Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere – Part 2: Data analysis and calibration for long-term monitoring

T. Leblanc, I. S. McDermid, and T. D. Walsh

Table Mountain Facility, Jet Propulsion Laboratory, California Institute of Technology, Wrightwood, CA 92397, USA

Received: 18 May 2011 – Accepted: 5 August 2011 – Published: 10 August 2011

Correspondence to: T. Leblanc (leblanc@tmf.jpl.nasa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The well-recognized, key role of water vapor in the upper troposphere and lower stratosphere (UT/LS) and the scarcity of high-quality, long-term measurements triggered the development by JPL of a powerful Raman lidar to try to meet these needs. This development started in 2005 and was endorsed by the Network for the Detection of Atmospheric Composition Change (NDACC) and the validation program for the EOS-Aura satellite. In this paper we review all the stages of the instrument data acquisition, data analysis, profile retrieval and calibration procedures, as well as selected results from the recent validation campaign MOHAVE-2009 (Measurements of Humidity in the Atmosphere and Validation Experiments). The stages in the instrumental development and the conclusions from three validation campaigns (including MOHAVE-2009) are presented in details in a companion paper (McDermid et al., 2011). In its current configuration, the lidar demonstrated capability to measure water vapor profiles from $\sim 1$ km above the ground to the lower stratosphere with an estimated accuracy of 5%. Since 2005, nearly 1000 profiles have been routinely measured with a precision of 10% or better near 13 km. Since 2009, the profiles have typically reached 14 km for 1 h integration times and 1.5 km vertical resolution, and can reach 21 km for 6-h integration times using degraded vertical resolutions.

1 Introduction

Water vapor is a key constituent of the atmosphere due to its radiative, chemical, and thermodynamic properties. In the troposphere, water vapor is a primary greenhouse gas, and in the stratosphere it is produced by methane oxidation, thus linking it to ozone chemistry. Therefore, to fully understand, quantify, and predict future water vapor-related radiative and chemical processes impacting climate change, accurate, long-term water vapor measurements throughout the troposphere and stratosphere are required (Forster and Shine, 1999). Despite water vapor's recently observed trends
(e.g. Hurst et al., 2011), many instruments today cannot achieve the required accuracy without thorough calibration and validation. To help address this issue, the Network for the Detection of Atmospheric Composition Change (NDACC, formerly known as NDSC) included water vapor Raman lidar in its suite of long-term monitoring techniques.

A high capability water vapor Raman lidar was therefore built at the Jet Propulsion Laboratory (JPL) Table Mountain Facility (TMF) in California (34.4° N, 117.5° W, elevation 2285 m), with the primary objective of reaching the upper troposphere and lower stratosphere (UTLS) with the best possible accuracy (5% or better). The instrument, referred to as “TMW” in the remainder of this paper, has been optimized over the years and is now capable of producing routine measurements of water vapor between 4 km and 15–20 km with a precision of 10% or better, and 5% accuracy. The instrumentation setup and optimization over the past five years is fully described in a companion paper (McDermid et al., 2011). The present paper describes in detail the data acquisition and analysis, the calibration procedures, and reviews some of the main results from the latest validation campaign MOHAVE-2009.

A brief review of the Raman lidar technique is first given in Sect. 2. The lidar data acquisition setup and analysis will then be described and results from the MOHAVE-2009 campaign will be presented in Sect. 3. The last section presents additional considerations chosen to guarantee long-term stability of the routine measurements for future use by NDACC.

2 Water vapor Raman lidar measurement principle

The Raman lidar measurement technique is relatively simple in principle and easy to implement (e.g. Vaughan et al., 1988; Sherlock et al., 1999a). A laser pulse is emitted into the atmosphere and scattered by the molecules and particles. A fraction of the laser light is collected back on the ground with a telescope, where it is geometrically and spectrally separated, and sampled in time (i.e. distance). In the case of the
Raman technique, the light scattered by a specific molecule is shifted by an amount that depends on the energy difference between its vibrational and/or rotational states (Hinckley, 1976; Measures, 1992). For water vapor, the Stokes Q branch occurring in the OH-stretching band near the frequency shift of \( 3654 \text{ cm}^{-1} \) is most often used. As will be seen shortly, the water vapor Raman technique also makes use of scattering by a reference molecule with a well-known mixing ratio throughout the altitude range of interest, for example the well-mixed and abundant gas nitrogen. In this case, the strongest Stokes Q-branch occurring at a frequency shift of \( \nu_1 \approx 2330 \text{ cm}^{-1} \) (transition from the ground-state to the first vibrational state) is normally used. The fraction of the total energy scattered at the shifted wavelengths is typically three orders of magnitude smaller than that for elastic scattering. The backscatter coefficient can be expressed as the product of the molecule’s Raman backscatter cross-section \( \sigma \) and its number density \( N \), which leads to the following form of the lidar equation:

\[
P_M(r) = \frac{P_E \kappa_M O_M(r) A_M \delta r}{r^2} \sigma_M(r) N_M(r) \exp \left[ - \int_0^r \left( \alpha_E(r') + \alpha_M(r) \right) dr' \right].
\]

(1)

\( P_E \) is the number of photons emitted, per laser shot (at emission wavelength \( \lambda_E \)).
\( r \) is the distance between the laser and the backscattering layer being considered.
\( \delta r \) is the thickness of the backscattering layer being considered.
\( P_M \) is the number of photons detected, per laser shot, at the wavelength shifted by the target molecule \( M (M = \text{H}_2\text{O or N}_2) \).
\( \kappa_M \) is the overall optical transmittance and quantum efficiency for the channel corresponding to the molecule \( M \).
\( O_M \) is the telescope field-of-views and laser beam overlap factors for the channel corresponding to the molecule \( M \).
\( A_M \) is the receiving area coupled with the channel corresponding to the molecule \( M \).

The terms \( \alpha_E \) and \( \alpha_M \) are the total atmospheric transmittances along the beam path from the lidar to the scattering layer and back to receiver channel corresponding to the molecule \( M \). The above equation describes the collected signal in one individual
channel with ideal noise-free instrumentation. In reality, the total acquired signal is a combination of the collected light backscattered in the atmosphere and noise originating from both residual sky background light and from the instrumentation. The signal can also be subject to non-linearity, especially at very high-count rates. After noise extraction and correction for non-linearities, the ratio of the corrected signals $\bar{P}$ collected in the water vapor ($M = H_2O$) and nitrogen ($M = N_2$) channels can be written:

$$R(r) = \frac{\bar{P}_{H_2O}(r)}{\bar{P}_{N_2}(r)} = \left[\kappa_{eff}(r) \kappa_O(r) \kappa_\sigma(r) \kappa_\alpha(r)\right]^{-1} \frac{N_{H_2O}(r)}{N_{N_2}(r)}. \quad (2)$$

$\kappa_{eff}$ is a constant expressing the ratio of all the optical and quantum efficiencies of the receivers as well as other constant terms.

$\kappa_O(r)$ is the ratio of the overlap functions of the nitrogen and water vapor channels.

$\kappa_\sigma(r)$ is the ratio of the nitrogen and water vapor Raman cross-sections. $\kappa_\alpha(r)$ is the ratio of the particulate extinction along the return path of the beam at the nitrogen and water vapor wavelengths (often referred to as “extinction differential”).

The extinction terms can be separated into molecular and particulate extinction. Molecular extinction can be calculated for each channel prior to computing the ratio $R(r)$ using climatological, modeled, or measured profiles of the air number density and the density of the atmospheric absorbers, leaving only the particulate extinction contribution in Eq. (2). This latter equation can be related to water vapor mixing ratio expressed as a function of number density:

$$q(r) = 0.781 \frac{N_{H_2O}(r)}{N_{N_2}(r)}. \quad (3)$$

Replacing $q(r)$ into Eq. (2) and reverting yields:

$$q(r) = \kappa_{eff} \kappa_O(r) \kappa_\sigma(r) \kappa_\alpha(r) R(r). \quad (3)$$
Depending on the lidar instrument setup, the four multiplicative terms in front of the term $R(r)$ in Eq. (3) have a varying degree of dependency on altitude. Their estimation is known as the “lidar calibration”, and is discussed in Sect. 3.4.

3 Instrument, data acquisition setup and profile retrieval

3.1 Instrumentation overview

A complete description of the instrumentation, including upgrades, is presented in our companion paper (McDermid et al., 2011). The laser is a high pulse energy (650 mJ/pulse) Nd:YAG laser transmitting at 355 nm at a repetition rate of 10 Hz. The primary telescope is a 91 cm-diameter Newtonian telescope for the far-range (high- and mid-intensity channels) supplemented by three (later four) 7.5 cm-diameter refractive telescopes for the near range (low intensity channels). As mentioned in Sect. 2, TMW makes use of the vibrational Raman shift at 2330 cm$^{-1}$ for nitrogen and 3654 cm$^{-1}$ for water vapor, which for an emission at 355 nm, translates to collecting at 387 nm and 407.5 nm respectively (407.5 thereafter referred to as “407” for brevity). The small receivers are each designed to detect only a single wavelength (355 nm, 387 nm and 407 nm), while light from the large telescope is sent to a polychromator where it is geometrically and spectrally separated in five channels: high-intensity 355 nm, 387 nm and 407 nm, and mid-intensity 387 and 407 nm.

3.2 Data acquisition system and profile retrieval

The signals from the detectors (photomultipliers) of all channels are input to a series of photon counting multi-channel-scalers (Licel Transient Recorder). The Licel units, triggered at 10 Hz, collect the signal in 16 384 bins of 7.5 m width each (dwell time of 50 ns). The photon counts are then summed and stored into 5-min-integrated data files with a 75-m bin width, before being analyzed.
The signals from all 3 pairs of 387–407 nm channels are analyzed for the retrieval of water vapor (3–20 km) by our lidar analysis software LidAna, versions 6.2 or 7.0 (version depending on the application). These two versions are the latest optimized versions of the software which has been used for over a decade to analyze all existing JPL lidar products archived at NDACC. Temperature (10–90 km) and backscatter ratio (4–40 km) profiles can also be retrieved, though no routine temperature or aerosol product is currently derived from the TMW lidar signals (these products exist from the other JPL lidars). Nevertheless, preliminary temperature results obtained from TMW during the MOHAVE-2009 campaign are shown in McGee et al. (2011).

For all TMW channels, the collected signals are first corrected for background noise. The raw signals are fitted over an altitude range where only noise is known to be present. For water vapor retrieval, a simple linear fit with a zero-slope (constant noise) is usually sufficient as no signal-induced noise is present in any of the Raman channels. At the bottom of the channels’ useful range, signal non-linearities (pulse pile-up effect) are corrected either empirically using the method described in Donovan et al. (1993), or experimentally using the non-saturated signals from the lower intensity channels. The high intensity pair is optimized to provide water vapor measurements between 8 km and 20 km. The useful range can be extended downward to 5 km with proper correction of saturation using the non-saturated low intensity pair coupled to the large telescope, and which also provides valid measurements between 4 km and 8 km. The two pairs of channels coupled out of the large telescope were designed so that the contribution of their overlap functions in Eq. (3) is nearly independent of height above 7 km. Below this altitude, this contribution is occasionally verified experimentally using the pair of channels coupled out of the small telescopes. The correction however is not used in the normal data processing since it sometimes yields larger errors than if we assume that the ratio of the overlap functions in the narrow field-of-view water vapor and nitrogen channels cancel out. This latter assumption is verified by simply comparing the profiles obtained from the narrow and wide field-of-view channels (see Fig. 2 later in this section). The low intensity pair of wide fov channels provides water vapor
profiles between altitudes of 3 km (600 m above ground) and 7 km, with a contribution of the overlap functions negligible down to the lowest useful data bins (2.8 km altitude).

In addition to background correction, saturation correction, and the optional overlap correction, the signals are corrected for atmospheric extinction along the laser beam path using the density profiles computed from the National Centre for Environmental Prediction (NCEP) temperature and pressure profiles interpolated at the measuring site. These profiles are made available on a daily basis to all NDACC participants at the NDACC Data Handling Center (http://www.ndacc.org/). Though mostly negligible at these wavelengths, a correction for ozone absorption is further applied to all channels before water vapor is computed from the uncalibrated ratio following Eq. (3).

At this stage of the analysis, the signal ratios for each of the three 387/407 nm pairs of channels should be smoothed to mitigate the random noise due to photon counting. It is made using a height dependent smoothing scheme that limits random uncertainties to less than 10 % at all altitudes except in the very highest altitude bins (14 km and above). A Kaiser filter with a fixed attenuation of −50 dB is used to smooth the data (Kaiser and Reed, 1977). The attenuation is the only fixed input parameter of the filter. The other parameters (cut-off frequency and number of coefficients) are automatically computed as a function of height to provide a minimum impact of smoothing, yet insuring precision of 10 % of better throughout the profile below about 14 km. Above this altitude, the maximum number of points used for filtering is capped to 97 (7.2 km full-width), which causes precision to degrade from 10 % to 25 % or more. Vertical resolution is reported in the data files following the definition of the cutoff frequency of the filter. Figure 1 illustrates the effect of the signal filtering expressed as a function of the Kaiser filter’s cutoff frequency and number of coefficients (top panel), and expressed as a function of the full-width at half-max (FWHM) of an Impulse Response (Dirac’s Delta function) (bottom panel). Following the definition based on cutoff frequency, unsmoothed signals are reported with a vertical resolution of 150-m (two sampling bins, corresponding to the Nyquist frequency), and a cutoff frequency of 0.05 yields a vertical resolution of 20 bins (1.5 km). The typical variation with height of vertical resolution and the
impact of the smoothing on precision is shown in Fig. 2 and Table 1 for a 2-h integrated measurement.

The last stage of the analysis is the calibration of the signal ratios of all 3 pairs of channels, as well as the combination of the pairs into one single profile. This procedure was carefully optimized to insure a minimum impact on the stability of the future lidar long-term time series, and is discussed next.

3.3 Calibration and profile assembly

Calibration of water vapor Raman lidar measurements has been extensively discussed in the past (e.g. Vaughan et al., 1988; Sherlock et al., 1999b; Whiteman et al., 2003; Leblanc and McDermid, 2008). There are two main approaches: One approach consists of calculating every single term of Eq. (3) linking \( R(z) \) and \( q(z) \). Since this task is complex and has many sources of uncertainty (including – but not limited to – the accuracy of the lidar parts’ manufacturer specifications and the determination of the Raman water vapor cross-section), the resulting calibration overall accuracy using this approach is rarely found to be better than 10 %.

A second approach consists of estimating and/or minimizing any height-dependent term in Eq. (3) (namely, the ratio of the overlap functions, differential aerosol extinction, and temperature dependence of the ratio of the Raman water vapor and nitrogen cross-sections), and reduce all the terms of this equation to a single, height-independent proportionality constant. This constant can then be deduced by scaling the lidar ratios to one (or a set of) well-known water vapor mixing ratio value(s) measured by another technique. Radiosonde measurement in the troposphere is the most common source used today. Another common source of calibration is the Total Precipitable Water (TPW) measurement from a co-located GPS or microwave instrument. When using an external measurement, the accuracy of the calibration procedure follows that of the measurement used. Today the accuracy of the best quality radiosondes, GPS, and microwave measurements is estimated to be 5 %, 7 % and 10 % respectively. The accuracy of the lidar calibration using this approach also depends on the quality of the
spatio-temporal coincidence between the lidar and the correlative measurement. In the case of the TMW lidar, we carefully evaluated the calibration accuracy using several sources of measurement, several coincidence criteria, and several normalization techniques. The external source consists of Vaisala RS92 water vapor profiles corrected using the method described by Miloshevich et al. (2004, 2009). A combination of the matching and normalization methods producing the least temporal variability (on a yearly basis or longer) was considered the most accurate and was eventually retained for use in the LidAna v7 standard analysis program. The results of this work are summarized below.

Four different coincidence criteria were tested. Figure 3 shows a schematic of each of the four methods for a 2-h-long lidar measurement. On each figure, the lidar measurements are denoted by red open rectangles symbolizing a series of 24 consecutive 5-min datasets (time-altitude). The external source of calibration (in this case radiosonde) is symbolized by a one-time flight launched at $t = 0$ and during which water vapor is measured quasi-instantaneously (blue tilted striped line). The coincidence criterion for each method leads to a set of coincident lidar-radiosonde data pairs used to scale the uncalibrated lidar profile. These data pairs are represented on each figure by the green thick circles. No point above 10 km altitude is used in the calibration process due to increasing random noise of the lidar signals. For each calibration method the lidar measurements are partially integrated over the time window for which coinciding pairs were found. For “Matching Method 1” (top-left panel) this results in scaling, for each altitude bin, the 2-h lidar average to the instantaneous radiosonde measurement. For “Matching Method 2” (bottom-left panel), the averaging time window is restricted to the duration of the radiosonde flight below 10 km, which is approximately 30 min. For “Matching Method 3”, the entire 2 h window is used, but only the altitude points where water vapor variability over the 2-h period was less than 20% are used. Finally, in “Matching Method 4” (bottom-right panel), only the data points strictly coinciding both in time and altitude are used. For each of the above matching methods, three different scaling algorithms were used. In “Scaling Method 1”, a simple average of the
ratio calculated for all matching pairs is calculated. For “Scaling Method 2” the mean value is replaced by the median value. In “Scaling Method 3” a Gaussian distribution of the ratios is computed and the calibration constant is set to the center value of this Gaussian function. Eventually 12 cases (3 matching methods multiplied by 4 scaling methods) were evaluated.

The standard deviations in the lidar calibration constant obtained from each method over a test-period of 16 months (October 2007–April 2009) are compiled in Table 2, and time series of the calibration constants for six of the twelve cases studied are plotted in Fig. 4. In this figure, the calibration constant is represented by a vertical bar for each of the 118 measurement nights used. First, the uncalibrated lidar ratio for each altitude bin between 3.5 km and 10 km is scaled to the radiosonde value. The mean value, median value, or center of Gaussian distribution value is calculated over all available altitude bins. The resulting value is located at the center of the plotted vertical bar. Each bar extends one standard deviation up and down from the mean (for the Gaussian distribution method, it is the full-width at half-maximum). The length of each bar therefore gives an indication of the stability of the calibration process when different altitude bins are used. However it does not provide an indication of its stability in time. To investigate the stability in time, the standard deviations (in time) of the daily calibration values are computed. Their values are displayed in Fig. 4 as percents of the mean over the period considered. The full time series is divided in 4 uneven periods, indicated by black horizontal bars. The first two periods are short and correspond to the MOHAVE-II campaign (October 2009). They are not used to compute the percentages displayed on the figure. The standard deviations for each of the two other periods, as well as the average of the two, are displayed on the figure. They range from 6.8 % to 14.6 %, depending on the method used. Out of all the methods, the “Matching Method 2” (restricted time and unrestricted altitude coincidence) and “Scaling Method 2” (median) turned out to provide the best results. They are now used systematically in our standard data processing. The standard deviation of several other methods remain in the same order of magnitude (10 % or better), and could therefore be used as well.
In addition to the spatio-temporal match of the lidar and radiosonde profiles, one must pick the best combination of the three ranges (high-intensity, mid-intensity, and low-intensity) in order to insure the best accuracy and stability in time. The profiles from the high-intensity channels contain less random noise at higher altitudes but have a larger uncertainty associated with overlap and saturation at the bottom. On the other hand, the low-intensity channels coupled from the small telescopes are less sensitive to overlap issues at the lowest levels but quickly become noisy above 5 km altitude. The accuracy degradation of the low-and high-intensity ranges is mitigated for the mid-intensity range once corrected for saturation and overlap. We therefore use this range to calibrate the lidar profiles to the radiosonde measurements, typically at altitudes between 4 km and 7 km. The other two ranges are then normalized to the mid-intensity range using the altitudes of best overlap (typically 6–8 km for the high-intensity range, and 3–5 km for the low-intensity range). Though we use only one range, calibrating all three ranges independently using the radiosonde profiles yields very similar results in most cases and could be used likewise. An example of profile with all three ranges combined is shown in Fig. 5.

4 Results from the MOHAVE-2009 campaign

The latest and most important change in the TMW instrument configuration occurred during the summer of 2009. No further changes occurred after that and the instrument remained untouched throughout the MOHAVE-2009 campaign with the exception of the calibration lamp, whose purpose is described in the next section.

Two types of frost-point hygrometers (CFH and NOAA-Frost Point), two types of radiosonde (Vaisala and InterMet), two microwave radiometers (NRL and University of Bern), two Fourier-Transform spectrometers (the JPL MkIV and FTUVS), and two GPS receivers participated to the campaign. Coincident measurements from several satellite instruments (Aura-MLS, Aqua-AIRS, Aura-TES, ENVISAT-MIPAS and ACE-FTS)
were also compared to the balloon-borne and ground-based measurements. The campaign lasted for approximately two weeks between 11 October and 27 October 2009. A overview of the campaign operations and results is presented in Leblanc et al. (2011). Documentation and selected results can be found on the campaign website: http://tmf-lidar.jpl.nasa.gov/campaigns/mohave2009.htm. A total of 44 balloons were launched. Each balloon payload contained a minimum of one single radiosonde and a maximum of two radiosondes, one ozonesonde, and one hygrometer sonde. A total of 58 RS92 radiosondes, 16 CFH, 4 NOAA-FPH, and 16 ECC ozonesondes were launched over the duration of campaign. Approximately 300 h of water vapor lidar measurements were acquired. Results from MIPAS, the MkIV spectrometer, balloon measurements, and two lidar instruments are presented in Stiller et al. (2001), Toon et al. (2011), Hurst et al. (2011), Whiteman et al. (2011), and McGee et al. (2011) respectively.

Water vapor has different scales of natural variability in the troposphere and stratosphere. For this reason the coincidence criteria used for the comparisons presented here are different above and below 14 km: all profiles coinciding within 1 h and 100 km were used for altitudes below 14 km, and all profiles coinciding within 6 h and 250 km were used for all altitudes above 14 km.

Figure 6 shows the campaign average of all TMW and all RS92 relative humidity (RH) profiles measured simultaneously (i.e. within 1-h of balloon launch), and their mean difference. The comparisons are shown for the RS92 uncorrected (top row panels) and corrected (bottom row panels) versions. As mentioned before, the TMW profiles were calibrated using the radiosonde profiles between 4 and 7 km. The corrected version of the RS92 profiles was used, leading obviously to a near-zero difference in the resulting RH over the pressure range 600–400 hPa (Fig. 5, bottom-right panel). At these pressures, a small negative bias with the uncorrected RH is observed, a direct consequence of the effect of the RS92 correction (corrected RH values smaller than uncorrected values). Most interestingly, the mean bias between the lidar and the corrected RH remains almost negligible (below 3 %) and not statistically significant all
the way up to 100 hPa. In the upper troposphere, a mean bias of 3–5 % is observed, the radiosonde being drier than the lidar.

Figure 7 (top row panels) shows the campaign average of all coincident water vapor profiles (left panel) measured by TMW and Aura MLS (version 3), and their difference (right panel). As we are in the UTLS, the coincidence criterion was relaxed from that in the troposphere. Nevertheless, only 3 profiles were found coincident. MLS shows a 7–10 % dry bias in the lower stratosphere (200–30 hPa) with respect to TMW. However the bias is not statistically significant. The only significant difference is a large dry bias for MLS at 250 hPa. This feature is caused by the retrieval in response of the very fast transition from the dry stratosphere to the wet troposphere. The bottom row panels of Fig. 7 shows the campaign average of all coincident water vapor profiles (left panel) measured by TMW and Aqua-AIRS, and their difference (right panel). Despite the very different sampling type, the two instruments remain in very good agreement. A 5 % bias can be observed (AIRS being wetter) between 400 hPa and 150 hPa. Below 400 hPa the somewhat larger differences are not statistically significant.

The campaign-mean profiles measured simultaneously by the TMW lidar and CFH, and their difference, are plotted as a function of pressure in Fig. 8. This figure summarizes well the performance achieved by the TMW lidar during MOHAVE-2009. The comparisons were made using two different coincidence criteria. In the troposphere, i.e. below approximately 100 hPa (bottom panels), atmospheric variability is very high and only the profiles coinciding strictly were used to compute the mean (i.e. 12 CFH flights and the 12 corresponding 1-h integrated lidar profiles). In the UTLS, i.e. above approximately 200 hPa (top panels), the lidar profiles integrated all-night were used (mean of 8 nights during which 9 CFH were launched). The TMW lidar and CFH profiles are in excellent agreement throughout the troposphere and the UTLS. The mean differences range from −5 % in the lowermost troposphere (TMW drier) to +2 % in the lower stratosphere (TMW wetter). These values are well below the reported uncertainties of both instruments.
A closer inspection of Fig. 8 top-left (UTLS, 200–50 hPa) reveals that the standard deviation of the lidar measurements above 100 hPa is two to three times larger than that of the CFH. This number provides a good indication of the number of additional lidar measurements needed to achieve the same precision as CFH. With the assumption that lidar data contains random noise following a Poisson distribution, using one monthly CFH profile yields a precision in the UTLS equivalent to that of the integrated lidar measurement of one full night per week, or four 2-h-long measurements per week. These estimations show that with a typical routine schedule of 2 h per night, 4 nights per week, the TMW lidar can achieve a precision in the UTLS equivalent to that achieved when launching one CFH per month.

5 Additional considerations for long-term monitoring

5.1 Alternative calibration using total precipitable water (TPW)

In order to ensure the proper long-term monitoring of water vapor mixing ratio, the calibration must remain as stable as possible on large time scales (several years). Therefore, an alternative source of external measurement for the calibration of TMW is now systematically used, namely the TPW measurements from a co-located GPS. Since the lidar cannot sample all altitudes down to the ground, the scaling of total water must be made carefully and the lidar-blind lowest atmospheric layers must be taken into account. To achieve this, the lidar profiles are extrapolated down to the ground using ground measurements obtained every 5 min from a co-located Vaisala Weather Station (MAWS-1). The scaling process is illustrated in Fig. 9 with the same coloring conventions as Fig. 3 (i.e. red for lidar, blue for external source, and green for coincident data pairs). Note that the GPS measurements are sampled every 15-min, which corresponds to using 3 correlative 5-min lidar datasets. The schematic on the right side of the figure illustrates the scaling method, referred to as “Stick-and-Slide”: the uncalibrated lidar mixing ratio profile is glued to the ground (“stick”) using the measurements...
by the Vaisala MAWS station. The contribution from the lowermost layers is taken into account by interpolating the mixing ratio between the ground value and the uncalibrated lidar profile bottom value. The uncalibrated profile is shifted (“slide”) until the corresponding TPW exceeds or falls below the value measured by the collocated GPS. For each sliding interval, the contribution of the lowermost layers must be re-interpolated. The lidar-computed TPW converges to the GPS value by reversing the sliding direction multiple times and by dividing the sliding interval by two each time a change of direction occurs (dichotomic algorithm). The algorithm stops when the difference between the lidar-computed and GPS-measured TPW values falls below a user-specified residual (for example 0.1 %). Sensitivity tests have shown that the method is most efficient when the uncalibrated lidar profile is cut-off at a bottom altitude where a compromise is reached between the impact of signal saturation and/or overlap and the impact of missing measurements between the ground and the first useful lidar sampling bin. In particular, it is more accurate to cut-off the lidar profile a few hundred meters farther up and avoid any profile segment contaminated by saturation or overlap, than starting the profile in its lowermost sampling bins where contamination may occur. Assuming no contamination by saturation or overlap, the accuracy of the “Stick-and-Slide” method was found to be around 5 % when the (contamination-free) lidar profile is cut-off 500 meters above ground, and 15 % when it is cut-off 1000 m above ground. This degradation follows from the accumulated uncertainty associated with unaccounted water vapor in the lowermost layers. Additional uncertainty comes from the inherent accuracy of the GPS (or microwave) measurements, i.e. 7 % or 10 %. Overall this calibration method is a good alternate to the radiosonde calibration method, though not as accurate. It has however the advantage of being more stable over longer periods of time (several years), because it is not subject to manufacturer changes like those experienced over the past few decades with radiosonde (several Vaisala radiosonde versions, each having different observed biases).
5.2 Hybrid calibration

The requirement for long-term stability of the lidar calibration has been mentioned several times in this paper already. Indeed even after the calibration method was optimized for our TMW lidar, natural variability of tropospheric water vapor can lead to calibration changes of 15% or larger from night to night, which does not reflect the actual changes of the lidar system but simply the fact that a different region of the atmosphere was sampled by lidar and radiosonde. The only fix to this problem is to launch not one, but several radiosondes throughout the lidar experiment (typically four radiosondes for a 2-h long lidar experiment). Though radiosondes are affordable such an intensive launch plan is too expensive and therefore inapplicable to the routine, long-term measurements of water vapor by lidar. However, and as explained below, a well chosen combination of radiosonde and partial calibration experiments can bring a solution to the problem.

Since 2007, we have systematically used a calibration lamp to monitor the changes in the calibration of the lidar receiver. The lamp is currently mounted on the receiver module next the large telescope Newtonian focus and illuminates mostly downward towards the primary mirror and partially upward towards the roof hatch. Following the method described by Leblanc and McDermid (2008), signals coming from the illumination of the lamp with the hatch closed and with the laser turned off are acquired for 15 min before and after a regular atmospheric water vapor data acquisition experiment. The 387/407 nm ratio obtained during these routine “lamp-runs” is used to monitor any changes in the receiver transmittance, then compared to that obtained during occasional, intensive calibration campaigns (typically once a year). If this ratio has not changed significantly, then an absolute calibration can be applied retrospectively to all experiments acquired between the two campaigns. This procedure allows saving time and money since routine radiosonde launches are no longer required. It also potentially reduces uncertainty associated with spatio-temporal matching since...
the calibration constant can be averaged using many launches made during a campaign instead of relying on individual coincidences. Another advantage is the flexibility to choose only the most stable nights, variability-wise, during which absolute calibration is performed. A critical requirement for the Hybrid method to be valid is that the lamp must remain fixed at the same location and undisturbed throughout the period between two consecutive campaigns. If the lamp is moved, or if any change in the partial calibration constant is detected between absolute calibration campaigns, then radiosondes must be launched immediately to quantify the impact of these changes on the absolute calibration constant. The Hybrid method is described in details in Leblanc and McDermid (2008). Though radiosondes are used in the hybrid method described here, any source of accurate measurement may be used for the absolute calibration campaigns.

Two lamps (200 W and 45 W Quartz-Tungsten Halogen) have been used since the hybrid method was introduced in October 2007. A complete review of the signal ratios of the 387 and 407 nm channels obtained during the routine lamp runs is presented in Fig. 10. Three periods are presented, which corresponds to the different receiver configurations already discussed. The top figure shows the signal ratios for all three ranges (high-, mid- and low-intensity) between October 2007 (MOHAVE-II, when the lamp runs started) and summer 2009. This figure is indeed an extension of Fig. 8 of Leblanc and McDermid (2008), in which the observed jumps and drops in the channel ratios until April 2008 are discussed. These features will not be discussed again here, though the instrumentation changes are noted on the figure. The second panel (middle) highlights the major configuration changes made between June and October 2009: receiver redesign, field stop optimization, and most importantly, lamp change (from 200 W to 45 W) during MOHAVE-2009. The third period (bottom panel), extending from MOHAVE-2009 to present time, shows the channel ratios since the 45 W lamp has been in operation. No instrumentation change was made throughout this period. Inspection of all three plots leads to several important conclusions. First, and as anticipated in Leblanc and McDermid (2008), the channel ratios obtained
during lamp runs remain very stable over time (standard deviation mostly below 1.5 % over timescales of a year) unless an instrumentation change occurs. Second, no apparent drift is observed at these timescales with the exception of the mid-range channel ratio during period 3, which shows a 5 % drop in late spring 2010 not associated with any instrumentation change. Third and most importantly, all observed standard deviations as well as the spring 2010 mid-range channel ratio drop are closely related to the magnitude of the lamp irradiance with respect to the magnitude of these channels’ background noise. During period 1, the lamp is brighter (200 W), and provides channel illumination far above all channels’ internal noise. The observed standard deviations in this case reflect only the spectral stability of the lamp irradiance, i.e. about 1.5 %. During period 3, the lamp is dimmer (45 W), and only the standard deviation of the ratio of the high-intensity channels reflects the lamp spectral stability. The standard deviation of the other two ranges reflects a combination of the lamp spectral stability and non-negligible internal background noise. Indeed the mid-range ratio drop in spring 2010 is a direct consequence of the mean background noise reduction in the 407 nm channel. This apparent reduction was associated with the presence (before spring 2010), then absence (after spring 2010) of undesired isolated spikes in the background noise. Due to of its subtle nature, this feature was never investigated and there is no explanation for it today.

During MOHAVE-2009, the 200W lamp initially used was replaced by a dimmer one because the signals acquired by the high-intensity channels were potentially affected by saturation. This hypothesis was indeed verified as the change in the channel ratio associated with the lamp replacement turned out to be very different from those observed on the other two ranges (ratio multiplied by 4 for the high intensity range as opposed to a 20 % increase for the other two ranges, see middle panel of Fig. 10). This finding therefore leads to an important conclusion regarding the general use of a lamp: neither 200 W nor 45 W is currently appropriate to obtain an optimized channel ratio for all ranges. In our case, we can (and shall) safely use the Hybrid technique for the mid- and low-intensity ranges only before MOHAVE-2009, and we can (and
shall) safely use it for the high-intensity range only after MOHAVE-2009. Ideas for an improved methodology that could be used simultaneously for all ranges are currently being explored.

6 Conclusions

Accurate measurements of water vapor throughout the troposphere and UTLS have been made routinely for several years by a high capability water vapor Raman lidar developed at the Jet Propulsion Laboratory Table Mountain Facility, California. The lidar system, referred to as “TMW”, was built in 2005 and optimized over the past four years. Because the overall goal of the TMW instrument is to provide long-term monitoring, several additional considerations (compared to traditional water vapor Raman lidar) were undertaken to insure optimal lidar calibration stability with time. Besides radiosonde, an alternative source of calibration is now systematically used, namely the TPW measurements from a co-located GPS receiver. Furthermore, a dedicated calibration method combining a laboratory lamp and radiosondes, and referred to as “hybrid calibration” method, is now used to minimize the cost of launching radiosondes and increase the accuracy and stability of the absolute calibration.

To support and achieve a full optimization (from measurement to retrieval), several validation campaigns took place, during which the lidar measurements were extensively compared to operational radiosonde (Vaisala RS92) and Research-grade instruments (CFH). The latest comparisons with RS92 and CFH measurements (MOHAVE-2009) showed excellent agreement throughout the troposphere and UTLS, with mean biases well below the reported measurements uncertainties. The TMW lidar shows a 2% mean dry bias with CFH in the lower troposphere and a 3% mean wet bias in the UTLS, and virtually no bias with the corrected RS92 measurements throughout the troposphere (ground to 18 km).

Inspection of the CFH and lidar measurement standard deviations above 16 km during MOHAVE-2009 showed that the lidar signals are still noise limited in the UTLS,
and that the standard deviation values mainly reflect the precision of the lidar measurements. It therefore provides a good indication of the number of additional lidar measurements needed to achieve the same precision as that of CFH. It was estimated that one monthly CFH profile yields a precision in the UTLS equivalent to that of the integrated lidar measurement of one full night per week, or four 2-h-long measurements per week. With our present target to routinely operate TMW 2 h per night, 4 nights per week, we can achieve measurements with a precision in the UTLS equivalent to that achieved if launching one CFH per month. It is not unlikely that future instrumentation upgrades (for example a laser replacement) will enhance further the performance of the lidar. All in all, the TMW water vapor Raman lidar has now achieved a comfortable level of maturity and is expected to contribute high-quality long-term, routine profiles of water vapor to the NDACC database, starting in summer 2011.

Acknowledgements. The work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under agreements with the National Aeronautics and Space Administration. The authors are very grateful to Tony Grigsby and Jeffrey Howe for assisting in the collection of the data presented here.

References


Table 1. Vertical filtering scheme (Kaiser filter, 50 dB attenuation) applied to the high-intensity range for a typical 2-h integrated measurement, and resulting vertical resolution and precision.

<table>
<thead>
<tr>
<th>Alt. (km)</th>
<th>Cutoff Freq.</th>
<th>2N + 1 (FW)</th>
<th>Vert. Resol. (km)</th>
<th>I. R.** (km)</th>
<th>Prec. (%) before</th>
<th>Prec. (%) after</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.46</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>5.06</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>6.035</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>7.01</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>8.06</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>9.035</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>10.01</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>11.06</td>
<td>0.5</td>
<td>1*</td>
<td>0.15</td>
<td>0.075</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>12.035</td>
<td>0.428</td>
<td>3</td>
<td>0.15</td>
<td>0.086</td>
<td>6.2</td>
<td>6.1</td>
</tr>
<tr>
<td>13.01</td>
<td>0.173</td>
<td>7</td>
<td>0.45</td>
<td>0.2</td>
<td>9.5</td>
<td>5.0</td>
</tr>
<tr>
<td>14.06</td>
<td>0.078</td>
<td>13</td>
<td>0.9</td>
<td>0.4</td>
<td>14.8</td>
<td>5.9</td>
</tr>
<tr>
<td>15.035</td>
<td>0.041</td>
<td>25</td>
<td>1.8</td>
<td>0.8</td>
<td>22.0</td>
<td>7.2</td>
</tr>
<tr>
<td>16.01</td>
<td>0.023</td>
<td>45</td>
<td>3.3</td>
<td>1.4</td>
<td>32.4</td>
<td>8.2</td>
</tr>
<tr>
<td>17.06</td>
<td>0.013</td>
<td>77</td>
<td>5.7</td>
<td>2.4</td>
<td>48.3</td>
<td>10.7</td>
</tr>
<tr>
<td>18.035</td>
<td>0.010</td>
<td>97</td>
<td>7.2</td>
<td>3.1</td>
<td>69.0</td>
<td>15.7</td>
</tr>
<tr>
<td>19.01</td>
<td>0.010</td>
<td>97</td>
<td>7.2</td>
<td>3.1</td>
<td>97.2</td>
<td>22.1</td>
</tr>
<tr>
<td>20.06</td>
<td>0.010</td>
<td>97</td>
<td>7.2</td>
<td>3.1</td>
<td>&gt;100</td>
<td>26.0</td>
</tr>
</tbody>
</table>

* 2N + 1 = 1 (i.e. N = 0) means no smoothing.
** I. R. is the filter's corresponding Impulse Response (to Dirac's Delta Function) full-width at half-maximum (FWHM).
Table 2. Standard deviation of the lidar calibration constant calculated 120 2-h-long samples over the period October 2007–April 2009 using 12 different calibration methods (single radiosonde coincidences). See text for details.

<table>
<thead>
<tr>
<th>Scaling Matching</th>
<th>Method 1 (Mean)</th>
<th>Method 2 (Median)</th>
<th>Method 3 (Gaussian)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1 (all datasets)</td>
<td>16 %</td>
<td>12 %</td>
<td>10 %</td>
</tr>
<tr>
<td>Method 2 (time-coincident)</td>
<td>8 %</td>
<td>7 %</td>
<td>8 %</td>
</tr>
<tr>
<td>Method 3 (min. variability)</td>
<td>12 %</td>
<td>11 %</td>
<td>10 %</td>
</tr>
<tr>
<td>Method 4 (time + altitude coincident)</td>
<td>11 %</td>
<td>11 %</td>
<td>11 %</td>
</tr>
</tbody>
</table>
Fig. 1. Effect of the TMW lidar signals vertical filtering based upon two different vertical resolution definitions. Top panel: definition based on the cut-off frequency of a digital filter (Kaiser filter with 50 dB attenuation). Bottom panel: definition based on the full-width at half-maximum (FWHM) of an Impulse Response to a Dirac Delta Function.
Fig. 2. Example of the effect of vertical filtering on the TMW water vapor profiles. Left panel: mixing ratio profiles for all three ranges (pink: low-intensity, green: mid-intensity, and blue: high-intensity range). Right panel: statistical uncertainty (%) before (solid curves) and after (dashed curves) filtering, and corresponding vertical resolution (km) (dotted curves) following the definition based on digital filter cutoff frequency.
Fig. 3. Schematics of the four methods tested to optimize the calibration of the TMW lidar profiles using co-located radiosondes. Lidar data points are symbolized by red open rectangles, radiosonde data points by small blue open squares tilted with time, and the actual data pairs used for the normalization by filled green circles.
Fig. 4. Six examples of the 12 calibration methods tested. The percentages represent the standard deviations of the calculated calibration constants over 2 sub-periods and averaged over the entire 118 samples (October 2007 to April 2009). See text for details on each method, and Table 2 for a compilation of the standard deviations of all 12 methods.
Fig. 5. An example of the construction of the final (black) water vapor (left panel) and Relative Humidity (right panel) profiles from the three channel ranges (pink: low-intensity, green: mid-intensity, and blue: high intensity). The lidar profiles total uncertainty is over-plotted using thin solid curves, and a co-located radiosonde profile is over-potted using a dotted orange curve.
Fig. 6. Comparison of the MOHAVE-2009 campaign-averaged relative humidity profiles calculated from the 28 simultaneous profiles measured by the TMW lidar and Vaisala RS92 radiosondes. Top panels: mean profiles (left panel) and difference with the uncorrected RS92 data (right panel). Bottom panels: mean profiles (left panel) and difference with the corrected (Miloshevich et al., 2004, 2009) RS92 data (right panel).
Fig. 7. Top panels: averaged mixing ratio profiles (left panel) calculated from the 3 coincident profiles measured in the UTLS by the TMW lidar and Aura-MLS during MOHAVE-2009, and their difference (right panel). Bottom panels: averaged mixing ratio profiles (left panel) calculated from the 4 simultaneous profiles measured by the TMW lidar and Aqua-AIRS during MOHAVE-2009, and their difference (right panel).
Fig. 8. Comparison of the MOHAVE-2009 campaign-averaged mixing ratio profiles calculated from all coincident profiles measured by the TMW lidar and the CFH. Top panels: mean profiles (left panel) and difference (right panel) computed in the UTLS using a loose time coincidence criterion. Bottom panels: mean profiles (left panel) and difference (right panel) computed in the troposphere using a strict time coincidence criterion.
Fig. 9. Schematics of the “stick-and-slide” method used to perform a calibration of the TMW lidar profiles using co-located and simultaneous GPS Total Precipitable Water measurements. Lidar data points are symbolized by red open rectangles, GPS column data by blue vertical bars, and the actual column data pairs used for the normalization by filled green vertical bars. The sliding uncalibrated lidar profile is shown in red, the constrained water vapor ground-measurement in blue, and the “sticky” interpolated layer in green. See text for details.
Fig. 10. Historical evolution of the 407 nm to 387 nm channel ratios (pink: low-intensity, green: mid-intensity, and blue: high-intensity) between October 2007 and present. All instrumentation changes are denoted by arrows and comments, including the change of lamp (from 200 W to 45 W) in October 2009.