studies of (de)hyrdration by overshooting convection is that ambient relative humidity determines the outcome: if high, leading to freeze-drying, if low, leading to hydration. While their latest additions to the text address dehydration, nothing is said to address hydration. In summary, while there are many reasons to publish this work, it will be a stepping stone, not close to any final answer, on a most intriguing problem. Publication is recommended.

We actually view hydration by overshooting convection as partially included in the vertical mixing term (technically the mixing is acting on total water in this case, which then becomes water vapor in the lowermost stratosphere through evaporation). We had mentioned this in the previous version (section 6, about half way through the paragraph discussing dehydration effects). Because of this we think that one shouldn't include an extra source term representing hydration in the 1-d model. But of course, an ad-hoc inclusion of a small source (e.g. due to hydration) during July-September together with a small sink (e.g. due to dehydration) during December-March would make our synthetic water vapor evolution agree with the observed one. Note however, that this is only true for MLS; our simulation already agrees very well with HALOE during July-September.

From a more philosophical point of view, one argument against inclusion of prescribed source / sink terms is the following. In the extreme case one could simulate the entire tape recorder purely by specifically designed sources and sinks, without any transport contributions – in strong conflict with what is well-known about lower stratospheric transport. In our view, it is preferable then to not include any prescribed source/sink term (as in Mote et al.) and rather diagnose and interpret missing tendencies after the fact.

Additional clarifying comments have been included in the text.

Minor suggestion: be sure to emphasize that only the lowermost part of the signal is analyzed. Whether your analysis benefits higher levels, with respect to annual & QBO influences on parameters, remains to be determined. I did some forward modeling prior to Mote et al. to ensure that vertical diffusivity must decrease rapidly with height above 80 hPa. Indeed, it must, otherwise the signal is too wide and decays much too fast. Frankly, if you have HALOE results not shown, I would add them prior to publication.

We feel that the text was already quite clear about the analyzed region: e.g. the title of the paper specifies that we study the "... tape recorder signal near the tropical tropopause". Qualifiers such as "in the lowermost stratosphere" or "just above the tropopause" are frequently used. Nevertheless, we have now clearly stated in the abstract that the main analyzed level is 80 hPa.

Note that we use the same functional vertical structure for all transport parameters as in Mote et al., so the vertical decay above 80 hPa is implicitly included. Our variations in K represent a vertically uniform scaling factor onto the Mote et al. profile.