

The relation between atmospheric humidity and temperature trends for stratospheric water

S. Fueglistaler,¹ Y.S. Liu,² T.J. Flannaghan,³ P.H. Haynes,³ D.P. Dee,⁴ W.J. Read,⁵ E.E. Remsberg,⁶ L.W. Thomason,⁶ D.F. Hurst,^{7,8} J.R. Lanzante,⁹ and P.F. Bernath¹⁰

Received 31 July 2012; revised 19 December 2012; accepted 2 January 2013; published 30 January 2013.

[1] We analyze the relation between atmospheric temperature and water vapor—a fundamental component of the global climate system—for stratospheric water vapor (SWV). We compare measurements of SWV (and methane where available) over the period 1980–2011 from NOAA balloon-borne frostpoint hygrometer (NOAA-FPH), SAGE II, Halogen Occultation Experiment (HALOE), Microwave Limb Sounder (MLS)/Aura, and Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) to model predictions based on troposphere-to-stratosphere transport from ERA-Interim, and temperatures from ERA-Interim, Modern Era Retrospective-Analysis (MERRA), Climate Forecast System Reanalysis (CFSR), Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC), HadAT2, and RICHv1.5. All model predictions are dry biased. The interannual anomalies of the model predictions show periods of fairly regular oscillations, alternating with more quiescent periods and a few large-amplitude oscillations. They all agree well (correlation coefficients 0.9 and larger) with observations for higher-frequency variations (periods up to 2–3 years). Differences between SWV observations, and temperature data, respectively, render analysis of the model minus observation residual difficult. However, we find fairly well-defined periods of drifts in the residuals. For the 1980s, model predictions differ most, and only the calculation with ERA-Interim temperatures is roughly within observational uncertainties. All model predictions show a drying relative to HALOE in the 1990s, followed by a moistening in the early 2000s. Drifts to NOAA-FPH are similar (but stronger), whereas no drift is present against SAGE II. As a result, the model calculations have a less pronounced drop in SWV in 2000 than HALOE. From the mid-2000s onward, models and observations agree reasonably, and some differences can be traced to problems in the temperature data. These results indicate that both SWV and temperature data may still suffer from artifacts that need to be resolved in order to answer the question whether the large-scale flow and temperature field is sufficient to explain water entering the stratosphere.

Citation: Fueglistaler, S., et al. (2013), The relation between atmospheric humidity and temperature trends for stratospheric water, *J. Geophys. Res. Atmos.*, 118, 1052–1074, doi:10.1002/jgrd.50157.

1. Introduction

[2] Ever since the famous deduction of the stratospheric circulation by Brewer in 1949 [Brewer, 1949] based on a

¹Dept. of Chemistry and Biochemistry, Old Dominion University, Norfolk, VA, USA.

²School of Mathematics and Statistics, University of St. Andrews, St. Andrews, UK.

³Dept. of Applied Mathematics and Theoretical Physics, Cambridge University, Cambridge, UK.

⁴ECMWF, Reading, UK.

⁵Jet Propulsion Laboratories, Pasadena CA, USA.

⁶NASA/Langley, Hampton VA, USA.

⁷NOAA ESRL, Boulder CO, USA.

⁸CIRES, University of Colorado, Boulder CO, USA.

⁹NOAA/GFDL, Princeton University, Princeton, NJ, USA.

¹⁰Old Dominion University, Norfolk, VA, USA.

Corresponding author: S. Fueglistaler, Dept. of Geosciences/AOS, Princeton University, Princeton, NJ 08544, USA. (stf@princeton.edu)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/jgrd.50157

handful of water vapor frostpoint measurements in the stratosphere over England, it has been known that the exceptionally low temperatures at the tropical tropopause constrain the amount of water entering the stratosphere. Most of the stratosphere is subsaturated due the radiative effects of ozone leading to temperatures higher than at the tropical tropopause. Only two additional processes affect the stratospheric humidity: a source from methane oxidation and a small sink due to ice cloud formation during the cold polar night vortex primarily in the Southern Hemisphere. Stratospheric water vapor (SWV) is in itself of interest due to its importance for stratospheric chemistry [Bates and Nicolet, 1950; Brasseur and Solomon, 2005] and the planetary radiative budget [Forster and Shine, 1999; Solomon et al., 2010]. Here, our interest in SWV arises from the possibility that the SWV long-term trend does not follow expectations based on temperature trends around the tropical tropopause [Rosenlof et al., 2001; Kley et al., 2000].

[3] The control exerted by temperatures in the tropical tropopause layer [TTL; see *Fueglistaler et al.* 2009a] on water entering the stratosphere ($[\text{H}_2\text{O}]_e$) is also evident in its temporal variability. There is a large annual cycle in tropical mean tropopause temperatures as well as in $[\text{H}_2\text{O}]_e$ [*Mote et al.*, 1996], and similarly interannual variability in tropical mean temperatures and $[\text{H}_2\text{O}]_e$ are well correlated for periods of several years [*Randel et al.*, 2004; *Fueglistaler and Haynes*, 2005].

[4] However, for cloud formation and subsequent dehydration local (in space and time), temperature minima are more important than the average over the tropics or any other subjectively determined region. Indeed, it was thought that the stratosphere is drier than expected from tropical average tropopause temperatures, and it was this apparent conundrum that led to the recognition of the importance of the spatial temperature structure [*Newell and Gould-Stewart*, 1981; *Holton and Gettelman*, 2001]. Today we know that this perception arose from averaging temperatures not at the coldest point of the temperature profile but at a pressure level of 100 hPa [*Dessler*, 1998], which is a little below, and hence warmer, than the coldest point. However, understanding the effect of the circulation through a spatially and temporally varying temperature field on humidity remains key to understanding water entering the stratosphere [*Liu et al.*, 2010] and atmospheric humidity in general.

[5] The long-term SWV trend as reported by [*Rosenlof et al.*, 2001; *Kley et al.*, 2000] corresponds to a temperature increase of +2 to +4 K since the 1960s [*Fueglistaler and Haynes*, 2005]. At the time, no analysis of temperature observations indicated a positive trend of this magnitude at the tropical tropopause. Hence, the question arises whether SWV is an example where observations contradict the expectation that atmospheric humidity trends should follow closely those of the saturation mixing ratio. Such a discrepancy could arise from changes in the locations of last dehydration (which is not captured by “tropical” means of temperature), aerosol-related changes in dehydration efficiency [*Sherwood*, 2002; *Notholt et al.*, 2005], or poorly quantified impacts of very deep convection [*Sherwood and Dessler*, 2000].

[6] In recent years, new efforts have been made to homogenize the temperature record, and new, improved meteorological reanalyses have become available that allow predictions of water entering the stratosphere based on advection through the temperature field. In parallel, instruments onboard satellites launched in the early 2000s have added almost another decade of high-quality water vapor observations. Finally, the NOAA in situ measurements have been corrected for small offsets in mirror thermistor temperatures [*Scherer et al.*, 2008; *Hurst et al.*, 2011].

[7] Here, we compare this extended record of SWV measurements against model predictions based on reanalysis data and several homogenized temperature time series. We use deliberately simple process models for dehydration in order to address the question whether there exists a fundamental mismatch between measured water vapor mixing ratios and those expected based essentially on temperature and transport pathways.

2. Data and Methods

2.1. Strategy for Comparison between Model and Observations

[8] The contribution from methane oxidation to stratospheric water complicates comparison between model predictions of

$[\text{H}_2\text{O}]_e$ and observations. Temporal variability in the stratospheric circulation leads to variability in the spectrum of age of air at a given location, with two effects on stratospheric water vapor: first, variations in stratospheric transport lead to variations in the time lag (i.e., variations in the stratospheric “age spectrum”) between observations at a given location in the stratosphere, and the $[\text{H}_2\text{O}]_e$ time series; and second, the contribution from methane oxidation varies (the older, the more methane has been oxidized). In much of the lower stratosphere, the latter difficulty is not present when considering the sum of water vapor mixing ratios and twice the methane mixing ratio HH ($[\text{HH}] \equiv [\text{H}_2\text{O}] + 2x[\text{CH}_4]$) [*Jones and Pyle*, 1984; *Le Texier et al.*, 1988]. However, some of the key observational data sets provide only water vapor.

[9] Hence, in this paper we employ the following strategy to compare model predictions of $[\text{H}_2\text{O}]_e$ to observations. First, we will compare $[\text{H}_2\text{O}]_e$ directly to measurements just a little above the tropical tropopause (we take the 82 hPa level from the MLS/Aura data), where the contribution from methane oxidation is small enough that possible variations therein do not lead to substantial variations in H_2O at that level (we estimate the mean contribution from methane to water vapor mixing ratios there to be about 0.1 ppmv with regional and temporal variations of around 0.05 ppmv; see Appendix B). Second, we will use a simplified representation of transport in the tropical lower/mid-stratosphere to compare $[\text{H}_2\text{O}]_e$ to observations at 10 hPa in the tropics. This allows a cross-check with results for the data at 82 hPa and allows a more reliable comparison with measurements that are affected by stratospheric aerosol (see also *Fueglistaler* [2012]). The comparison at 10 hPa is done for HH, and for H_2O when the observations do not provide methane data. Finally, we compare model predictions of H_2O to the NOAA balloon-borne frostpoint hygrometer (henceforth NOAA-FPH) measurements over Boulder, CO, whereby we focus on a layer around 100 hPa (which is roughly equivalent to the 82 hPa level in the tropics in potential temperature).

[10] In the following, section 2.2 presents the measurements of stratospheric water vapor considered in this study; section 2.3 describes the model used to predict $[\text{H}_2\text{O}]_e$, and water vapor and HH at 10 hPa. Section 2.4 discusses the temperature corrections applied to the model calculations to bracket the range of uncertainty in $[\text{H}_2\text{O}]_e$ due to uncertainties in the temperature record. Supporting calculations and data analyses are given in the Appendix.

2.2. Stratospheric Water Vapor and Methane Measurements

[11] Figure 1a summarizes the water vapor and methane measurements analyzed in this study. We use a combination of measurements from the Halogen Occultation Experiment (HALOE; data version 19) onboard the Upper Atmosphere Research Satellite [*Russell et al.*, 1993] and the Microwave Limb Sounder (MLS; data version 3) onboard the Aura satellite [*Read et al.*, 2007] to obtain a 20 year record of water entering the stratosphere. Similarly, we use HALOE and Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS; data version 3) [*Bernath et al.*, 2005] for the corresponding time series of HH, where we will focus on the 10 hPa level in the tropics. Further details of the merging of the data are given in Appendix A1. A

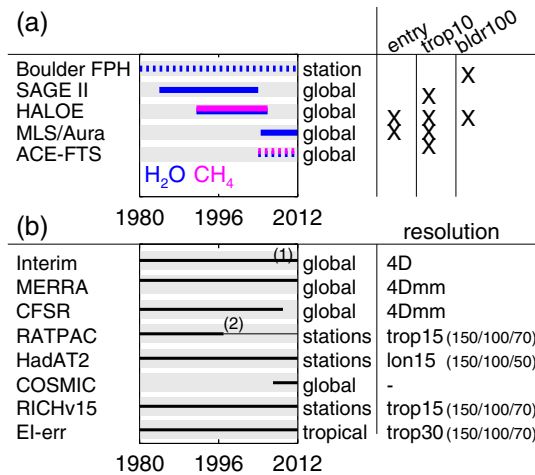


Figure 1. Overview of data used in this study. (a) Instrument name, period of available water vapor (blue) and methane (purple) measurements, spatial coverage, and specific metrics used in this study (“entry” refers to entry mixing ratios, for observations, this is the 25°S–25°N average at 82 hPa (slightly above the tropopause); “trop10” is the tropical average at 10 hPa; “bldr100” is the location of Boulder at 100 hPa). Dashed lines refer to data with low sampling rate at location of interest. (b) Temperature data name, period of available data, spatial coverage, and specific data used here (4D, full four-dimensional data; 4Dmm, monthly means of four-dimensional data; trop15, tropical average 15°S–15°N; lon15, longitude resolved average 15°S–15°N; trop30, tropical average 30°S–30°N). Label (1) refers to onset of assimilation of COSMIC-GPS temperature data; label (2) to the end of manual correction in RATPAC, and switch to IGRA data. See text for acronyms of data.

result of particular interest here is that HALOE and MLS/Aura measurements have not only different mean mixing ratios but also systematically different amplitudes of variations in particular around tropopause levels that cannot be reconciled simply by the different vertical resolution of the two measurements (Appendix A1). We will return to this point in the discussion.

[12] Further, we use the SAGE II version 6.20 water vapor data [Thomason *et al.*, 2004] and the homogenized NOAA-FPH in situ data over Boulder, CO [Scherer *et al.*, 2008; Hurst *et al.*, 2011]. The SAGE II and NOAA-FPH data are used as additional benchmark time series. Lack of methane data, issues with aerosol contamination for SAGE II (see Appendix A3), and, for the Boulder FPH data, sparse sampling (on average once per month at a Northern midlatitude location) render model-observation comparisons for these observations subject to additional uncertainties. However, the results shown below suggest that these additional uncertainties are smaller than the discrepancies between model and observations (and between observations). The SAGE II and NOAA-FPH data therefore provide important information in particular also for the period before HALOE (i.e., before September 1991). All observational data have been screened and processed following the recommendations available in the published literature (see also Appendix A).

2.3. Model Calculations

2.3.1. Model Calculations of $[H_2O]_e$ with Trajectories

[13] We predict $[H_2O]_e$ based on the temperature history of air entering the stratosphere derived from trajectory calculations. All model predictions in this paper are based on back trajectories based on European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim data [Dee *et al.*, 2011]. Trajectories may be calculated using the horizontal wind fields and pressure tendencies (“kinematic trajectories”) or diabatic heating rates (“diabatic trajectories”). The latter trajectories are less dispersive (also because the model diabatic heating rates are frequently averaged over the forecast period, whereas the pressure tendencies are the instantaneous fields), but Liu *et al.* [2010] showed that for the modern 4D-var ERA-Interim reanalysis, the differences between kinematic and diabatic trajectories are much smaller than for older reanalyses. We have calculated both kinematic and diabatic trajectories and discuss results from both methods below. However, we will show that the diabatic trajectories are affected by a very obvious erroneous drift. Since all other results are very similar between the two methods, we discuss results primarily based on the kinematic trajectories. Restriction to a single representation of transport may have implications for the time series of $[H_2O]_e$ if temporal changes in transport are not well represented in ERA-Interim. We will return to this point in the discussion.

[14] For the model prediction of $[H_2O]_e$, we use reverse domain filling trajectories started once per month from 1981 to 2011 at 82 hPa between 30°S and 30°N (on a grid with $2^\circ \times 2^\circ$ lon/lat spacing), and traced back in time 1 year (see [Liu *et al.*, 2010] for further details). Typically, about 95% of air parcels can be traced into the troposphere, and we calculate $[H_2O]_e$ from the minimum saturation mixing ratio encountered during ascent. We refer to this as the Lagrangian Dry Point (LDP) estimate.

[15] Similarly, we use reverse domain filling trajectories started once per month for the entire period to obtain a model estimate that can be compared to the NOAA-FPH measurements over Boulder, CO. For this case, trajectories are started on a finer grid ($0.25^\circ \times 0.25^\circ$ lon/lat) localized around Boulder (from 35°N to 45°N and 95°W to 115°W) on the pressure levels 110 and 90 hPa (potential temperatures similar to that of the 82 hPa level in the tropics). These model predictions are compared to the measurements in the same pressure range.

[16] In addition to the LDP estimate, we use a microphysical box model (results labeled “EI-cloud”) that evaluates the dehydration along the trajectory. The model used here is similar to that of Fueglistaler and Baker [2006] and calculates gravitational removal of condensate based on the fall speed of the ice particles size. Condensation and evaporation controlling the particle size (monodisperse) is based on the change in saturation mixing ratio (i.e., assumes thermodynamic equilibrium between gas and condensate). For the gravitational removal of the condensate, two closure schemes exist that both operate with an implicit “cloud layer depth” but redistribute the gravitationally removed water by either conserving particle number density or particle size. Also, a nucleation threshold can be prescribed (upon reaching the nucleation threshold, the system is assumed to return to thermodynamic equilibrium instantly).

[17] Figure 2 shows a schematic of the saturation mixing ratio along a trajectory with the position of the Lagrangian Dry Point. Further, the figure shows the evolution of the total water mixing ratio (the difference between total water and the saturation mixing ratio is the “cloud”) for a specific set of parameters. The figure illustrates that with the cloud box model, the water vapor mixing ratio is given by the position of “Last Saturation,” with some contribution from evaporated ice that has not fallen out.

[18] We find that the more realistic representation of dehydration affects the average entry mixing ratio but that model predictions with a given set of parameter combinations yield very similar results as the LDP estimate for the interannual variability of the period 1981–2011. Therefore, in this paper we will discuss results based on the cloud microphysical box model only in the context of the mean annual cycle.

2.3.2. Methane Entry Mixing Ratios

[19] We use the NOAA/Earth System Research Laboratory (ESRL) tropospheric methane measurements from Mauna Loa, Hawaii [Dlugokencky *et al.*, 1995] to estimate methane entry mixing ratios. The time series of methane entry mixing ratios is extended back in time with the methane mixing ratios measured in Antarctica [Etheridge *et al.*, 1992; Dlugokencky *et al.*, 1994], whereby we adjusted the data from Antarctica with a constant offset to achieve a good match with the tropics for the overlap period. The resulting time series of methane entry mixing ratio is shown in the Appendix (Figure B1a) for the period where HALOE and ACE-FTS methane measurements are available. Due to the long atmospheric lifetime of methane, this approach yields fairly accurate estimates of methane entry mixing ratio variations for the timescale of years and longer, which is also evident in the comparison against HALOE stratospheric methane measurements shown in the Appendix B (Figure B1a).

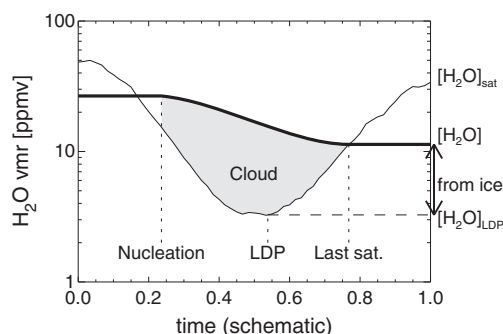


Figure 2. Schematic of dehydration model. The thin line is the saturation mixing ratio (prescribed for this example) along a trajectory, and the bold line is the total water mixing ratio (prescribed initial conditions for this example). Gravitational removal of condensate is calculated with a simple microphysical box model (see text; for this example, a nucleation threshold of 1.6 for the saturation mixing ratio is prescribed). The position of the minimum saturation mixing ratio of a trajectory is the Lagrangian Dry Point (LDP), which may be taken as an estimator of the water mixing ratio at the end point of the trajectory. The saturation mixing ratio difference between the point of last saturation and the LDP is the contribution from evaporating ice that is neglected in the LDP estimate.

2.3.3. Model for Transport in the Tropical Lower Stratosphere

[20] (Fueglistaler, 2012) showed that for the tropical average at 10 hPa, a simple representation of transport based on a mean age spectrum for this location can account for the leading-order effects of stratospheric transport. Fueglistaler [2012] fitted the age spectrum of the solution to the one-dimensional advection-diffusion equation to HALOE water vapor observations, with a modification of the tail of the distribution to an exponential fall-off to match the approximate mean age of air [Waugh and Hall, 2002]. Here, we use his age spectrum for the tropics at 10 hPa to predict H_2O and HH at that location. For HH, the prediction is simply

$$[\text{HH}]_{\text{entry}} \equiv [\text{H}_2\text{O}]_{\text{entry}} + 2 \cdot [\text{CH}_4]_{\text{entry}} \quad (1)$$

convolved with the age spectrum at 10 hPa $h(\tau)_{10\text{hPa}}$:

$$[\text{HH}]_{10\text{hPa}}(t) = [\text{HH}]_{\text{entry}}(t) * h(\tau)_{10\text{hPa}} \quad (2)$$

where the symbol $*$ denotes the convolution. For H_2O , we assume a constant fraction of oxidized methane f_{ox} , such that

$$[\text{H}_2\text{O}]_{10\text{hPa}}(t) = ([\text{H}_2\text{O}]_{\text{entry}}(t) + 2f_{\text{ox}} \cdot [\text{CH}_4]_{\text{entry}}(t)) * h(\tau)_{10\text{hPa}} \quad (3)$$

where f_{ox} is determined from the ratio between observed methane at 10 hPa and the convolution of $[\text{CH}_4]_{\text{entry}}$ with the age spectrum. Appendix B provides further details and shows that this approach also provides a sensible estimate for the contribution of oxidized methane to water vapor over Boulder.

2.4. Temperature Corrections

[21] The temperature structure in the tropical tropopause layer shows significant differences between different reanalyses, and temperature trends differ substantially between different reanalyses and radiosonde homogenization efforts (a point recently emphasized again by Wang *et al.* [2012]). A brief overview of key differences in temperatures between different data sets is given in section 3.1, and a detailed analysis is provided in Appendix C. In the following, we describe the procedures employed to estimate the impact of temperature differences between different data sets on water entering the stratosphere.

[22] In order to retain the vertical resolution of the ERA-Interim temperature data (with levels near 153, 132, 113, 96, 80, and 67 hPa in the layer of interest) and its higher-frequency variation, we substitute only the quasi-stationary (i.e., monthly mean) temperature field of ERA-Interim with that from the alternative temperature data. For temperature data available as anomalies from the mean annual cycle only (see below), the differences in the anomalies of monthly means are added to the ERA-Interim data, with the mean offset calibrated against the average of the period 2007–2011 where ERA-Interim temperatures in the Tropical Tropopause Layer have less biases than before 2007 (see below). Where necessary, data are linearly interpolated in log-pressure space (note that the differences in temperature profile do not have the strong curvature of the temperature profile near the tropopause, and hence linear interpolation is sufficient).

[23] Figure 1b summarizes the temperature data used in this study. The first temperature correction considered is based on the time-mean difference between ERA-Interim and COSMIC temperature data [Rocken *et al.*, 2000] and the time series of the ECMWF temperature error statistics (analysis minus input radiosonde data) between 30°S and 30°N. Calculations based on this correction are labeled “EI-corr.” The introduction of COSMIC temperature data to the ERA-Interim assimilation system toward the end of 2006 led to a temperature shift in particular at 100 hPa of order 0.5 K. This occasion of an obvious discontinuity in the ERA-Interim temperature data is further discussed in detail in Appendix C2, and implications on diabatic heating rates used for trajectory calculations are discussed in section 3.

[24] Further, we use the four-dimensional monthly mean temperature from NASA’s Modern Era Retrospective-Analysis (MERRA) [Rienecker *et al.*, 2011] and the Climate Forecast System Reanalysis (CFSR) [Saha *et al.*, 2010]. Results based on substitution of the quasi-stationary temperature field by these reanalyses are labeled “EI-MERRA” and “EI-CFSR,” respectively. (We also considered National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR), but its temperature trend around the tropical tropopause is so different from that of all other data sets (see section 3.1) that we do not show $[H_2O]_e$ based on these data sets.)

[25] Finally, we use the monthly mean data from three different homogenization efforts of the radiosonde record, namely, Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) [Free *et al.*, 2005], HadAT2 [Thorne *et al.*, 2005], and RICHv1.5 [Haimberger *et al.*, 2012]. Since $[H_2O]_e$ is predominantly controlled by temperatures in the inner tropics [Fueglistaler *et al.*, 2005; Fueglistaler and Haynes, 2005], we only use stations between 15°S and 15°N. The sparse and uneven sampling of this latitude belt by radiosondes does not allow a four-dimensional temperature correction. Instead, we average the differences between the RATPAC stations and ERA-Interim at the same location to produce a “tropical mean” average temperature difference profile. For HadAT2, we construct a zonally resolved inner tropical temperature correction profile (see Appendix C4), while for RICHv1.5, we use the zonal means (at the 10° resolution in latitude of the RICHv1.5 data) of the differences in the grid cells where RICHv1.5 reports data. Results based on these corrections are labeled “EI-RP,” “EI-HadAT2,” and “EI-RICHv15.”

3. Results

3.1. Uncertainties in Trends and Spatial Structure of Temperatures

[26] Figure 3a shows the ERA-Interim temperature profile averaged from 15°S to 15°N from the upper troposphere to the lower stratosphere. The amplitude of the mean annual cycle (approximately the difference between the February and August profile) peaks just above the cold point tropopause with a peak-to-peak amplitude of about 8 K. The figure also shows the climatological mean cumulative distribution function of the pressure of the Lagrangian Dry Point of air entering the stratosphere (as approximated by the back trajectories started at 82 hPa), which confirms our assertion that $[H_2O]_e$ is controlled by conditions in a thin layer around the tropopause.

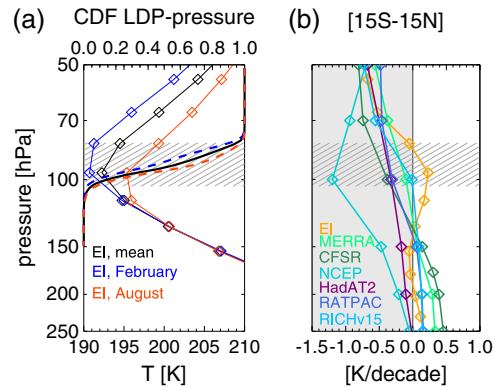


Figure 3. (a) Mean temperature profiles from ERA-Interim (diamonds indicate data levels) of the period 2006–2011 (annual mean, and means for February and August). The thick lines are the cumulative distributions (with x axis at top of figure) of the pressure of the Lagrangian dry points from trajectory calculations over the same period. The gray shaded layer is the [5%, 95%] interval of the distribution. (b) Profiles of linear temperature trends within 15°S to 15°N over the period 1980–2009 for the temperature data used in this study. (For HadAT2, RATPAC, and RICHv1.5, the trends are the averages over all stations (valid grid cells) within the latitude range 15°S–15°N).

[27] The figure also shows how the LDP pressure distribution responds to the mean annual cycle in temperatures, with a shift toward lower pressures during the cold season. Figure 3b shows the linear temperature trends of the temperature data used in this study for the period 1979–2009. The figure shows that in the layer of interest here (as indicated by the 5 to 95 percentile range of the LDP distribution), the trend differs substantially between the data sets.

[28] Figure 4a shows the time series of the temperature anomalies (mean annual cycle subtracted) at 100 hPa (the only level within the 5 to 95 percentile interval of the LDP pressure distribution where all data sets report data) at the locations of the stations included in RATPAC (for RICHv1.5, we take the average over all grid cells with valid data in the 15°S to 15°N band). Figure 4b shows the same information for the area average between 15°S and 15°N. The figure shows that the strong cooling of NCEP/NCAR temperatures at 100 hPa shown in Figure 3b is the result of a fairly erratic drift relative to the other data sets. Further, the figure shows that the largest differences between ERA-Interim and the remaining other data sets are the trend in the 1980s and the shift around the year 2006. It is these two periods that account for much of the differences in the trends over the full period shown in Figure 3b.

[29] Figure 4c shows the differences between the temperature average at the location of the stations and the area mean temperature (for the reanalyses only). The figure shows that the “station average” and the area average can diverge substantially for periods of a year or so. This effect is primarily due to the underrepresentation of the Eastern Pacific in the station data leading to a sampling bias of zonally asymmetric temperature variability. The latter is largely due to the El Niño/Southern Oscillation (ENSO), evident in the high correlation with the ENSO index shown in the figure. As such,

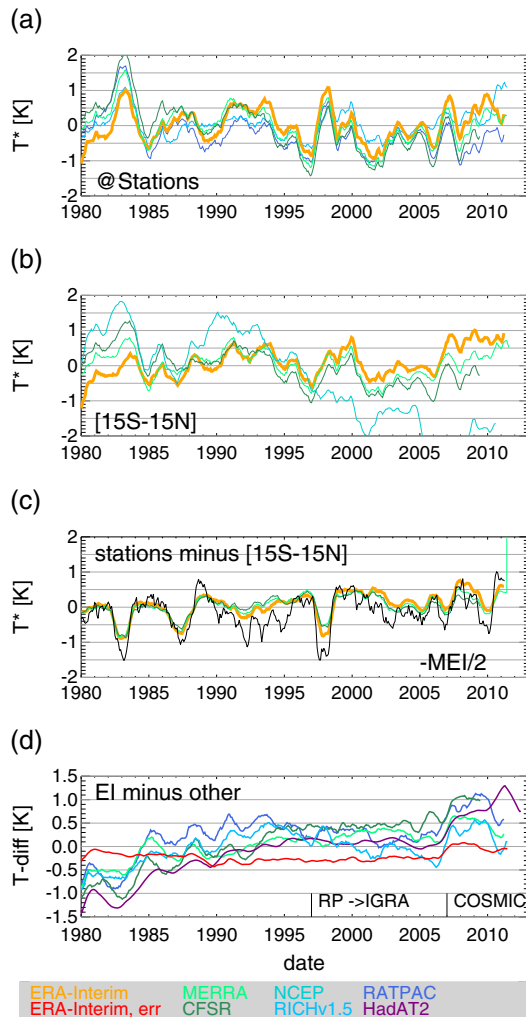


Figure 4. (a) Interannual variability of temperature at 100 hPa averaged at stations within 15°S–15°N listed in RATPAC (data are only taken when RATPAC reports valid data); for RICHv1.5, data are taken at all valid grid cells in the same latitude band. (b) Interannual variability of temperature at 100 hPa averaged between 15°S and 15°N. (c) Temperature difference between station and area mean for data sets with area mean. Also shown is the MEI index for ENSO (sign reversed and scaled with 1/2 to match temperature record). (d) Differences between ERA-Interim and the other data sets (Interim minus other data). For RATPAC, HadAT2, and RICHv1.5, the figure shows the average difference at valid locations within 15°S and 15°N. Also shown is the transition to IGRA data for RATPAC and onset of assimilation of COSMIC-GPS temperature data for ERA-Interim (labeled “COSMIC”). Data sets color coded as indicated at bottom of figure. All curves smoothed with 12 month running mean.

the station averages are quite poor surrogates for true tropical means in particular for ENSO-related variability. However, Figure 4d shows that the average temperature difference between two data sets is typically very similar for the area means (the solid lines) and the station averages (the dashed lines). Consequently, while the difference at the station

location averaged over all stations (as done for EI-corr, EI-RP, and EI-RICHv15) cannot account for the effect of differences in localized temperature trends on $[H_2O]_e$, it is a fair estimator of the true tropical average difference between two temperature data sets. Finally, note the shift at the end of the year 2006 in all differences which is the result of the previously mentioned temperature change in ERA-Interim following the assimilation of COSMIC-GPS temperature data.

3.2. The Annual Cycle of $[H_2O]_e$

[30] Figure 5 summarizes key properties of the annual cycle of the cold point temperatures in ERA-Interim and the Lagrangian model calculations. Figure 5a shows the mean and the amplitude of the annual cycle in ERA-Interim temperatures at the cold point tropopause for each grid point (spacing $1^\circ \times 1^\circ$) in the tropics, with the longitude of the position color coded. The black asterisks in the figure show the same information for the zonal mean quantities, with latitudes of 5°N/S, 10°N/S, 15°N/S, and 20°N/S labeled. Evidently, the properties of the cold point tropopause vary widely in the tropics, and without knowledge which areas are involved in the final dehydration of air entering the stratosphere, little can be said about $[H_2O]_e$.

[31] The LDP calculation provides an objective weighting of the temperature field (with the “weighting function” varying in space and time in response to variations in transport and temperature field [Fueglistaler and Haynes, 2005]), and Figure 5a shows that this estimate (black “X” and “+” for kinematic and diabatic trajectories, respectively) is, in the annual mean, several Kelvin lower than that of any local cold point tropopause temperature. This shows the strong prevalence of the Lagrangian estimate to sample times and positions where temperatures are anomalously low. We find this effect to be dominated by the temporal variability rather than the quasi-stationary temperature structure originally proposed by Newell and Gould-Stewart [1981] and Holton and Gettelman [2001]. Substitution of the LDP temperature with the monthly mean temperature at the LDP location (black diamond, for kinematic trajectories only) increases the average LDP temperature by 3.2 K. After substitution with the zonal mean, monthly mean (which eliminates the quasi-stationary zonal asymmetry) yields only an additional 1.3 K increase (the black square, for kinematic trajectories only). The LDP distribution is strongly focused on the inner tropics [Fueglistaler and Haynes, 2005], but the figure shows that the amplitude of the annual cycle is less for the LDP estimate (about 3.5 K for kinematic trajectories and about 4.2 K for diabatic trajectories) than that of the cold point tropopause in much of that region (equatorward of 10°, the zonal mean peak-to-peak amplitude is 4 K and more). The explanation for this result is that the tropical mean water vapor on a given level (here the 82 hPa level) reflects a distribution of times since last dehydration. Consequently, there is not only a phase lag but also a decrease in amplitude due to the width of the spectrum of times since the LDP. Figure 6a shows the annual cycle in cold point temperature and the annual cycle in LDP temperature (of the kinematic trajectories) at the time of the trajectory arriving at 82 hPa. The figure further shows that the amplitude of the annual cycle of LDP temperatures binned according to time of the LDP agrees fairly well with that of the 10°S–10°N average cold point. Figure 6b shows the annual mean LDP

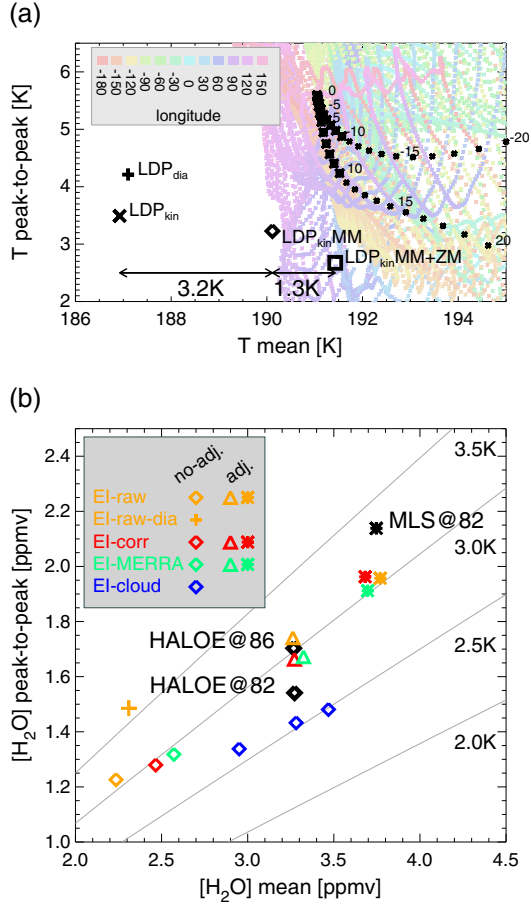


Figure 5. Characteristics of the mean annual cycle. (a) Average and peak-to-peak amplitude of the mean (1994–2005) annual cycle of monthly mean cold point temperatures from ERA-Interim between 20°S–20°N, color coded for longitude. The black asterisks show the corresponding information for the zonal means (larger symbols for latitudes equatorward of 10°; latitudes of 0°, ±5°, ±10° ... labeled). The black “X” shows the same information for the Lagrangian Dry Point (LDP) of troposphere-to-stratosphere transport from kinematic trajectories (black “+”: ditto for diabatic trajectories). The black diamond and square show the kinematic LDP estimate after substitution of LDP temperature with the monthly mean (MM) and monthly, zonal mean (MM+ZM). (b) Ditto for water vapor. Black diamonds and asterisk show HALOE (1994–2005) at 86 hPa, interpolated to 82 hPa, and MLS (2005–2011) at 82 hPa between 25°S and 25°N. LDP model results shown in color (inset; see text for details). Adjusted (“adj.”) data: with constant LDP temperature offset. Diamonds and triangles (“+” for diabatic trajectories) refer to the period 1994–2005; asterisks to 2005–2011.

time distribution. The distributions have typically a width of several months, with a broader distribution for trajectories ending at 82 hPa during boreal summer due to the Eulerian average temperature being lower during boreal winter. Figure 6c plots the cumulative distribution frequencies (CDF) of the LDP time spectra (as shown in Figure 6b) for the kinematic trajectories against that of the diabatic trajectories (i.e., each data point is the CDF of monthly binned LDP

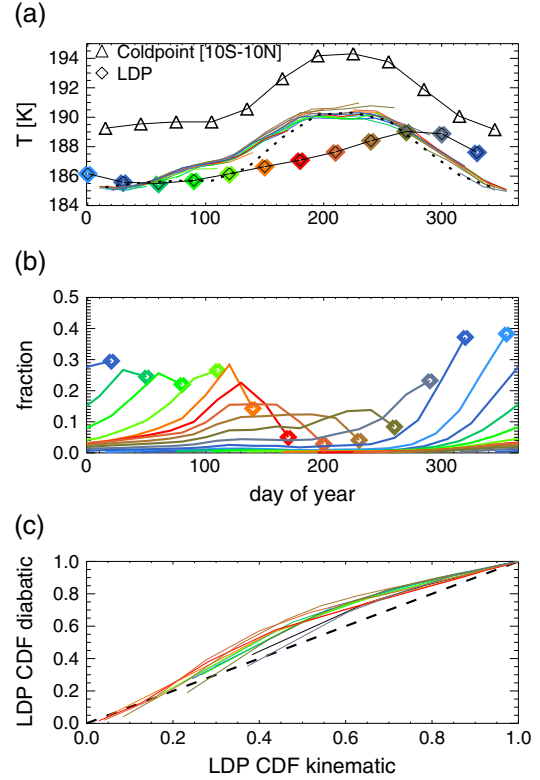


Figure 6. (a) The climatological mean annual cycle of LDP temperature of kinematic trajectories ending at 82 hPa (diamonds, months color coded). The thin colored lines show the LDP temperature averaged at the LDP time (color refers to month ending at 82 hPa). Black triangles show the mean annual cycle of ERA-Interim coldpoint temperatures averaged between 10°S and 10°N and with offset (dotted line) for better visual comparison with LDP results. (b) The corresponding climatological mean frequency distribution of LDP time for trajectories ending at 82 hPa (same color code as above, diamond for end month). The mean LDP temperature of a given month (diamond in Figure 6a) is the sum of the LDP temperature (the thin line shown in Figure 6a) times the frequency as shown in Figure 6b of the same month. (c) The cumulative distributions frequencies (CDF) of LDP frequency (as shown in Figure 6b) of kinematic versus diabatic trajectories. Black dashed line is 1:1 line; kinematic trajectories are more dispersive and have a broader LDP time distribution.

frequency). The figure shows that for each month at 82 hPa, the diabatic trajectories have a narrower distribution of time since the LDP, which leads to the larger amplitude of the mean annual cycle seen in Figure 5a. (The average temperature sampled each month is similar to that of the kinematic trajectories shown in Figure 6b.)

[32] Figure 5b compares the Lagrangian estimate in terms of water vapor (rather than frost point temperature) to observations by HALOE and MLS. We find that the large differences between HALOE and MLS are primarily due to instrumental differences (about 0.7 ppmv) and to a lesser extent (about 0.2 ppmv) due to true differences between the HALOE and MLS period (see details in Appendix A1). The figure shows that all model predictions are dry biased compared to both

HALOE and MLS, with the biases being smaller for EI-corr, EI-MERRA (and also EI-CFSR, not shown) as expected from the analysis of temperatures shown in section 2.4 and Appendix C. *Liu et al.* [2010] provide an error calculation that considers also small-scale temperature variability not resolved in the reanalysis data and show that the dry bias is significant even when uncertainties in the temperature field and transport timescale are considered. *Liu et al.* [2010] further show that a constant frostpoint temperature offset applied to the LDP temperatures brings the model $[\text{H}_2\text{O}]_e$ to good agreement with observations.

[33] In the following, we follow *Liu et al.* [2010] and apply a constant temperature correction to the LDP temperature of each trajectory (we refer to these model calculations as “adjusted”). Due to the nonlinear dependence of the vapor pressure to temperature, this adjustment also affects the amplitude of water vapor variability (i.e., evaluation of the water vapor pressure at higher average temperatures gives a larger difference for a given frostpoint temperature difference), a point that requires further discussion here. The gray lines in Figure 5b show the difference of the saturation mixing ratios (corresponding to the peak-to-peak amplitudes) for constant temperature differences of 2, 2.5, 3 and 3.5 K as a function of the saturation mixing ratio of the average temperature, evaluated at a pressure of 90 hPa (typical for the LDP pressure). The triangles and asterisk symbols in Figure 5b show the adjusted model results, whereby the temperature correction applied is chosen such that the *mean* value of the model results agrees with those of HALOE (triangles) and MLS (asterisks). The figure shows that the adjusted LDP estimates of $[\text{H}_2\text{O}]_e$ scale well along the gray lines of mean/amplitude for the amplitude of the modeled LDP temperatures. The adjusted LDP calculations follow the line for a peak-to-peak frost point temperature difference of 3 K, while the actual peak-to-peak amplitude is about 3.5 K (the “x” in Figure 5a). This difference is explained by the fact that the LDP estimate is an average over a distribution of temperatures, with the corresponding saturation mixing ratios not scaling linearly.

[34] The required temperature adjustment for EI-corr to match HALOE is about 1.7 K (for typical LDP temperature and pressure combinations, the saturation mixing ratio changes approximately by 0.5 ppmv per 1 K temperature change). The required adjustments for EI-MERRA and EI-CFSR are similar, which suggests that based on currently available reanalyses, the dry bias is not a consequence of errors in the quasi-stationary temperature field, supporting the conclusions of *Liu et al.* [2010]. Rather, the adjustment may reflect overestimation of dehydration in the LDP calculation or the effect of other processes missing in the model calculations. Indeed, the calculations using the cloud microphysical box model (blue diamonds in Figure 5b) with sensible values for particle number densities and “cloud vertical thickness” can reproduce the measurements well except for the difference in the scaling of the amplitude. We will return to this point in the discussion.

3.3. Interannual Variability

[35] In the following, we discuss the temporal variability of the data after subtracting the climatological mean annual cycle. The mean annual cycle is defined as the mean value for each month over the period considered. Note that the remainder both has variability on shorter than 12 months

and longer than 12 months; for simplicity, we refer to these deseasonalized time series also as “interannual variability.”

3.3.1. Artifacts in the Diabatic Trajectory Calculations

[36] Figure 7a shows the time series of the deseasonalized LDP temperatures for kinematic and diabatic trajectories. The figure shows that the amplitude of the variations are on the order of 1 K. The difference between the two model estimates shown in Figure 7b is about an order of magnitude smaller except at the end of the year 2006, when the LDP temperature of the diabatic trajectories decreases by about -0.4 K relative to that of the kinematic trajectories. The timing of this divergence is exactly when the ERA-Interim temperatures shift in response to the assimilation of the COSMIC-GPS temperature data. However, this change affects both the kinematic and diabatic trajectories, and it turns out that the cause for the shift is the impact of the temperature shift on the radiative heating rates [see also *Fueglistaler et al.*, 2009b]. The warming at 100 hPa induces a reduction in net radiative heating, which increases the residence time at levels around the tropopause. As discussed in *Liu et al.* [2010] (see their Figure A1), an increase in residence time induces a decrease in the LDP temperature due to the higher probability to sample regions and times of anomalously low temperatures.

[37] We find that this event in 2006 is the most obvious artifact that can be readily explained, but the fact that other larger deviations occur in periods of anomalous temperatures

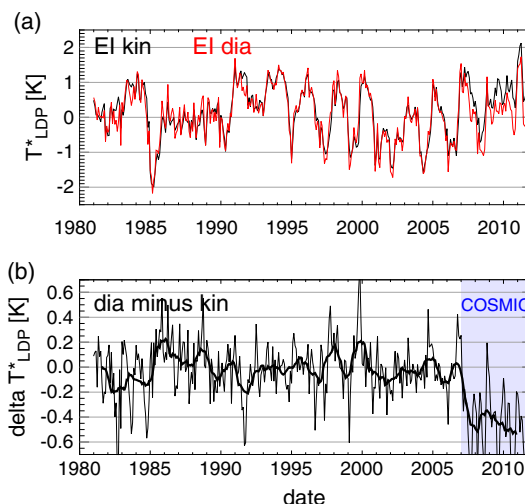


Figure 7. (a) Time series of LDP temperature anomalies (mean annual cycle subtracted) of trajectories ending at 82 hPa from kinematic (black) and diabatic (red) calculations (no temperature corrections). (b) Time series of the difference of LDP temperatures between kinematic and diabatic trajectories shown in Figure 7a; black solid line smoothed with 12 month running mean. COSMIC-GPS temperature data are assimilated to ERA-Interim from end of 2006 onward (blue shading). Note: The shift end of 2006 is an artifact in the diabatic trajectories. The temperature change with COSMIC assimilation affects both trajectory calculations, but diabatic trajectories are also affected by associated change in ERA-Interim radiative heating rates, leading here to slower ascent and consequently to lower LDP temperatures (for discussion, see *Liu et al.* [2010]).

(see, e.g., the deviation during the Pinatubo period) also follow the expectations. Since results based on kinematic and diabatic trajectories are otherwise fairly similar, we use in the following the results based on the kinematic trajectories.

3.3.2. Interannual Variability of $[\text{H}_2\text{O}]_e$ (at 82 hPa)

[38] Figure 8a shows the time series of tropical, interannual variability of water vapor from the merged HALOE/MLS data from slightly below the tropopause up to 1 hPa. The figure shows the upward propagation of $[\text{H}_2\text{O}]_e$ anomalies similarly to those of the annual cycle [Mote *et al.*, 1996]. Figure 8b shows the observations and model calculations at 82 hPa. The labels “E2” and “E3” highlight the timing of two events we will refer to frequently, namely the eruption of Mt. Pinatubo in June 1991 (E2) and the marked drop in water vapor in October 2000 (E3); the event “E1” will be introduced later.

[39] The results confirm the earlier finding [Fueglistaler and Haynes, 2005] of an overall high correlation between

observations and LDP model results. The much longer time series available now in conjunction with the model calculations based on the improved reanalysis data allows for the first time a close inspection of the differences between observations and model predictions.

[40] Figure 8c shows the high-pass filtered model predictions of $[\text{H}_2\text{O}]_e$ and the observations at 82 hPa. Figure 8d shows the corresponding low-pass filtered data. The filter employed is a simple rectangular window with duration 30 months, which separates variability on shorter timescales from variability on multiyear timescales. Sensitivity calculations with filters with smaller side lobes and with variations in the width of the window by ± 6 months yield very similar results and are not further discussed. The two figures show that the higher-frequency variations are very well captured by all model calculations, with correlation coefficients of around 0.9. For all model calculations, the amplitude of the variations

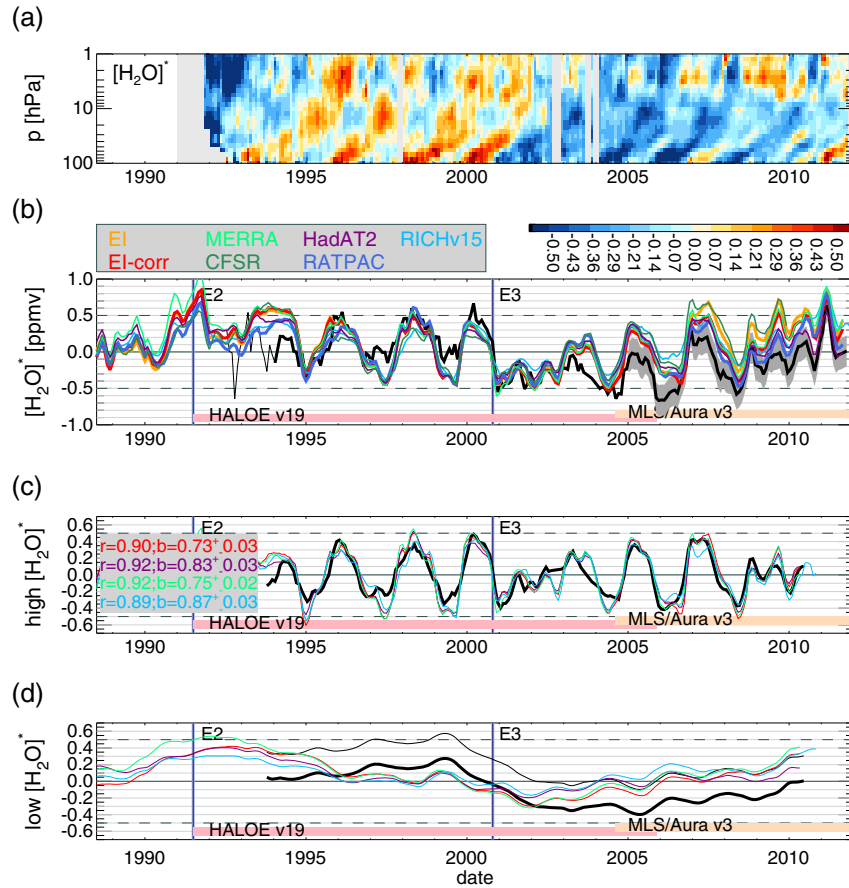


Figure 8. (a) Interannual variability of tropical (25°S – 25°N) water vapor (H_2O^* , color coded, in ppmv) from HALOE (up to 2005) and MLS (from 2005 onward), with MLS adjusted (see text). (b) H_2O^* from HALOE and MLS (as in Figure 8a) at 82 hPa (black line); HALOE interpolated linearly in log-pressure. Dark gray for MLS is 1 std. deviation uncertainty in the merging of HALOE and MLS (see Appendix). Lagrangian model estimates (color code shown in figure) based on ERA-Interim, ERA-Interim corrected, and with temperatures corrected based on the quasi-stationary temperature field of MERRA and CFSR, and tropical means for homogenized radiosonde data (see text for details). All calculations adjusted with constant frostpoint offset to match mean HALOE data. (c) High-pass filtered (running mean over 30 months subtracted; see text) observations and model calculations. Inset: correlation coefficients and slope b of linear regression. (d) Corresponding low-pass filtered data. For visual clarity, only EI-corr, MERRA-, HadAT2, and RICHv1.5-corrected data shown in Figures 8c and 8d. Thin black line is merged HALOE/MLS time series shifted to qualitatively match model predictions for MLS period.

is slightly larger (the slope of the linear regression is $b \sim 0.75$, corresponding to an overestimation of the amplitude by about 25%) than in the observations. Recall that the model calculations have been adjusted with an offset to match the mean and amplitude of the mean annual cycle, and, while not dramatic, this amplitude mismatch therefore deserves being noted.

[41] For the low-pass filtered data, the situation is rather different. All model calculations show a drying tendency relative to HALOE from around 1994 until about the end of 1996. A reverse drift (model is moistening relative to HALOE) is observed in the early 2000s. These drifts roughly cancel such that the pre-1995 and post-2005 data are on approximately the same level, whereas the drop in October 2000 so prominent in the HALOE data is far less pronounced in the model predictions. We will return to this point in the discussion.

3.3.3. Interannual Variability of [HH] at Entry and at 10 hPa in the Tropics

[42] Figure 9a shows the time series of model [HH]_e for the period 1981–2011, and Figure 9b shows the model predictions

for 10 hPa in the tropics (discussed in section 2.3.3) together with the corresponding HH measurements from HALOE and ACE-FTS. Also shown are the events E2, E3, and E1, where E1 refers to the large oscillation in entry mixing ratios in 1984–1985. Note that for 10 hPa, the timing of the events E1 to E3 has been shifted by the climatological mean phase lag at 10 hPa. Similarly, the time axis of Figures 9b–9d is shifted relative to that of Figures 9a to facilitate visual inspection of the correspondence of variations at entry and at 10 hPa.

[43] Figure 9a shows a remarkable change in the characteristics of [HH]_e (Figure 9a; these variations are dominated by variations in [H₂O]_e, and methane plays a negligible role) in the late 1980s when the quasi-periodic variability (with period around 2 years) is weak in all model calculations. This comparatively quiescent period was preceded by a large-amplitude oscillation (event E1).

[44] As in Figure 8, the figure also shows the high- and low-pass filtered data (Figure 9c and 9d). The correlation coefficients for the high-pass filtered data of deseasonalized HH at 10 hPa is less than for [H₂O]_e (Figure 8c), but still high (around 0.8). Given that we have used a representation

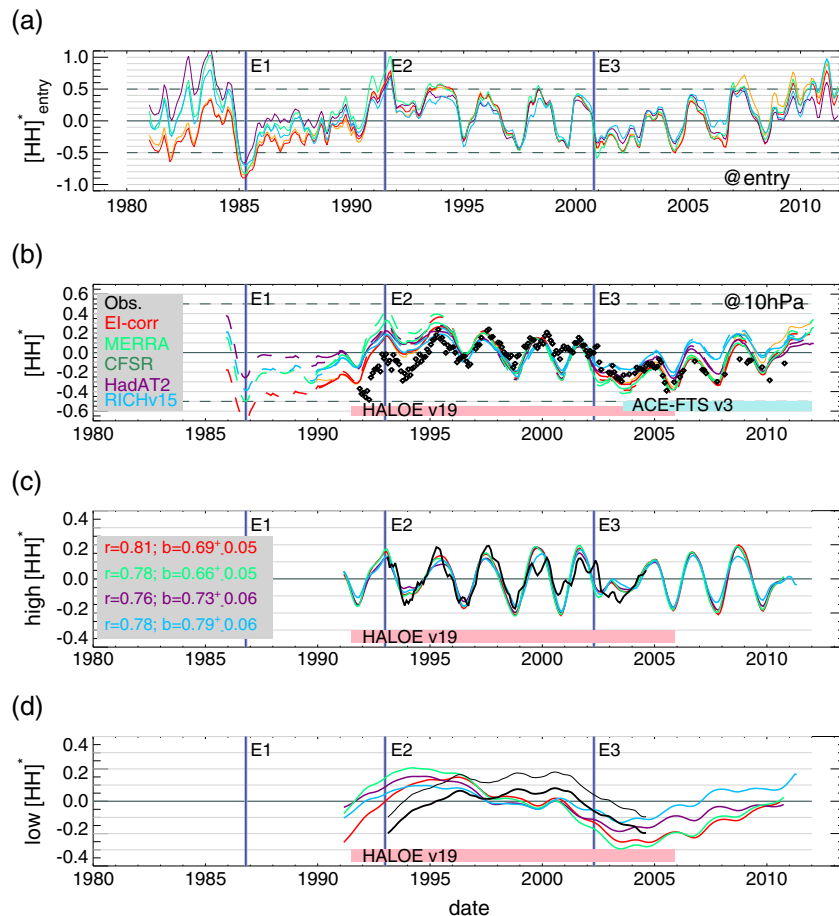


Figure 9. (a) Interannual variability of model HH entry mixing ratios (for colors, see inset in Figure 9b). (b) Interannual variability of HH observations from HALOE and ACE-FTS at 10 hPa (merged data, ACE-FTS data not continuous), and model predictions with age kernel truncated at 9 years (solid lines) and 5 years (dashed lines). (c) High-pass filtered (running mean over 30 months subtracted) interannual variability at 10 hPa. Inset: correlation coefficient and slope of linear regression (“ b ”). (d) Corresponding low-pass data. The thin black line shows the HALOE anomalies shifted to approximately match the model results for 1995/1996. For visual clarity, only a subset of model results is shown in Figures 9b–9d. Events E1, E2, and E3 as in previous figure.

of stratospheric transport from the tropopause to 10 hPa that does not vary with time, this correlation coefficient is actually very good and confirms our earlier assertion that leading-order variations in stratospheric transport can be neglected for the tropical mean at 10 hPa. The slope of the linear regression is similar to that for $[\text{H}_2\text{O}]_e$. This would be expected for the situation where stratospheric transport is captured accurately, and the sole deficiency is an error in the amplitude of variations in modeled $[\text{H}_2\text{O}]_e$. However, visual inspection of Figure 8c suggests that this agreement in amplitude mismatch may be also fortuitous inasmuch as in this case, the slopes may be affected strongly by the oscillations around the year 2000, whereas the amplitude mismatch of $[\text{H}_2\text{O}]_e$ can be observed over much of the period 1994–2010.

[45] The low-pass filtered data at 10 hPa show a similar behavior as the corresponding data at 82 hPa. The divergence between observations and model calculations at 82 hPa seen in 1994–1996 (Figure 8d) is seen at 10 hPa (Figure 9d) at the expected phase lag (i.e., around 1996/1997). The data at 10 hPa further suggest that the drift in $[\text{H}_2\text{O}]_e$ in 1994–1996 might extend further back in time (which cannot be observed at 82 hPa because of lack of reliable data for this period).

[46] The second divergence noted at 82 hPa between models and observations in the early 2000s cannot be assessed with the data at 10 hPa due to lack of continuous HALOE data beyond 2005, but we note that the difference between models and the merged HALOE/ACE-FTS data (Figure 9b) for that period is qualitatively consistent.

[47] Finally, we note that the steep increase in HH entry mixing ratios beginning in mid-1990 (i.e., before the eruption of Pinatubo) seen in the model calculations (Figure 9a) is in qualitative agreement with the evolution at 10 hPa seen in the HALOE data. However, even for the model with best agreement (EI-corr), the late 1980s are too moist (temperatures too high) relative to the early/mid-1990s. We will return to this issue, also raised by Fueglistaler [2012], in the discussion.

3.3.4. Interannual Variability of $[\text{H}_2\text{O}]$ at 10 hPa

[48] Figure 10 shows the model predictions for water vapor in the tropics at 10 hPa (discussed in section 2.3.3; assuming a contribution to water vapor from constant fraction of oxidized methane) and the corresponding observations from SAGE II and MLS/Aura. The results are generally similar to those for HH at 10 hPa. Notable differences are (i) the correlations of the high-pass filtered data are lower than for HH (as expected since the methane oxidation fraction is held constant, see section 2.1); (ii) the SAGE II data at 10 hPa

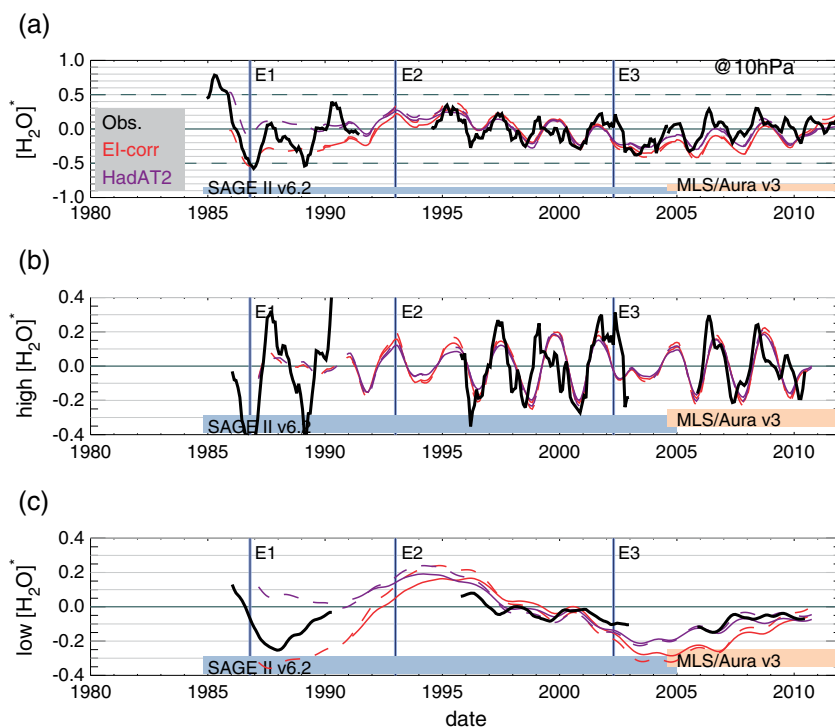


Figure 10. (a) Anomalies of SAGE II and MLS/Aura observations (black) and model predictions (colors, as labeled; for clearer display, only EI-corr and EI-HadAT2 are shown) of water vapor in the tropics at 10 hPa. Model predictions assume a constant fraction of oxidized methane (see text). Model predictions shown with dashed lines are based on age spectrum truncated at 5 years. For SAGE II, data between June 1991 and June 1994 are excluded due to artifacts from Pinatubo aerosol. (b) The high-pass filtered data (running mean over 30 months subtracted); (c) the corresponding low-pass filtered data. Note: SAGE II and MLS/Aura have insufficient overlap for reliable merging of data, and anomalies shown are with respect to different periods. A constant offset of -0.1 ppmv is applied to the low-pass filtered MLS anomalies in Figure 10c to demonstrate agreement with HadAT2-corrected model predictions. Events E1, E2, and E3 as in previous figures.

(Figure 10a) do not indicate the quiescent period in the later 1980s predicted by the model calculations (i.e., Figure 9a); and (iii) the drift around 1995/1996 noted against HALOE HH (Figure 9c) is less pronounced for SAGE II H_2O .

[49] From about 2005 onward, the low-pass filtered model predictions and (MLS/Aura) observations agree fairly well, whereas in the mid/late 1980s the SAGE II observations are about halfway between the model predictions based on the corrected ERA-Interim temperatures (EI-corr) and the HadAT2-corrected temperatures (EI-HadAT2).

3.3.5. Comparison with Boulder FPH Measurements

[50] Figure 11a shows the NOAA frostpoint hygrometer data and the model predictions for Boulder (with a contribution from methane oxidation, see Appendix B). Because of the low sampling frequency, direct comparison of observations with the model predictions is done only for a 2 year running mean of the data (the black solid line). Results using more sophisticated techniques, for example fitting of the irregularly sampled measurements with harmonics to separate

the mean annual cycle and anomalies thereof, yield very similar results and are therefore not shown.

[51] The low-pass filtering of the observations is also done separately (where available, see *Hurst et al.* [2011]) for ascending and descending data, because some measurements during ascent may be contaminated as the balloon surface sheds any accumulated water vapor or condensate, and for data averaged over the pressure range 120 to 80 hPa. The figure shows that the differences between these time series are small compared to the difference to the model predictions.

[52] Visual inspection of the figure (further substantiated in the discussion of the residuals below) suggests that there is a major mismatch between observations and models from about 1995 to 2005. We further note that the difference between ascending and descending measurements (i.e., the difference between the blue and green line) is generally small but is largest during this period of largest mismatch between NOAA-FPH measurements and model predictions.

[53] Figure 11a further shows the HALOE and MLS/Aura measurements in the vicinity of Boulder (also filtered with running mean over 2 years). Compared to the difference between model and NOAA-FPH data, the difference between model and HALOE in the 1990s–2000s is substantially smaller, which is also true for the difference to ascending or descending FPH data separately. For the period of overlap with MLS/Aura, the NOAA-FPH measurements are slightly drier but show a similar trend as the MLS/Aura measurements.

[54] For the period 1981 to 1995, the ERA-Interim model results (both uncorrected and corrected) agree better with the NOAA-FPH observations than the model predictions based on the HadAT2 temperature-corrected data. Due to the strong smoothing, the NOAA-FPH data cannot provide further information about the previously noted “quiescent” period in the model calculations, but it confirms the existence of the large oscillation around 1984–1985 (event E1). At higher levels, the NOAA-FPH data (not shown) also show the increase in the early 1980s, but is more ambiguous whether there was a decrease around 1984/1985 (see Figure 2 of *Hurst et al.* [2011]). Finally, from about 2005 onward models, NOAA-FPH and MLS/Aura observations agree reasonably. However, there is a mean offset between NOAA-FPH and EI-corr for the two periods with little drift (the 1980s and the 2000s), a point further discussed in section 3.4.

[55] Figure 11b shows, for reference, the difference in the model predictions for 82 hPa in the tropics and 100 hPa over Boulder and the corresponding difference for the HALOE observations. The figure shows that, both in models and observations, the two time series can differ over longer periods and that this difference is not well correlated between HALOE and the model estimates. While the former is expected, the latter indicates additional challenges in modeling interannual variability of extratropical water vapor, where latitudinal transport plays an important role (which renders interpretation of these data subject to additional uncertainties related to latitudinal transport and processes specific for this location.)

3.4. Residuals

[56] Figure 12 summarizes the residuals between anomalies (i.e., after subtraction of the mean annual cycle) of observations and model predictions, with Figures 12a–12d referring

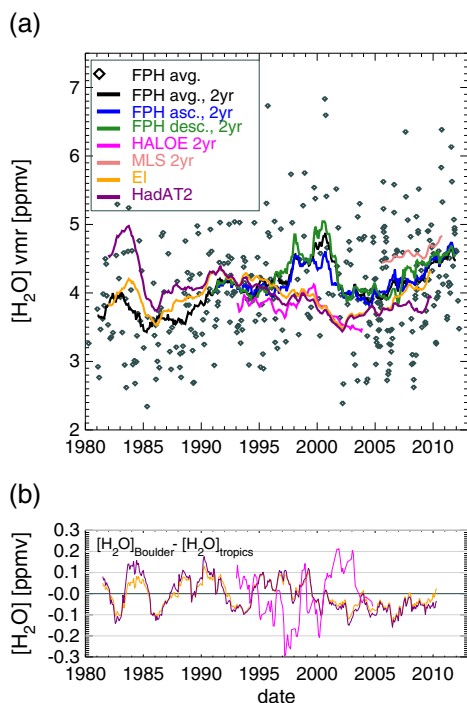


Figure 11. (a) NOAA-FPH water vapor measurements over Boulder, CO (black diamonds “FPH avg”: average of ascending and descending measurements) averaged over the pressure range 108–92 hPa (120–80 hPa data similar). Black solid line is the running mean with a window of 2 years. The blue and green lines are the 2 year running mean of ascending and descending measurements separately (where available). Also shown are HALOE and MLS data (2 year running mean) near Boulder, CO. Model calculations (2 year running mean) for Boulder (assuming 11% methane oxidation fraction, see text) are shown for “EI” and “HadAT2” (which bracket other model results). (b) The difference (smoothed with 1 year running mean) between the tropics at 82 hPa and Boulder at 100 hPa in the model predictions and in HALOE data. Colors as above (magenta, HALOE; orange, EI; purple, HadAT2).

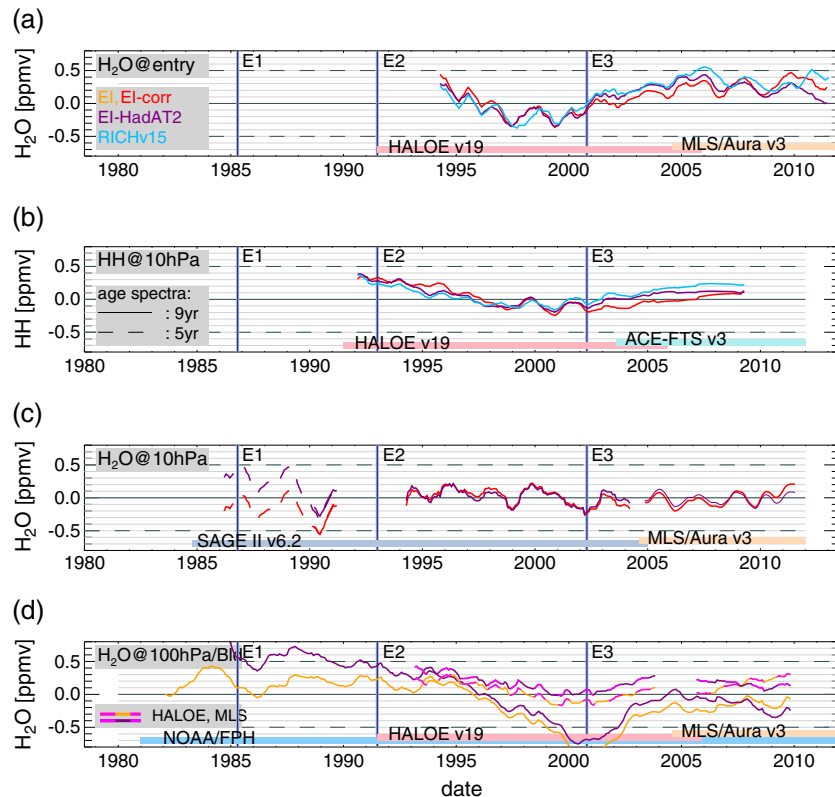


Figure 12. Residuals (model minus observation) for model calculations EI-corr, EI-HadAT2, and EI-RICHv15 (for colors, see inset of Figure 12a; EI only shown in Figure 12d). (a) $[H_2O]_e$ (deseasonalized, 6 month running mean) for merged HALOE/MLS anomalies. (b) HH at 10 hPa (deseasonalized, 6 month running mean; age kernel truncated at 9 years, see inset Figure 12b) for merged HALOE/ACE-FTS anomalies. (c) H_2O at 10 hPa (deseasonalized, 6 month running mean; age kernel truncated at 9 (solid line) and 5 years (dashed line)) for SAGE II and MLS (not merged, offset between measurements visually adjusted for smooth continuation). (d) H_2O at 100 hPa/Boulder, CO (2 year running mean, see text). Also shown are differences to HALOE and MLS near Boulder at 100 hPa (see labels in panel), offset between NOAA-FPH and HALOE residuals visually adjusted for agreement in the first years of common data. Events E1, E2, and E3 as in previous figures. Note: The time axis for Figures 12b and 12c is shifted by the mean phase lag at 10 hPa (18 months) relative to that of Figures 12a and 12d.

to the results shown in Figures 8–10, and 11. For clearer display, we only show the residuals for the model calculations that roughly bracket the range of uncertainty in the temperature record, namely, ERA-Interim and the HadAT2-corrected model predictions. Residuals that are common between all model/observation pairs may indicate the effect of a process not considered in the model calculations, whereas deficiencies specific to a pair may help to identify an error in either the specific temperature or water vapor data.

[57] Inspection of Figure 12 shows that the differences between models, between models and observations, and between observations do not follow such simple patterns and that an extended discussion is required:

- (i) Over the HALOE period (1991–2005), differences between model predictions are smaller than the difference to HALOE for water vapor at 82 hPa (Figure 12a; only data from 1994 onward shown due to previously mentioned concerns about aerosol contamination following the eruption of Mt. Pinatubo), HH (Figure 12b) and H_2O (Figure 12c) at 10 hPa, and H_2O over Boulder

(Figure 12d). For this period, all model calculations show, relative to HALOE, a negative drift (model is drying relative to the observation) in the mid-1990s, followed by a positive drift (model is moistening relative to observation), with little net bias remaining from about 2004 onward. Due to this evolution of the residual, the model predictions cannot fully reproduce the sharp, step-like decrease in entry mixing ratios in October 2000 that is so prominent in the HALOE data. Substantial differences between HALOE and NOAA-FPH data were reported by Randel *et al.* [2004] and Scherer *et al.* [2008], but it is remarkable that the residuals relative to the NOAA-FPH data (Figure 12d) over the period 1994–2004 show the same shape as the residuals to HALOE, but with much larger amplitude, and a remaining net dry bias of about 0.4 ppmv for the difference between post-2004 and pre-1995. The residual of this period relative to SAGE II H_2O (Figure 12c) shows less drift. The difference in the residual is not due to possible problems with the model's estimate of the contribution from methane oxidation to water vapor but due to the

difference between HALOE and SAGE II H₂O measurements (see Appendix, Figure A4c).

- (ii) From about 2005 onward, the model predictions show generally little drift relative to observations, and some changes in the residual can be tracked back to problems in the temperature data. For example, the uncorrected ERA-Interim model prediction produces a change in the residual at the end of 2006 when the introduction of COSMIC-GPS temperature data to the ERA-Interim assimilation system produced a temperature increase around 100 hPa (see Figures C1 and C2). Similarly, the residuals to RATPAC and HadAT2 show some changes in the residual that can be tracked back to suspicious temperature drifts at specific radiosonde stations (see Appendix C4).
- (iii) Some of the oscillations in the residual at 82 hPa are also observed at 10 hPa with the appropriate lag (roughly 1.5 years). For example, the oscillation seen around mid-1998 at 82 hPa corresponds to the oscillation seen around 1999/2000 (Figures 12a and 12b). This shows that the errors introduced by the simplification of stratospheric transport are secondary to the errors in [H₂O]_e. It also shows that the HALOE data are self-consistent at two widely spaced pressure levels, but it does not give conclusive evidence that the model [H₂O]_e is wrong since we cannot exclude here that HALOE may under- or over-represent the real variability at both levels. For the comparison with MLS/Aura data, oscillations at the two levels are less conclusive since our estimate of the contribution from methane oxidation cannot take year-to-year variability into account. It is therefore not surprising that the H₂O residuals at 82 and 10 hPa (Figures 12a and 12c) do not have much correspondence.
- (iv) For the period before mid-1991, only SAGE II and NOAA-FPH H₂O data are available. The conspicuous absence of year-to-year variability in the model predictions for the late 1980s is not confirmed by the SAGE II observations, whereas the large-amplitude oscillation in [H₂O]_e around 1984–1985 (E1) is seen also in both SAGE II and NOAA-FPH measurements. The possibility of aerosol artifacts in SAGE II water vapor measurements renders the model residuals relative to the SAGE II data inconclusive at this point (applies also to the following item).
- (v) The strong increase in water vapor mixing ratios at the beginning of the HALOE measurements noted by Fueglistaler [2012] is qualitatively reproduced by the model calculations as a consequence of a large increase in water entry mixing ratios in 1990, peaking just at about the time of the onset of the Pinatubo eruption (see event E2 in Figures 8b and 9a), but it is unclear whether the model predictions of the 1980s have a general moist bias relative to the 1990s. Relative to the SAGE II data at 10 hPa, the model calculations for the 1980s based on HadAT2 are indeed moist biased, while those based on the ERA-Interim temperatures are slightly dry biased. Compared to the NOAA-FPH measurements over Boulder, all model predictions are moist biased, and differences in particular for the earlier 1980s are largest for the HadAT2-corrected model prediction, a point we will return to below.

[58] This summary of the residuals shows that there is no single aspect of the residuals that is present in all model calculations (this is also true for the model calculations not shown in Figure 12) and all observations. Hence, it is difficult to make a strong case for a missing process in the model calculations. Our subjective assessment, supported by the analysis of the credibility of HALOE trace gas trends by Gordley *et al.* [2009], the consistency checks by Fueglistaler [2012], and the fact that the HALOE measurements in the mid-IR is less sensitive to aerosol than the SAGE II measurements in the near-IR, would tend to trust the HALOE water vapor more than that of SAGE II. As such, the drifts noted in the mid-1990s and early 2000s may reflect a real deficiency in the model calculations. Whether this deficiency reflects a missing process in the calculation for [H₂O]_e, a problem in the temperature record (although the differences between model predictions are smaller in the HALOE period than before, the observed spread in Figure 8b indicates that the temperature record in this period cannot be considered as settled), or temporal variations in transport pathways not captured by ERA-Interim remains to be seen.

[59] Similarly, the lack of global, continuous water vapor measurements without possible aerosol artifacts before, during, and after the eruption of Pinatubo in 1991 renders reliable assessment of the evolution of temperatures and water vapor over this important period very difficult. Evidently, the drifts of models against HALOE HH at 10 hPa seen in the early 1990s cannot have lasted indefinitely in the past, but whether the model-observation difference would have been constant before the early 1990s or whether they would have reverse sign (like the drift in the early 2000s) remains open due to lack of data.

4. Summary and Outlook

[60] In this paper, we have shown a comparison between observations of stratospheric water vapor and model predictions, with the objective to determine whether temperatures (and transport) in the TTL can explain stratospheric water vapor or whether there are indications for missing processes as suggested by the analyses of Kley *et al.* [2000] and Rosenlof *et al.* [2001] more than a decade ago.

[61] The model predictions use trajectories of troposphere-to-stratosphere transport based on ERA-Interim data and estimate the amount of water entering the stratosphere based on the temperature history along the trajectory. The impact of uncertainties in the temperature record on [H₂O]_e is evaluated by replacing the quasi-stationary temperature field of ERA-Interim with that of a range of alternative temperature data. As such, the model results presented here reflect the current state of art. Differences in troposphere-to-stratosphere transport between ERA-Interim and other reanalyses exist [see Liu *et al.*, 2010; Schoeberl *et al.*, 2012] but have primarily an impact on mean entry mixing ratios, not trends. It remains to be seen whether this is also true for future reanalyses.

[62] We find that the preference of the Lagrangian Dry Point to sample positions in space and time that are anomalously cold is particularly effective for submonthly temporal variability, more so than for the quasi-stationary structure of the temperature field hypothesized in earlier work to be key

to stratospheric water vapor. The model predictions for $[\text{H}_2\text{O}]_e$ are consistently dry biased for all temperature data sets. Adjusting the LDP temperature with a constant temperature offset allows to tune the models' mean annual cycle (mean and amplitude) to good agreement with observations (see also discussion in *Liu et al.* [2010]). The model calculations that employ a cloud microphysical box model for dehydration can produce the mean $[\text{H}_2\text{O}]_e$ with sensible model parameter combinations but tend to underestimate the amplitude of the annual cycle compared to both MLS and HALOE. Results based on diabatic trajectories are very similar but have a slightly larger amplitude due to a narrower LDP time distribution (a consequence of the less dispersive transport). Further discussion of this aspect is deferred to a separate publication.

[63] We have argued that the amplitudes of variations on different timescales may provide more conclusive constraints on model results than mean entry mixing ratios. We found that the model predictions tuned for the mean annual cycle tend to overestimate the amplitudes of interannual variability on timescales of 2–3 years. Whether this is also true for lower frequency variations and trends cannot be determined due to several occasions where model predictions clearly drift relative to measurements. We further noted systematic differences in the amplitude of variations in HALOE and MLS/Aura water vapor measurements, which also limit the use of arguments concerning model deficiencies based on amplitudes.

[64] Apart from a slight amplitude mismatch, the variability on timescales of 2–3 years is very well reproduced by the model calculations (with correlation coefficients of 0.9 and larger). The model calculations show that the character of these variations can substantially change on longer timescales: A large-amplitude oscillation in 1984–1985 is followed by several quiescent years. About a year before the eruption of Pinatubo, entry mixing ratios increase abruptly and subsequently show large oscillations until the marked drop in October 2000. After the drop, amplitudes are again smaller and recover in the mid 2000s.

[65] The differences between the temperature data sets used here are largest in the 1980s, whereas from about the 1990s onward the differences in model predictions of $[\text{H}_2\text{O}]_e$ due to the temperature differences are generally smaller than the differences to observations. The large differences in the temperature data of the 1980s may be related to different handling of corrections related to changes in solar heating of radiosonde temperature sensors with changes in equipment discussed in *Sherwood et al.* [2005] and remaining artifacts in “corrected” data as discussed in *Randel and Wu* [2006] and *Lanzante* [2009].

[66] The large differences in model predictions of $[\text{H}_2\text{O}]_e$ and the differences in the two measurement time series that cover the 1980s and 1990s, namely, SAGE II and NOAA-FPH, render assessment of the trend from the 1980s to the 1990s difficult. Based on the data analyzed here, the strongest conclusion that can be drawn for this period are that (i) the temperature history for the period of the 1980s to the 1990s as represented in the HadAT2 data is inconsistent with basically all water vapor data of that period (this is also true for the ATMOS version 3 data presented by *Michelsen et al.* [2000]), whereas (ii) the temperature record of ERA-Interim (which is colder in the early 1980s than any other temperature data set) yields model

predictions of $[\text{H}_2\text{O}]_e$ for the 1980s that are still slightly too moist (relative to the post 1990 data) when compared to NOAA-FPH and slightly too dry when compared to SAGE II. Compared to the HALOE data at 10 hPa (available from September 1991 onward), all model calculations of $[\text{H}_2\text{O}]_e$ suggest issues for the period late 1980s/early 1990s, but the available water vapor measurements do not allow to locate the problem more precisely.

[67] An interesting result is the failure of *all* model calculations to capture the magnitude of the drop in water vapor in October 2000 so prominent in the HALOE data. There is a cooling in that period in all temperature data sets, and this cooling has been noted before (see *Randel* [2006]; *Fueglistaler and Haynes* [2005]; *Fueglistaler* [2012]), but the amplitude is too small to produce a drop in $[\text{H}_2\text{O}]_e$ that could explain the HALOE observations. This mismatch between the model results and the HALOE observations (but not the SAGE II observations) is connected to the drifts noted in the model results relative to HALOE in the years before, and after the drop in October 2000. The question therefore arises whether (in addition to the discussed possibility of a spurious drift of HALOE, for which there are, however, at present no indications) there is a problem not directly related to $[\text{H}_2\text{O}]_e$. For example, changes in latitudinal mixing between the tropics and extratropics could result in changes in the contributions from oxidized methane to water vapor in the tropical stratosphere. Inspection of the HALOE methane time series in the tropics, however, does not show variations of the required magnitude (about 0.2 ppmv methane) in this period. We have also inspected variations in trajectory pathways and fraction of trajectories involved in mixing with the extratropics but found no obvious indications of changes linked to this event.

[68] Our work demonstrates that only very long observational time series (with several years of overlap between different instruments) may provide the necessary constraints to address the fundamental question whether interannual variability and trends in stratospheric water vapor can be understood from those in the temperature field around the tropical tropopause under consideration of transport pathways. Despite an additional decade of observations and improved reanalyses and homogenized radiosonde temperature records since the SPARC Water Vapor Report [*Kley et al.* 2000], our results show that this fundamental question remains open due to uncertainties in the temperature and water vapor observations alike. The extensive, detailed comparison between observations and model predictions presented here, however, led to the identification of specific, well-defined periods where observations and model predictions systematically diverge.

[69] Of particular interest is that the divergences occur on comparatively short timescales rather than manifest themselves as slow drifts over multidecadal timescales. Our analyses of the observational data for these periods did not reveal obvious problems. However, the discrepancies within the observational records of temperature and water vapor, respectively, are a clear indication that artifacts must exist in the observational records, whereby we have shown that for the model predictions not only errors in monthly mean temperatures must be considered but also erroneous trends in the temperature variance and in the timescale of troposphere-to-stratosphere transport. Special attention should

be paid to the periods identified here in future efforts to further scrutinize the observational records of temperature and water vapor.

Appendix A

Stratospheric Water Vapor Measurements

[70] This section discusses the merging of HALOE and MLS/Aura data (Appendix 5.1), the merging of HALOE and ACE-FTS data (Appendix 5.2), and differences in HALOE and SAGE II data (Appendix 5.3).

A1. Merging HALOE and MLS/Aura Water Vapor

[71] The measurements by HALOE and MLS have a temporal overlap of only a little more than a year, and we therefore use two different methods to estimate the required adjustment for a homogenized time series. The first method compares the climatological mean annual cycles of the two measurements. These mean annual cycles can be evaluated over the full period when good data are available (we exclude the pre-1994 HALOE data because of issues with aerosol contamination) and as such are the statistically most robust quantities that can be derived for the two instruments. However, because the two periods barely overlap, the two mean annual cycles may be different. The second method compares colocated profiles, whereby we require MLS profiles (which has the much higher sampling rate) to be within 6 h of the time of the HALOE profile, and within 15° in longitude and 5° in latitude of a HALOE profile. All MLS profiles fulfilling these criteria for a HALOE profile are averaged, resulting in one difference per HALOE profile. These differences are then gridded and averaged to monthly means. This method directly evaluates the offset of the two instruments but is statistically less accurate as there exist only a limited number of colocated profiles.

[72] Figure A1a shows that at 82 hPa both mean and amplitude of the MLS mean annual cycle for the MLS period are larger than those of HALOE for the HALOE period. Figure A1b shows that there is a good linear relation between the two mean annual cycles, with the linear regression only weakly dependent on whether an ordinary least squares (OLS) calculation or total least squares (TLS) calculation is used.

[73] Figure A1c shows that results based on the colocated data are similar to that of the first method. The MLS measurements are on average moister and have a larger amplitude of variation with time. Figure A1d shows the results of the linear regressions as done for the comparison of the mean annual cycle. The slopes of the two fits are very similar.

[74] The observed linear relation seems associated primarily with temporal variability. The inset in Figure A1d shows the relation of the colocated measurements if we subtract the means of all measurements in a given month (i.e., the figure shows the deviations due to spatial variability only). The figure shows only weak correlations, which implies that around 80 hPa the difference in the measurements of the two instruments seems to be systematic only for time variations. Given the large changes in vertical gradients due to the seasonal cycle in $[H_2O]_c$, one may suspect that the difference

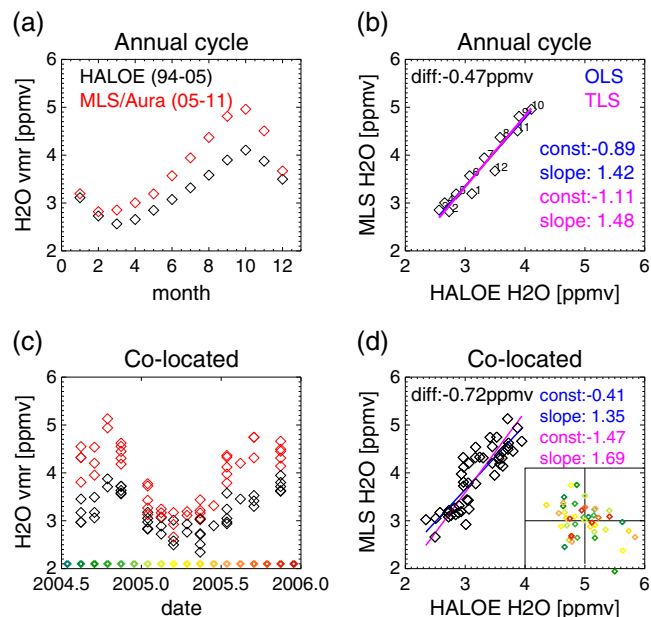


Figure A1. Relation between HALOE version 19 and MLS/Aura version 3 data in the tropics (25°S – 25°N) at 82 hPa. (a) The climatological mean annual cycle, determined for each instrument separately, and over different periods (HALOE, 1994–2005; MLS, 2005–2011). Colors as labeled in panel. (b) Scatterplot of HALOE and MLS climatological mean monthly means, with results for ordinary least square regression (OLS, blue), and total least squares linear regression (TLS, magenta). Values of fit coefficients given in panel. (c) As in Figure A1a, but for colocated profiles, with differences of colocated profiles gridded into monthly means on the latitude grid used throughout this work (width 10° , centered at 20°S , 10°S ... 20°N). (d) As in Figure A1b, but for colocated profiles. The inset centered at 5 ppmv/3 ppmv shows the same data after removal of the average of all data points within a given month (axes have same scaling). Colors refer to date of data points as indicated by color of diamonds along time axis in Figure A1c; note lack of strong correlation between MLS and HALOE or a pattern reflecting date.

could arise from the difference in vertical resolution of HALOE and MLS. However, convolving the HALOE profiles with the MLS kernel (an upper limit, as HALOE is assumed to be the true profile) does not substantially alter the amplitude of variations at 82 hPa.

[75] The lack of an explanation for the source of the different characteristics of the instruments is unsatisfactory, but the similarity of the results based on the two different methods suggests that an empirical fit can merge the two data sets. Closer inspection of Figure A1b shows that the points of the mean annual cycle form an ellipse (i.e., the deviation from the linear fit is not random but a function of the month). Consequently, merging the two data sets with a single linear fit strongly distorts the interannual anomalies.

[76] Therefore, we subtract the mean annual cycles separately of each data set, instead of a linear scaling. The true offset of the two averaging periods is then determined by the average of the differences in the interannual anomalies in the period of overlapping measurements. Note that this last step

builds on the data of the overlap period and therefore has not the same statistical accuracy as the mean annual cycle.

[77] Figure A2 shows that the mean offset (HALOE minus MLS) of the anomalies is -0.24 ± 0.21 ppmv, and the absolute offset (i.e., including the offset of the mean annual cycle) is -0.70 ± 0.21 ppmv. For comparison, the mean difference based on the colocated measurements (with each valid grid cell having equal weight) is -0.72 ppmv. Hence, the two methods yield almost identical results: The instrument offset is ~ 0.7 ppmv (MLS minus HALOE), and the true difference between the period 1995–2005 and 2005–2011 is ~ -0.2 ppmv (latter period minus former period). The difference in the mean annual cycle determined separately for each instrument and corresponding period is the sum of the instrument difference and true period difference, i.e., $\sim 0.7 + (-0.2) \sim 0.5$ ppmv.

A2. Merging HALOE and ACE-FTS

[78] Figure A3 shows the HALOE and ACE-FTS measurements of HH ($\equiv [\text{H}_2\text{O}] + 2[\text{CH}_4]$) at 10 hPa. The figure shows that the two measurements have an offset, but no further systematic differences (for example, time dependence, dependence on latitude, or difference in the amplitude of variations) are evident in the arguably few available data points. We therefore subtract the mean difference from the ACE-FTS data to obtain a homogeneous time series and take 1 standard deviation of the residual as the uncertainty in the merging of the two data sets. This uncertainty is very small compared to the mean offset and also when compared to the uncertainty in the data arising from sampling and/or random errors (see scatter of gridded data shown in Figure A3). This uncertainty is therefore omitted in the data shown in the paper.

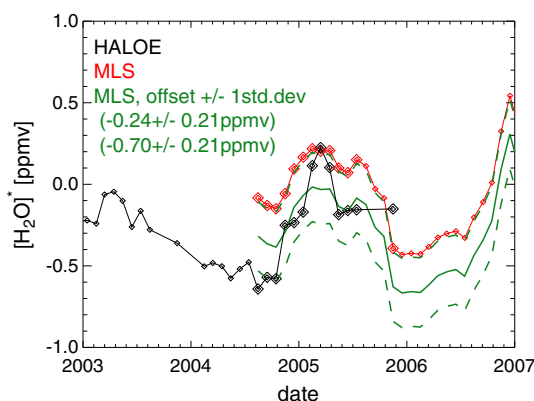


Figure A2. The anomalies after subtracting the mean annual cycles of each data set separately (black, HALOE; red, MLS). Large symbols show months where both measurements have sufficient data to form a tropical mean. The average and standard deviation of these months is taken to adjust the MLS anomalies to the HALOE anomalies (green lines; dashed lines show ± 1 std.dev.). The difference between the two periods is estimated as -0.24 ± 0.21 ppmv and the instrument offset as -0.70 ± 0.21 ppmv (HALOE minus MLS).

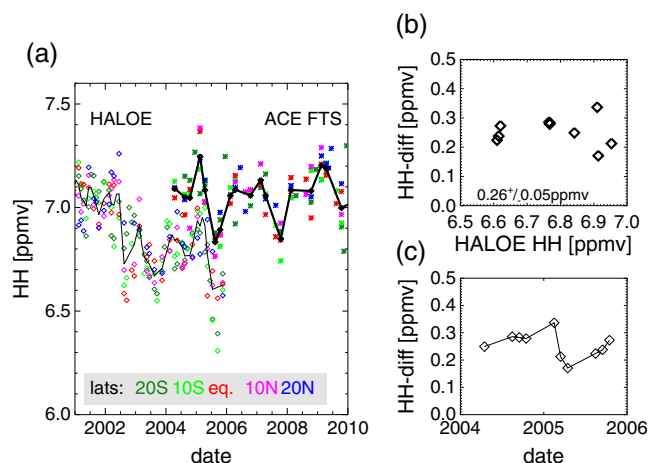


Figure A3. (a) HALOE version 19 (diamonds) and ACE-FTS version 3 (asterisks) measurements of HH ($\equiv [\text{H}_2\text{O}] + 2[\text{CH}_4]$) at 10 hPa. Profiles are binned into monthly means in latitude bins of width 10° centered at 20°S , 10°S , equator, 10°N , and 20°N (colors as indicated). Black symbols and lines (thin, HALOE; bold, ACE-FTS) are the means over the latitude bins. ACE data are linearly interpolated in log-pressure onto 10 hPa. (b) Difference in HH (ACE-FTS minus HALOE) at months where ACE-FTS has data as function of HALOE HH (HALOE linearly interpolated onto dates of ACE-FTS data). The average difference between HALOE and ACE-FTS HH at 10 hPa is 0.26 ± 0.05 ppmv. (c) Difference in HH as function of date.

A3. SAGE II Data

[79] We use the SAGE II version 6.2 water vapor data as described in Thomason *et al.* [2004] and Taha *et al.* [2004]. This version is a massive improvement over prior version, but there remains some concern (as with HALOE) that the SAGE II water data, which are based on an extinction measurement, may be affected by the presence of stratospheric aerosol. In this paper, we only use profiles that satisfy the quality requirements suggested by Taha *et al.* [2004] (error no larger than 50% and extinction at 1020 nm no larger than $2 \times 10^{-4} \text{ km}^{-1}$). Figure A4a shows the anomalies of tropical mean water vapor mixing ratio, and Figure A4b shows the difference between SAGE II and HALOE. Also shown are the aerosol surface areas densities derived from the SAGE II extinction measurements (black contour lines). The figure shows anomalously large anomalies in the wake of the eruption of Mt. Pinatubo. This period also clearly stands out in the comparison with HALOE. However, it is not clear from this figure whether there is a correlation between elevated aerosol loading and a positive bias in water vapor also at moderate aerosol levels. Figure A4 shows that the stratospheric aerosol loading was higher in the period before the eruption of Mt. Pinatubo than in the years after about 1995. Therefore, it cannot be excluded (but is also not proof) that the oscillations seen in SAGE II water vapor in the period of quiescent model $[\text{H}_2\text{O}]_c$ are aerosol artifacts.

[80] Finally, Figure A4c shows the difference between HALOE and SAGE II tropical water vapor at 10 hPa. The figure shows for the period 1995–2000 oscillations

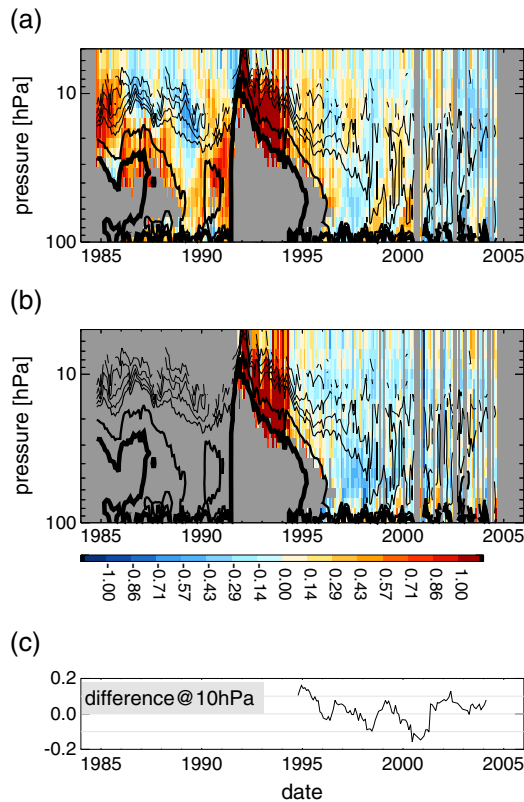


Figure A4. (a) Anomalies (relative to period 1995–2004) of tropical water vapor mixing ratios from SAGE II v6.2 (color coded), and aerosol surface area density (black lines corresponding to $[0.1, 0.2, 0.3, 0.4, 0.5, 1, 2] \mu\text{m}^2/\text{cm}^3$; with thicker lines for the higher values). (b) The difference SAGE II minus HALOE v19 water vapor mixing ratios, and aerosol surface area density as in Figure A4a. (c) The difference SAGE II minus HALOE water vapor mixing ratios at 10 hPa (smoothed with 6 month running mean, with arbitrary offset; only data from June 1994 onward shown).

superposed on a steady decrease. It is this steady decrease that produces the difference in the residuals (Figure 12), with SAGE II (and the model results) steadily getting drier relative to HALOE over the period 1995–2000, followed by a smaller “drop” in October.

Appendix B

The Contributions from Methane Oxidation to Stratospheric Water Vapor

[81] For the water vapor measurements where no corresponding methane data are available, we estimate the contribution from oxidized methane based on the long-term average of the fraction of oxidized methane at the corresponding location (i.e., in the tropics at 82 and 10 hPa and over Boulder at 100 hPa).

[82] For the tropics at 82 hPa, we estimate the contribution to SWV from the difference between the tropospheric methane time series and the (monthly averaged) ACE-FTS observations (which may be more reliable than HALOE at this location) in the tropics at that pressure level (with a

2 month time lag; results not sensitive to lag). For the period 2004–2009, we find a difference of 0.041 ± 0.014 ppmv, corresponding to a contribution of 0.082 ± 0.028 ppmv to SWV (data after 2010 are even lower; at present, it is not clear whether this is a sampling issue, and we therefore rely on the period 2004–2010). In situ methane measurements [e.g., Tuck *et al.*, 2003] give similar values. This mean contribution and variability is much smaller than all other terms and is therefore neglected for the comparisons of SWV at 82 hPa.

[83] For the tropics at 10 hPa, this fraction is determined based on the methane measurements by HALOE (here we choose HALOE over ACE-FTS for its longer time series) and a model estimate of methane at 10 hPa without chemical sink based on the convolution of the methane entry mixing time series (section 2.3.2) with the idealized age spectrum for this region (section 2.3.3).

[84] Figure B1a shows the tropical methane mixing ratios at 10 hPa measured by HALOE and ACE-FTS, the methane entry mixing ratios, and the model prediction for 10 hPa

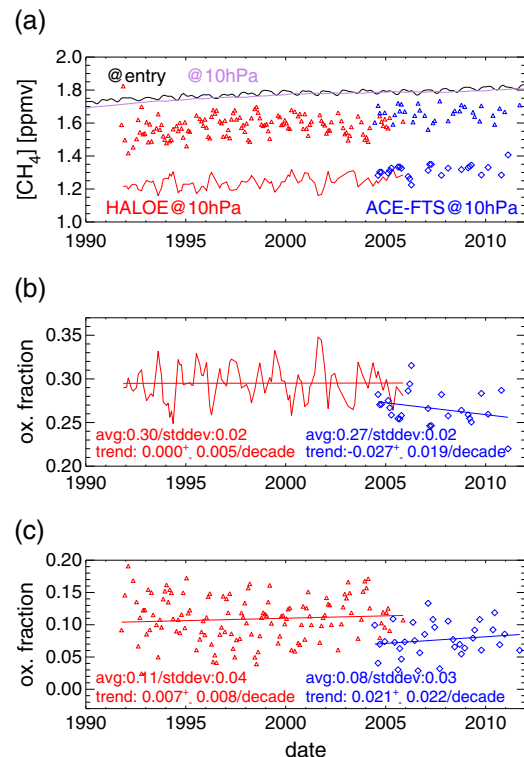


Figure B1. (a) Tropical mean methane mixing ratios at 10 hPa of the model prediction (purple line, based on convolution of the methane entry mixing ratio (black line) with the idealized age spectrum, ignoring chemical losses) and as measured by HALOE (red line) and ACE-FTS (blue diamonds). The red and blue triangles show the HALOE and ACE-FTS methane measurements near Boulder, CO at 100 hPa. (b) The fraction of oxidized methane (i.e., $1 - [\text{methane measured at 10 hPa}] / [\text{methane expected without loss}]$) for HALOE (red line) and ACE-FTS (blue line). For both data sets, the standard deviation and linear trend with 1-sigma uncertainty are given. (c) As in Figure B1b but for measurements near Boulder, CO at 100 hPa. See text for discussion of possible biases.

without chemical loss. Figure B1b shows the fraction of oxidized methane (i.e., $1 - [\text{methane measured at 10 hPa} / \text{methane expected without loss}]$). The figure shows that this ratio is very stable for the HALOE measurements; the oscillations arise from transient variations in the age of air spectrum at this location (which are not captured with a constant age of air spectrum). The ACE-FTS methane mixing ratios at this location are slightly higher than those of HALOE, and correspondingly the fraction of oxidized methane is slightly smaller. Also, there is a small trend in the estimated fraction of oxidized methane for ACE-FTS. Given the relatively short period of ACE-FTS measurements and the relatively poor sampling of the tropics, we rely for this paper on the results from HALOE (with a mean fraction of oxidized methane of 30%) but recommend further attention to the behavior of the methane record from ACE-FTS.

[85] Figure B1c shows the fraction of oxidized methane estimated for Boulder, CO at 100 hPa, whereby we directly use the methane entry mixing ratio time series (the typical phase lag at Boulder is a few months and may be ignored for this calculation). As in the case for 10 hPa in the tropics, the fraction of oxidized methane is fairly stable over the HALOE period but shows a weakly positive trend for the ACE-FTS data.

[86] For both cases, the fraction of oxidized methane is lower for the ACE-FTS data than for the HALOE data (0.27 versus 0.30 at 10 hPa in the tropics, and 0.08 versus 0.11 at 100 hPa over Boulder). HALOE methane is low biased also against other measurements [see *Rohs et al.*, 2006], and correspondingly the estimates of the fraction of oxidized methane are likely overestimating the contribution from methane. However, the impact of this error on water vapor trends is small compared to other errors. Figure B1 suggests that the oxidized fraction is no more than 5% too large for both the tropics at 10 hPa and Boulder at 100 hPa. With a methane trend from the 1980s to the 2000s of about 0.1 ppmv/decade, this gives an error in the water vapor trend of ± 0.01 ppmv/decade (i.e., $2 \cdot 0.05 \cdot 0.1$ ppmv/decade).

Appendix C

Temperature Data

C1. Tropical Mean Temperature Differences

[87] Figure C1 shows the mean annual cycle and the interannual anomalies in the differences for the tropical average temperature at 70, 100, and 150 hPa, complementing the information given in Figures 3b and 4. The figure shows that absolute differences are largest at 100 hPa, where ERA-Interim is about 0.3 K colder than the input radiosonde data, and about 1.1 K colder than CFSR. Similarly, the interannual variability in the temperature differences (indicating drifts) is largest also for CFSR at 100 hPa. A large drift between CFSR and the other data sets is also seen at 70 hPa around 1987. The smallest differences are the ERA-Interim error statistics, which implies that ERA-Interim temperatures follow closely the input radiosonde temperatures.

[88] As discussed in the main text, the large increase in temperature data assimilated from COSMIC-GPS measurements in 2006 leads to a shift in ERA-Interim temperatures,

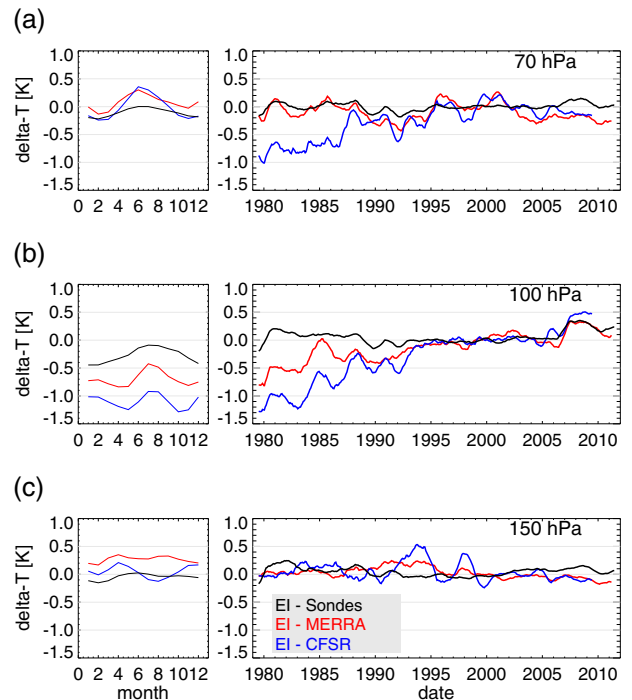


Figure C1. The differences in the tropical (30°S – 30°N) mean (period 1995–2005) annual cycle (left panels) and anomalies thereof (right panels, smoothed with 12 month running mean) between ERA-Interim and the ECMWF assimilated input radiosonde temperatures (black), MERRA (red), and CFSR (blue) for the pressure levels of (a) 70 hPa, (b) 100 hPa, and (c) 150 hPa. Note: figure shows Interim minus other data; negative values imply Interim being colder than the other data set.

which is also seen in Figure C1. In the following, we discuss this shift and implications for the absolute temperature error in more detail.

C2. Comparison with COSMIC-GPS, CHAMP-GPS, and the Absolute Bias of ERA-Interim

[89] Figure C1 shows ERA-Interim at 100 hPa to be colder than the input radiosonde data by about 0.3 K up to 2006, whereas the difference to the other reanalyses is larger. This implies that either the other reanalyses are warm biased or the ERA-Interim error statistics underestimate the bias. Such an underestimation is likely since the error statistics is based on a comparison to input data to the assimilation system, at the positions of the input data.

[90] In order to better estimate the mean temperature bias profile of ERA-Interim, we compare the ERA-Interim temperatures also to temperature data from COSMIC-GPS [Rocken et al., 2000], CHAMP-GPS [Wickert et al., 2001], and HIRDLS [Gille et al., 2008]. We find that HIRDLS is warm biased against all other data sets around 100 hPa. At 95 hPa between 10°S and 10°N , we find that the bias is largest and is in the range of +2.5 to +3 K compared to CHAMP-GPS data and about +3.5 K compared to ERA-Interim (due to ERA-Interim being cold biased in that period). These numbers are similar to values published

previously [e.g., Gille *et al.*, 2008] and are too large for HIRDLS temperature data to be of relevance for this study.

[91] Figure C2 shows that for the period 2007–2011, ERA-Interim in the inner tropics (most relevant for stratospheric water vapor) is cold biased against COSMIC-GPS by -0.42 ± 0.12 K at 100 hPa. For the same period, the ERA-Interim error statistic (for all stations within 30°S – 30°N) indicates almost zero bias. Combining the change in bias in 2006 in the ERA-Interim error statistics (a warming of Interim of about 0.3 K) with the absolute offset against COSMIC-GPS for the post-2006 period (the -0.4 K above) gives a cold bias of Interim for the pre-2006 period of around -0.7 K, which is similar to the bias against MERRA (Figure C1b). For the 70 and 150 hPa levels, the biases are smaller, and ERA-Interim shows less drift with onset of assimilation of the COSMIC-GPS temperature data.

[92] The shift in 2006 in ERA-Interim may be also estimated for the full area average (rather than station average as in the error statistics) using the CHAMP-GPS data as reference for the period 2002–2007. The sampling frequency of CHAMP-GPS is much lower than that of COSMIC-GPS, and we therefore only compare annual means (with the year 2006 excluded due to a longer data gap).

[93] Figure C2 shows that the absolute offset to ERA-Interim is similar for the CHAMP data and the COSMIC data in 2007, and the change in 2006 based on the difference to the CHAMP-GPS data is similar to that seen in the Interim error statistics. With only 1 year of data for the post-2006 period

and the differences for 2002–2005 varying by ± 0.1 K, the slightly larger difference in the CHAMP-GPS data at 100 hPa for 10°S – 10°N are within the statistical uncertainty of the difference seen in the Interim error statistics. Hence, our best guess for the tropical mean profile of ERA-Interim temperature bias is based on the absolute difference to the COSMIC-GPS measurements over the period 2007–2011, and we use the ERA-Interim error statistic for the temporal evolution of the tropical temperature bias.

C3. The Spatial Structure near the Tropopause

[94] Figure C3 shows the climatological mean (period 1995–2005) temperature structure at 100 hPa in January and July for ERA-Interim and the differences in the pattern to MERRA and CFSR. In order to emphasize the spatial pattern, the area mean of each data set (discussed in the previous section) has been subtracted. The differences in the spatial pattern are of order 1 K but can be up to 2 K locally during boreal summer, when the three reanalyses represent the temperature structure in particular south of the Indian/Southeast Asian monsoon region differently.

[95] For a reliable estimate of $[\text{H}_2\text{O}]_e$, not only has the “tropical average” temperature to be correct, but also the three-dimensional spatial structure. The differences seen in Figure C3 are large enough to be of concern, and we therefore substitute the full quasi-stationary temperature where possible (i.e., for the other reanalyses).

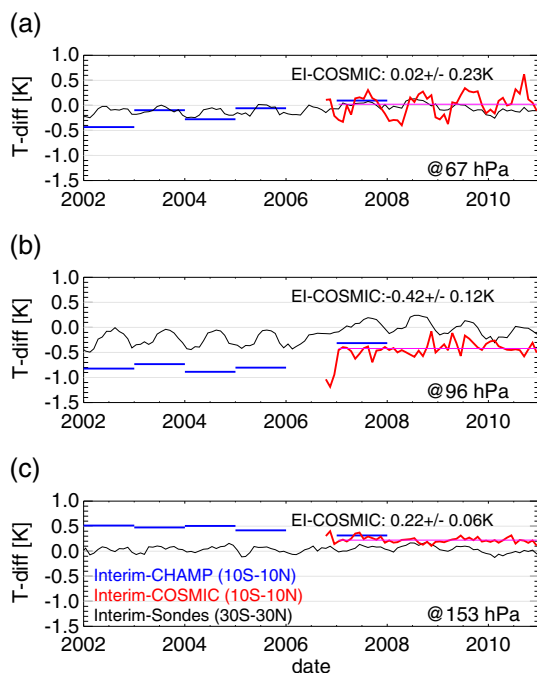


Figure C2. Differences between COSMIC and ERA-Interim monthly mean inner tropical mean (10°S – 10°N) temperatures (red); ditto, but for CHAMP and ERA-Interim annual mean temperatures (blue; no data for 2006 due to larger period without data); ERA-Interim internal error statistics for 30°S – 30°N (same data as shown in Figure C1). (a) 67 hPa; (b) 96 hPa; (c) 153 hPa. The means and standard deviations of the period 2007–2011 for the difference between Interim and COSMIC are given in the panels.

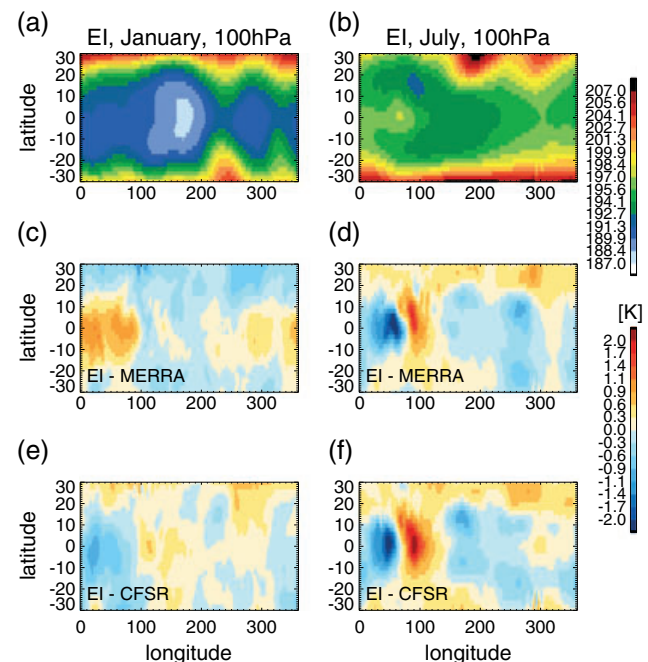


Figure C3. (a, b) Climatological mean (1995–2005) January and July temperatures (in Kelvin, upper color bar) at 100 hPa of ERA-Interim. (c, d) Differences to MERRA (Interim minus MERRA, in Kelvin, lower color bar); (e, f) differences to CFSR (Interim minus CFSR). For the difference plots, the area mean has been subtracted from each data set to emphasize the spatial structure of the differences.

[96] Analysis of the temperature structure differences at the pressure levels of 150 and 70 hPa (not shown) shows that the differences are smaller. Moreover, we found that the changes in the spatial structure over time play a secondary role compared to the changes in the area mean, and to first order the differences in $[\text{H}_2\text{O}]_e$ based on Interim, and the model calculations with substituted quasi-stationary temperature fields can be recovered with a tropical average temperature correction. This lack of different trends in the spatial temperature structure is important as it implies that the model calculations that have only tropical-mean corrections are not much inferior to those based on the reanalyses. This does not, however, preclude that all temperature corrections miss important aspects of the spatial structure, as the reanalyses around 100 hPa are also constrained primarily by the inhomogeneously distributed radiosonde data.

C4. The Temperature Correction for HadAT2 and the Impact of Spurious Temperature Drifts at Individual Stations

[97] The number of tropical stations in the HadAT2 data set allows a coarse resolution of the zonal structure of the temperature difference to ERA-Interim. The correction applied to ERA-Interim is based on the (monthly mean, smoothed with 12 month running mean) temperature difference of all stations within 15°S – 15°N , with the zonal structure calculated with a convolution with a Gaussian of width 10° longitude (ignoring the latitude of the stations). For wider widths of the Gaussian, the correction approaches the plain station average; varying the width by a factor of 2 or 3 has negligible impact on results (not shown). The correction is applied uniformly between 15°S and 15°N , with linear relaxation to no correction poleward of 30°S and 30°N .

[98] Figure C4 shows the temperature correction for 100 hPa, the pressure level where the differences are largest.

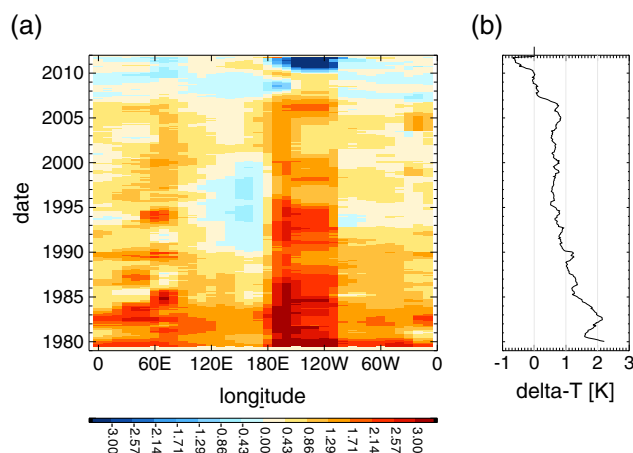


Figure C4. (a) The temperature correction for the HadAT2-corrected temperature data at 100 hPa, based on the difference HadAT2 minus ERA-Interim at the locations of stations within 15°S – 15°N smoothed in time with 12 month running mean and convolved in the zonal direction with a Gaussian of width 10° . Note that sign is opposite to temperature difference plots. (b) The corresponding zonal mean.

With the focus on $[\text{H}_2\text{O}]_e$ in this paper, we cannot comment on the interesting evolution of the temperature difference in detail, except for two aspects with immediate relevance to $[\text{H}_2\text{O}]_e$.

[99] First, we note the change in the difference over the Western Pacific (roughly 120°E – 180°E) in the 1990s relative to the overall decrease in the difference. This is evidently of interest in the context of the models' residuals in the 1990s, although the temperature differences are too small to have an influence on $[\text{H}_2\text{O}]_e$ from ERA-Interim and HadAT2 that would render one model calculation clearly superior for this period.

[100] Second, we note a strong decrease in the correction (i.e., a cooling of HadAT2 relative to Interim) over the Eastern Pacific around the year 2010. Inspection of the original data shows that this is due to the evolution at the only two stations in the Eastern Pacific, namely Pago Pago (WMO ID 91765) and Atuona (WMO ID 91925). In 2008, the temperature difference shows a large oscillation over Pago Pago but not Atuona (leading to the small drop in 2008 around 170°E seen in Figure C4a), whereas in 2010 both Pago Pago and Atuona drop several Kelvin relative to ERA-Interim. The same pattern is seen in the difference to MERRA (not shown), indicating that the problem may be the HadAT2 data rather than ERA-Interim. With the only two stations in the Eastern Pacific showing both the same behavior, the drifts at these stations have a substantial impact on $[\text{H}_2\text{O}]_e$ and account for the decrease in $[\text{H}_2\text{O}]_e$ of the HadAT2-based model calculation relative to all other model calculations around 2010 (see Figures 8, 10a, and 12a).

[101] **Acknowledgments.** We would like to thank three reviewers for their helpful comments and thank the following groups and agencies for providing data and support for this study. HALOE data were downloaded from <http://haloe.gats-inc.com/home/index.php>. The MLS data used herein were produced at the Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautics and Space Administration. The MLS data are available at <http://mirador.gsfc.nasa.gov>. The Atmospheric Chemistry Experiment (ACE), also known as SCISAT, is a Canadian-led mission mainly supported by the Canadian Space Agency and the Natural Sciences and Engineering Research Council of Canada. We thank the NASA Langley Research Center (NASA-LaRC) and the NASA Langley Radiation and Aerosols Branch for providing the SAGE II data, available at http://eosweb.larc.nasa.gov/PRODOCS/sage2/table_sage2.html. We thank NOAA/ESRL for tropospheric Methane measurements, available at <http://www.esrl.noaa.gov/gmd/obop/mlo/programs/esrl/methane/methane.html>. We thank ECMWF for providing the ERA-Interim data. MERRA data used in this study/project have been provided by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center through the NASA GES DISC online archive. We thank NCEP and NCAR for providing the CFSR data, available at <http://dss.ucar.edu/pub/cfsr.html>, and for the NCEP/NCAR reanalysis available at <http://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml>. We thank NOAA for providing the RATPAC data, available at <http://www.ncdc.noaa.gov/oa/climate/ratpac>. RICHv1.5 data were downloaded from <http://www.univie.ac.at/theoret-met/research/raobcore>. CHAMP and COSMIC-GPS temperature data were downloaded from <http://cdaac-www.cosmic.ucar.edu/cdaac/postProcess.html>. HadAT2 temperature data were downloaded from www.metoffice.gov.uk/hadobs. During the initial stages of this work, S.F. has been supported by a NERC Advanced Researcher Fellowship at Cambridge University, and Y.S.L. has been supported by a NERC PhD studentship at Cambridge University as part of the ACTIVE consortium project. T.J.F. has been supported by a NERC PhD studentship, with additional support from the EU-funded SHIVA project.

References

- Bates, D.R., M. Nicolet (1950), The photochemistry of atmospheric water vapor, *J. Geophys. Res.*, **55**, 301–327.
- Bernath, P. F., et al. (2005), Atmospheric Chemistry Experiment (ACE): Mission overview, *Geophys. Res. Lett.*, **32**, L15S01, doi:10.1029/2005GL022386.

- Brewer, A.W. (1949), Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, *Q.J.R. Meteorol. Soc.*, **75**, 351–363.
- Brasseur, G.P., S. Solomon (2005), *Aeronomy of the Middle Atmosphere*, 3rd edition, Springer, Dordrecht.
- Dee, D.P., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, **137**, 553–597.
- Dessler, A.E. (1998), A reexamination of the ‘stratospheric fountain’ hypothesis, *Geophys. Res. Letts.*, **25** (22), 4165–4168.
- Dlugokencky, E.J., L.P. Steele, P.M. Lang, K.A. Masarie (1994), The growth rate and distribution of atmospheric methane, *J. Geophys. Res.*, **99** (D8), 17021–17043.
- Dlugokencky, E.J., L.P. Steele, P.M. Lang, K.A. Masarie (1995), Atmospheric methane at Mauna Loa and Barrow observatories: Presentation and analysis of in-situ measurements, *J. Geophys. Res.*, **100**, 23103–23113.
- Etheridge, D.M., G.I. Pearman, P.J. Fraser (1992), Changes in tropospheric methane between 1841 and 1978 from a high accumulation-rate Antarctic ice core, *Tellus B-Chem. Phys. Meteorol.*, **44** (4), 282–294, doi:10.1034/j.1600-0889.1992.t01-3-00006.x
- Forster, de P.M. F., K.P. Shine (1999), Stratospheric water vapour changes as a possible contributor to observed stratospheric cooling, *Geophys. Res. Letts.*, **26** (21), 3309–3312.
- Free, M., D.J. Seidel, J.K. Angell, J. Lanzante, I. Durre, T.C. Peterson (2005), Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series, *J. Geophys. Res.*, **110** (D22), D22101.
- Fueglistaler, S. (2012), Step-wise changes in stratospheric water vapor? *J. Geophys. Res.*, **117**, D13302, doi:10.1029/2012JD017582.
- Fueglistaler, S., P.H. Haynes (2005), Control of interannual and longer-term variability of stratospheric water vapor, *J. Geophys. Res.*, **110** (D24), D24108, doi:10.1029/2005JD006019.
- Fueglistaler and Baker (2006), A modelling study of the impact of cirrus clouds on the moisture budget of the upper troposphere, *Atmos. Chem. Phys.*, **6**, 1425–1434.
- Fueglistaler, S., M. Bonazzola, P.H. Haynes, T. Peter (2005), Stratospheric water vapor predicted from the Lagrangian temperature history of air entering the stratosphere in the tropics, *J. Geophys. Res.*, **110**, D08107, doi:10.1029/2004JD005516.
- Fueglistaler, S., A.E. Dessler, T.J. Dunkerton, I. Folkins, Q. Fu, P.W. Mote (2009a), Tropical tropopause layer, *Rev. Geophys.*, **47**, RG1004, doi:10.1029/2008RG000267.
- Fueglistaler, S., B. Legras, A. Beljaars, J.-J. Morcrette, A. Simmons, A.M. Tompkins, S. Uppala (2009b), The diabatic heat budget of the upper troposphere and lower/mid stratosphere in ECMWF reanalyses, *Q. J. R. Meteorol. Soc.*, **135**, 21–37.
- Gille, J. C., et al. (2008), The High Resolution Dynamics Limb Sounder (HIRDLS): Experiment overview, recovery, and validation of initial temperature data, *J. Geophys. Res.*, **113**, D16S43, doi:10.1029/2007JD008824.
- Gordley, L., E. Thompson, M. McHugh, E. Remsburg, J. Russell III, and B. Magill (2009), Accuracy of atmospheric trends inferred from the Halogen Occultation Experiment data, *J. Appl. Remote Sens.*, **3**, 033526, doi:10.1117/1.3131722.
- Haimberger, L., C. Tavaloto, S. Sperka (2012), Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations, *J. Clim.*, **25** (23), 8108–8131, 2012.
- Holton, J.R., A. Gettelman (2001), Horizontal transport and the dehydration of the stratosphere, *Geophys. Res. Lett.*, **28**, 2799–2802.
- Hurst, D. F., S. J. Oltmans, R. Voemel, K. H. Rosenlof, S. M. Davis, E. A. Ray, E.G. Hall, A.F. Jordan (2011), Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record *J. Geophys. Res.*, **116**, D02306, doi:10.1029/2010JD015065
- Jones, R.L., J.A. Pyle (1984), Observations of CH₄ and N₂O by the NIMBUS 7 SAMS: A comparison with in situ data and two-dimensional numerical model calculations, *J. Geophys. Res.*, **89**, 5263–5279.
- Kley, D., et al. (2000), SPARC assessment of upper tropospheric and stratospheric water vapour, *WCRP 113, WMO/TD No. 1043*, SPARC Rep. No. 2, World Meteorol. Organ., Geneva.
- Lanzante, J.R. (2009), Comment on ‘Trends in the temperature and water vapor content of the tropical lower stratosphere: Sea surface connection’ by Karen Rosenlof and George C. Reid, *J. Geophys. Res.*, **114**, D12104, doi:10.1029/2008JD010541.
- Le Texier, H., S. Solomon, R. Garcia (1988), The role of molecular hydrogen and methane oxidation in the water vapor budget of the stratosphere, *Q. J. R. Meteorol. Soc.*, **114**, 281–295.
- Liu, Y.S., S. Fueglistaler, P.H. Haynes (2010), Advection-condensation paradigm for stratospheric water vapor, *J. Geophys. Res.*, **115**, D24307.
- Michelsen, H.A., F.W. Irion, G.L. Manney, G.C. Toon, M.R. Gunson (2000), Features and trends in Atmospheric Trace Molecule Spectroscopy (ATMOS) version 3 stratospheric water vapor and methane measurements, *J. Geophys. Res.*, **105**, 22713–22724.
- Mote, P.W., K.H. Rosenlof, M.E. McIntyre, E.S. Carr, J.C. Gille, J.R. Holton, J.S. Kinnarsley, H.C. Pumphrey (1996), An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, *J. Geophys. Res.*, **101**, 3989–4006.
- Newell, R.E., S. Gould-Stewart (1981), A stratospheric fountain?, *J. Atmos. Sci.*, **38**, 2789–2796.
- Notholt, J., et al. (2005), Influence of tropospheric SO₂ emissions on particle formation and the stratospheric humidity, *Geophys. Res. Lett.*, **32**, doi:2004GL022159.
- Randel, W.J., F. Wu (2006), Biases in stratospheric and tropospheric temperature trends derived from radiosonde data, *J. Clim.*, **19**, 2094–2104.
- Randel, W.J., F. Wu, S.J. Oltmans, K. Rosenlof, G.E. Nedoluha (2004), Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, **61**, 2133–2148.
- Randel, W. J., et al. (2006), Decreases in stratospheric water vapor since 2001: links to changes in the tropical tropopause and the Brewer-Dobson circulation, *J. Geophys. Res.*, **111** (D12), D12312, doi:10.1029/2005JD006744.
- Read, W.G., et al. (2007), Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H₂O and relative humidity with respect to ice validation, *J. Geophys. Res.*, **112**, D24S35, doi:10.1029/2007JD008752.
- Rienecker, M., et al. (2011), MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1
- Rocken, C., Y.-H. Kuo, W. Schreiner, D. Hunt, S. Sokolovskiy (2000), COSMIC system description, *Terrestrial Atmospheric And Oceanic Sciences*, **11** (1), 21–52.
- Rohs, S., C. Schiller, M. Riese, A. Engel, U. Schmidt, T. Wetter, I. Levin, T. Nakasawa, S. Aoki (2006), Long-term changes of methane and hydrogen in the stratosphere in the period 1978–2003 and their impact on the abundance of stratospheric water vapor, *J. Geophys. Res.*, **111**, D14315, doi:10.1029/2005JD006877.
- Rosenlof, K.H., et al. (2001), Stratospheric water vapor increases over the past half-century, *Geophys. Res. Lett.*, **28**, 1195–1198.
- Russell, J.M., III, A.F. Tuck, L.L. Gordley, J.H. Park, S.R. Drayson, J.E. Harries, R.J. Cicerone, and P.J. Crutzen (1993), The Halogen Occultation Experiment, *J. Geophys. Res.*, **98**, 10777–10797.
- Saha, S., et al. (2010), The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.*, **91** (8), 1015–1057, doi:10.1175/2010BAMS3001.1.
- Scherer, M., H. Vömel, S. Fueglistaler, S.J. Oltmans, and J. Staehelin (2008), Trends and variability of midlatitude stratospheric water vapour deduced from the re-evaluated Boulder balloon series and HALOE, *Atmos. Chem. Phys.*, **8**, 1391–1402.
- Schoeberl, M.R., A.E. Dessler, T. Wang (2012), Simulations of stratospheric water vapor and trends using three reanalyses, *Atmos. Chem. Phys.*, **12**, 6475–6487.
- Sherwood, S.C. (2002), A microphysical connection among biomass burning, cumulus clouds, and stratospheric moisture, *Science*, **295**, 1272–1275.
- Sherwood, S.C., and A.E. Dessler (2000), On the control of stratospheric humidity, *Geophys. Res. Lett.*, **27**, 2513–2516.
- Sherwood, S.C., J.R. Lanzante, C.L. Meyer (2005), Radiosonde daytime biases and late-20th century warming, *Science*, **309**, 1556–1559.
- Solomon, S., K.H. Rosenlof, R.W. Portmann, J.S. Daniel, S. M. Davis, T.J. Sanford, G.-K. Plattner (2010), Contributions of stratospheric water vapor to decadal changes in the rate of global warming, *Science*, **327**, 1219–1222, doi:10.1126/science.1182488.
- Taha, G., L.W. Tomason, S.P. Burton (2004), Comparison of Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 water vapor with balloon-borne and space-based instruments, *J. Geophys. Res.*, **109**, D18313, doi:10.1029/2004JD004859.
- Thomason, L.W., S.P. Burton, N. Iyer, J. M. Zawodny, and J. Anderson (2004), A revised water vapor product for the Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 data set, *J. Geophys. Res.*, **109**, D06312, doi:10.1029/2003JD004465.
- Thorne, P. W., Parker, D. E., Tett, S. F. B., Jones, P. D., McCarthy, M., Coleman, H., and Brohan, P. (2005), Revisiting radiosonde upper-air temperatures from 1958 to 2002, *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.
- Tuck, A.F., S.J. Hovde, K.K. Kelly, M.J. Mahoney, M.H. Proffitt, E.C. Richard, T.L. Thompson (2003), Exchange between the upper tropical troposphere and the lower stratosphere studied with aircraft observations, *J. Geophys. Res.*, **108**, D23, 4734, doi:10.1029/2003JD003399.

- Wang, J.S., D.J. Seidel, M. Free (2012), How well do we know recent climate trends at the tropical tropopause, *J. Geophys. Res.*, *117*, D09118, doi:10.1029/JD017444.
- Waugh, D.W., T.M. Hall (2002). Age of stratospheric air: Theory, observations, and models, *Rev. Geophys.* *40* (4), 1010, doi:10.1029/2000RG000101.
- Wickert, J., et al. (2001), Atmosphere sounding by GPS radio occultation: First results from CHAMP. *Geophys. Res. Lett.*, *28*, 3263–3266, doi:10.1029/2001GL013117.