Ground-Based Water Vapor Raman Lidar Measurements up to the Upper Troposphere and Lower Stratosphere 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}

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

Abstract

Recognizing the importance of water vapor in the upper troposphere and lower stratosphere (UTLS) and the scarcity of high-quality, long-term measurements, JPL began the development of a powerful Raman lidar in 2005 to try to meet these needs. This development 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 the stages in the instrumental development, data acquisition and analysis, profile retrieval and calibration procedures of the lidar, as well as selected results from three validation campaigns: MOHAVE (Measurements of Humidity in the Atmosphere and Validation Experiments), MOHAVE-II, and MOHAVE 2009.

In particular, one critical result from this latest campaign is the very good agreement (well below the reported uncertainties) observed between the lidar and the Cryogenic Frost-Point Hygrometer in the entire lidar range 3-20 km, with a mean bias not exceeding 2% (lidar dry) in the lower troposphere, and 3% (lidar moist) in the UTLS. Ultimately the lidar has demonstrated capability to measure water vapor profiles from ~1 km above the ground to the lower stratosphere with a precision of 10% or better near 13 km and below, and an estimated accuracy of 5%. Since 2005, nearly 1000 profiles have been routinely measured, and since 2009, the profiles have typically reached 14 km for one-hour integration times and 1.5 km vertical resolution, and can reach 21 km for 6-hour integration times using degraded vertical resolutions.

These performance figures show that, with our present target of routinely running our lidar two hours 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

1 Introduction

Due to its radiative, chemical, and thermodynamic properties, water vapor has long been identified as a key constituent of the atmosphere. The water vapor molecule strongly absorbs infrared radiation and in the troposphere it is therefore a primary greenhouse gas. In the stratosphere it is produced by methane oxidation thus linking it to ozone chemistry. It was shown that a global increase in lower stratospheric H_2O mixing ratio similar to that observed locally since 1981 (Oltmans and Hofmann, 1995; Hurst et. al., 2011a) would contribute to a surface warming reaching 40% of that responsible from CO_2 increases over the same period (Forster and

Shine, 1999). The resulting lower stratospheric cooling would be of same order of magnitude as that caused by changes in ozone concentrations. Therefore, to fully understand, quantify, and predict future water vapor-related radiative and chemical processes impacting climate change, high accuracy water vapor measurements throughout the troposphere and stratosphere are required (typically 3-10%). Despite water vapor's recently observed trends (e.g., Hurst et al., 2011a), 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 (34.4ºN, 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) (Leblanc et al., 2008a). 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 present paper reviews the TMW instrumentation setup and optimization over the past five years, describes the data acquisition and analysis, the calibration procedures, and presents results from the latest validation campaign MOHAVE-2009, which best characterize the present performance of the system. A brief review of the Raman lidar technique is first given in **section 2**. The instrument technical description and achievements, including multiple upgrades over the period 2005-present and the various stages of its validation, are presented in **sections 3 and 6**. The lidar data analysis and profile retrieval are described in **section 4**. Results from the most recent validation campaign MOHAVE-2009 are presented in **section 7**. The last section reviews 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 (Melfi et al., 1969; 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 cm^{-1} is most often used. 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. In the case of the well-mixed and abundant gas nitrogen, the strongest Stokes Q-branch occurring at a frequency shift of $\nu_1 \approx 2330$ 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 σ and its number density N, which leads to the following form of the lidar equation:

$$P_M(r) = P_E \kappa_M \frac{O_M(r) A_M \delta r}{r^2} \sigma_M(r) N_M(r) \exp\left[-\int_0^r \left(\alpha_E(r') + \alpha_M(r')\right) dr' \right] \quad (1)$$

P_E is the number of photons emitted, per laser shot (at emission wavelength λ_E)

r is the distance between the laser and the backscattering layer being considered

δr 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 = H2O or N2)

κ_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 α_E, and α_M are the total atmospheric transmittances along the beam path from the lidar to the scattering layer and back to the 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 $\overline{P}$ collected in the water vapor (M=H2O) and nitrogen (M=N2) channels can be written:

$$R(r) = \frac{\overline{P}_{H2O}(r)}{\overline{P}_{N2}(r)} = \kappa_{eff}(r) \kappa_O(r) \kappa_\sigma(r) \kappa_\alpha(r) \frac{N_{H2O}(r)}{N_{N2}(r)} \quad (2)$$

κ_{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:

$$q(r) = 0.781 \frac{N_{H2O}(r)}{N_{N2}(r)}$$

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 **sections 4.2 and 7**.

3 Instrument Initial Setup (2005-2007)

3.1 Transmitter

Throughout the development of the TMW lidar system (2005-present), the laser transmitter, telescope, and the data acquisition system have remained essentially unchanged. The laser is a high pulse energy Nd:YAG laser (Continuum) operating at 355 nm at a repetition rate of 10 Hz. While specified to produce more than 900 mJ/pulse, in long-term operation it consistently provides around 650 mJ/pulse. The beam is expanded 7.5x using a refractive beam-expander (CVI) before being steered into the atmosphere. Alignment of the laser to the receiving telescope is performed automatically and the program and mechanics of this part of the system has been fully described by Aspey et al (2008). Figure 1 of this reference also shows the basic arrangement of the lidar transmitter and receiver telescope.

3.2 Receiver

TMW makes use of the vibrational Raman shift at 2330 cm^{-1} for nitrogen and 3654 cm^{-1} for water vapor, as mentioned in **section 2**. For an emission at 355 nm, this translates to receiving at 387 nm and 407 nm respectively. The primary telescope is a 91 cm-diameter Newtonian telescope for the far-range channels supplemented by three (later four) 7.5 cm-diameter refractive telescopes for the near range. The small receivers are each designed to detect only a single wavelength and, as shown in **figure 1 (top)**, comprise a 7.5 cm fused silica lens (L1), a variable diaphragm field stop (FS), collimating and refocusing lenses sandwiching an interference filter (one each for 355 nm [F3], 387 nm [F1], and 407 nm [F2]). The received light is focused directly onto the photomultiplier (PMT) detector modules (P). Other than the addition of a fourth channel (very low intensity 355 nm for tropospheric aerosols information), the small receivers setup was basically unchanged since the lidar first light in 2005.

Coupling of the large telescope to the receiver polychromator is one of the arrangements that was changed as the lidar evolved. These changes will be described in **sections 5.2** and **5.4**. A diagram of the initial design of the channels coupled out of the large telescope is shown in **figure 1 (bottom)**. Light was sent from the telescope into the polychromator using a 1 mm diameter UV grade optical fiber (CeramOptec). The 0.22 numerical aperture of the fiber was matched to the f/2.7 focal ratio of the telescope and the diameter of the fiber also acted as a field-stop defining the field-of-view (fov) as 400 μrad. There were several reasons for choosing to use fiber coupling. It was not clear at the outset whether it would be required to install a chopper to counter signal-induced-noise. Should this have proved necessary, it would have been much easier to implement with a fiber. Also, we have successfully used similar fibers in other lidar systems (e.g., McDermid et al., 1999) and they minimize obscuration in the Newtonian configuration. Referring to **figure 1 (bottom)**, the light entering the polychromator from the fiber is collimated and then the first dichroic beamsplitter (B1) transmits 387 nm while reflecting

407 nm and 355 nm. An interference filter (F1) further selects only 387 nm and then the signal is divided using an anti-reflection coated fused silica plate (B3) into two parts with intensity ratio approximately 99:1. Using these 99:1 dual channels allows signal non-linearities, i.e. saturation, to be corrected and dynamic range to be extended. Similarly, beamsplitter (B2) reflects 407 nm and transmits the remaining 355 nm. As for the 387 nm channels, an interference filter (F2) is used to select only 407 nm and the signal is divided (B3) into high and medium intensity channels. Interference filter (F3) transmits only 355 nm onto a gated PMT.

In the original configuration, all 8 channels (5 channels from the large telescope and 3 from the small telescopes) utilized the same type PMT modules (Hamamatsu H5783P) except for the 355 nm high-intensity channel which used a bare R7400 PMT (Hamamatsu) in a home-built electronically gated base. This PMT was electronically gated to turn on at about 20 km in order to suppress some of the saturation and signal-induced-noise caused by the very high intensity Rayleigh returns from lower altitudes. All PMT signals were input to a series of photon counting multi-channel-scalers (Licel Transient Recorder). The Licel units, triggered at 10 Hz, collect the signal in 16384 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.

4 Data Analysis

4.1 Profile Retrieval

The signals from all 3 pairs of 387-407 nm channels are analyzed for the retrieval of water vapor (3-20 km) by our in-house lidar analysis software LidAna, version 7.0. This version is the latest optimized version 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, Manuscript in preparation).

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 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 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 **figure B2** 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 -50dB 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% up to 30%. Vertical resolution is reported in the data files following the definition of the cutoff frequency of the filter. **Figure 2** illustrates the effect of the signal filtering expressed as a function of the Kaiser filter's cutoff frequency and number of coefficients (top), and expressed as a function of the full-width at half-max (FWHM) of an Impulse Response (Dirac's Delta function) (bottom). 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 **Figure 3** and **table 1** for a 2-hour 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.

4.2 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 radiometer. 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 4** shows a schematic of each of the four methods for a 2-hour-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. “Matching Method 1” (top-left) consists of scaling, for each altitude bin, the 2-hour lidar average to the instantaneous radiosonde measurement. For “Matching Method 2” (bottom-left), the averaging time window is restricted to the duration of the radiosonde flight below 10 km, which is approximately 30 minutes. For “Matching Method 3”, the entire 2 hour window is used, but only the altitude points where water vapor variability over the 2-hour period was less than 20% are used. Finally, in “Matching Method 4” (bottom-right), 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 **figure 5**. 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 **figure 5** 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 (Oct 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 **figure 6**.

5 Results from the Initial Configuration (2005-2006)

The configuration described in **section 3.2** was used in 2005 and 2006 up to, and including the first MOHAVE campaign (October 2006). Data were acquired typically during two hours at the beginning of the night and simultaneously with the two other operating lidars on site (tropospheric ozone lidar, and stratospheric ozone/aerosols/temperature lidar). For calibration, the JPL lidar team uses its own Vaisala RS92 radiosonde station. One radiosonde was launched from TMF on each lidar measurement night. Typically the launch time is chosen to coincide with the first hour of the lidar measurement. More than 200 2-hour-long lidar profiles were analyzed during the June 2005 – October 2006 period. **Figure 7** shows the mean water vapor profile obtained simultaneously by the lidar and RS92 radiosondes during this period. Good agreement between the lidar and the radiosonde is found up to about 9-10 km, with an increasing difference is observed as we approach and cross the tropopause, the RS92 being too dry and/or the lidar being too wet. The excessive dryness of the RS92 in the upper troposphere has been known and well documented for many years (Miloshevich et al., 2004; 2009). However, the possible wet

bias of the lidar in the UTLS could not be confirmed until the lidar profiles could be compared to accurate measurements in this region.

These comparisons with accurate measurements took place during the Measurements Of Humidity in the Atmosphere and Validation Experiments (MOHAVE) campaign held in October 2006 with the primary objective of validating the Raman water vapor lidar measurement in the UTLS. Fifty RS92 radiosondes and 10 Cryogenic Frost-point Hygrometers (CFH) (Vömel et al., 2007) were launched from TMF during the 15 nights of the campaign. Two mobile lidars from the Goddard Space Flight Center (GSFC): Aerosol and Temperature Lidar (AT-Lidar) (McGee et al., 1991), and Scanning Raman Lidar (SRL) (Whiteman et al., 2006), were deployed at TMF for the campaign. Each of the three collocated water vapor lidars acquired over 150 hours of measurements (2 to 10 hours per night). At least one balloon per night was launched, each payload including one or two RS92. Ten payloads also contained a CFH and an ozonesonde. The lidar data corresponding to the first hour after launch were systematically processed and compared with the balloon measurements. In these 1-hour profiles, thin layered structures were well captured by all instruments but the lidar profiles begin to get noisy above 12 km (Leblanc et al., 2008b). With CFH taken as the reference profile, a dry bias was found in the upper troposphere for the RS92 and a wet bias was found for all three lidars, even though the lidars showed excellent agreement with each other (Leblanc et al., 2008b). After investigating the possible sources of the lidar wet bias and running additional test experiments, signal contamination by fluorescence in the lidar receivers was identified, which led all three lidar teams to modify their instrument configuration. For the JPL lidar, a 355 nm blocking filter was temporarily installed in front of the fiber optic. The instrument acquired 3 profiles simultaneously with a CFH in this configuration. The difference between the CFH and lidar profiles during these three flights was compared to that before the modification. The results are plotted in **figure 8** (left and center plots respectively). As anticipated, the wet bias completely disappeared after the modification (center plot). The nature of the fluorescence contamination was demonstrated by applying a physically-based empirical correction to the contaminated signals. The profiles retrieved from the corrected signals are shown in **figure 8** (right). The identification and removal of signal contamination by fluorescence was a major milestone in the TMF water vapor lidar project, and a major achievement of the MOHAVE campaign (Leblanc et al., 2008b)

6 Instrument Optimization and Validation

6.1 Receiver Second Configuration (2007-2009)

After MOHAVE, the front-end of the lidar receiver was redesigned to permanently suppress the fluorescence identified in the fiber optic during the campaign. This modification required custom-designed mechanical and optical components and was fully implemented in July 2007. Between MOHAVE and July 2007, only minor modifications were made (including sensitivity tests to fluorescence of several fiber optics). Routine operation continued alternately with and without fluorescence contamination. Signals knowingly acquired with contamination were empirically corrected to remove the effect of the fluorescence. However, because of the added uncertainties, the resulting corrected profiles are flagged and do not contribute to our long-term monitoring program. The new receiver configuration is shown in **Figure 9** (far-range channels only) and was fully implemented in July 2007. The strong 355 nm Rayleigh signal was

redirected out of the main optical path ahead of the focus of the large telescope using a long-wave pass beamsplitter (B4) that reflected 355 nm and transmitted 387 nm and 407 nm. An additional 355 nm blocking filter (F4) was placed ahead of the fiber connecting to the polychromator. A separate fiber was used to transmit the 355 nm to the gated PMT. This arrangement successfully eliminated fluorescence in the 387 nm and 407 nm signals (Leblanc et al., 2008b). However, the new and much stronger spectral selectivity resulted in an overall loss of signal by a factor of two, causing a decrease of 3 km (typically from 17 km to 14 km for 1-hour integration) of the uppermost altitude of the instrument.

6.2 The MOHAVE-II Campaign (October 2007)

To evaluate the effectiveness of the system modifications to remove fluorescence, a second MOHAVE campaign was carried out in October 2007. Again, two mobile lidars from GSFC were deployed for MOHAVE-II. However, the SRL system having sustained major damage during transportation for MOHAVE-I, was replaced by the Atmospheric Laboratory for Validation, Interagency Collaboration, and Education (ALVICE) trailer. ALVICE contains an upgraded version of the Raman Airborne Spectroscopic Lidar (RASL) adapted for ground-based measurements (Whiteman et al, 2007, 2010). The AT lidar had only minor modifications. MOHAVE-II was implemented following operational principles similar to that of MOHAVE-I (Leblanc et al., 2008b). **Figure 10** shows the average of the 10 profiles measured simultaneously by all participating instruments and techniques during the campaign. All the lidars agreed well with the CFH up to 12 km. Though the lidar profiles appear noisier, partly due to the signal decrease mentioned above, no wet bias appears anymore between the JPL lidar (magenta curve) and the CFH (green). On average, a 1-hour lidar integration reached ~15-16 km for the JPL system. A dry bias was still present on the uncorrected RS92 profile (red). This figure shows that the measurement accuracy of the Raman lidar is better than that of the RS92 in the UTLS. The CFH measurement remains the best quality among all instruments but its cost per profile is much higher (>$3000 per CFH launch) than that of the lidar and clearly prohibitive for long-term, routine monitoring, i.e. multiple measurements per week.

6.3 Receiver Current Configuration (2009-Present)

While the installation of the additional optics to divert the 355 nm returns into a separate fiber eliminated the fluorescence contamination and brought the lidar profiles into agreement with the CFH, the signal levels, and consequently the altitude range of the measurement were reduced. Simulation of the expected lidar returns and comparison with other systems suggested that significant signal was being lost somewhere in our receiver; most likely coupling in and out of the fibers. To try to recover the signal level and potentially increase it, it was decided to eliminate the fibers completely and to reposition the receiver at the telescope Newtonian focus.

Figure 11 is a schematic of the current receiver configuration (far-range channels only). To make the system more compact a new 387/407 nm beamsplitter (B1') operating at 45° angle of incidence was acquired to replace the previous one that operated at 30°. All of the other optical components were the same as described previously except for the field stops (FS). The new field stops were calibrated, ring-activated iris diaphragms that enabled the field-of-view to be varied as opposed to the fixed fov defined by the fiber diameter in the earlier configurations. This flexibility provided better optimization of laser/telescope overlap and sky background rejection. Values of 600 μrad and 800 μrad were finally selected for the Raman and 355H channels respectively. The various components of the receiver were aligned on top of each other to

minimize obscuration of the large telescope and a calibration lamp was permanently mounted to the side of the receiver module. The purpose of this lamp is discussed in **section 8.2**.

One other change in this final configuration was to the 407H photomultiplier detector. The original Hamamatsu H5783P module was replaced with a H10682-110 photon counting head device. Within this module, the PMT high voltage supply and discriminator level are factory preset to optimum values. Also, the PMT in this device has a super-bialkali (SBA) photocathode which has higher quantum efficiency than standard PMTs; 33% at 407 nm compared to 25% (eventually, when commercially available, ultra-biakali (UBA) photocathodes will further raise the QE to ~40%). The net result of all these changes was an order of magnitude increase in signal levels. However, such major changes required re-evaluation/validation and therefore another campaign, MOHAVE-2009 was organized.

7 The MOHAVE-2009 Campaign (October 2009)

The modifications described above occurred in the summer of 2009 (July, August and September). No further changes occurred after that and the instrument remained untouched throughout the MOHAVE-2009 campaign with the exception of the calibration lamp, which purpose is described in the next section. This third MOHAVE campaign was by far the largest of all three MOHAVE campaigns with many participating instruments: 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 in 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 October 11 and October 27, 2009. A overview of the campaign operations and results is presented in Leblanc et al. (2011; this issue). 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 hours 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, this issue), Toon et al. (2011, Manuscript in preparation), Hurst et al. (2011b, this issue), Whiteman et al. (2011, this issue), and McGee et al. (2011, Manuscript in preparation) respectively.

Water vapor has different scales of natural variability in the troposphere and stratosphere (see for example figure 8 of Leblanc et al. (2011)). For this reason the coincidence criteria used for the comparisons presented here are different above and below 14 km: all profiles coinciding within 1 hour and 100 km were used for altitudes below 14 km, and all profiles coinciding within 6 hours and 250 km were used for all altitudes above 14 km.

Figure 12 shows four examples of water vapor profiles measured simultaneously during the campaign. In all plots, the lidar profiles (TMW, ALVICE and STROZ) were obtained from a 1-hour integration starting at the corresponding balloon launch time. The CFH campaign-mean profile is provided on each plot for reference purposes only. October 18 and 20 (top-left and top-

right) illustrate the very large natural variability in the troposphere from one night to another: very humid upper troposphere on Oct. 18 and very dry on Oct. 20. Most of the very fine vertical structures (sometimes less than 1-km deep) are well captured by all balloon-borne and lidar instruments. A well-known RS92 dry bias is again systematically observed above 12 km (cyan curves). The RS92 profiles corrected following the method described by Miloshevich et al. (2009) show a better agreement with lidar and CFH above 13 km. The difference between corrected and uncorrected profiles can reach 35% and is reviewed in Leblanc et al. (2011, this issue). The bottom-left and bottom-right plots of **figure 12** show profiles measured on the night October 27. Besides showing very large vertical changes (from 200 ppmv to 2000 ppmv in less than 2 km), the two figures illustrate the very high temporal variability on short timescales, namely a change at 4 km altitude from 200 ppmv to 4000 ppmv in a 5-hour interval. The very fine and fast-changing structures observed on all plots of **figure 12** confirm the need for extra care when producing comparisons between the various instruments and techniques.

Figure 13 shows the campaign average of all TMW and all RS92 relative humidity (RH) profiles measured simultaneously (i.e., within 1-hour of balloon launch), and their mean difference. The comparisons are shown for the RS92 uncorrected (top row) and corrected (bottom row) 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 (figure 5, bottom-right). 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 14 (top row) shows the campaign average of all coincident water vapor profiles (left) measured by TMW and Aura MLS (version 3), and their difference (right). 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** of **figure 14** shows the campaign average of all coincident water vapor profiles (left) measured by TMW and Aqua-AIRS, and their difference (right). 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.

Finally the campaign-mean profiles measured simultaneously by the TMW lidar and CFH, and their difference, are plotted in **figure 15** as a function of pressure **(a-d)**, and as a function of geometric altitude **(e-f)**. Since the CFH is the most accurate instrument in the UTLS participating to the MOHAVE-2009 campaign, this figure is indeed the most important result of the campaign for the TMW lidar and it summarizes well the performance achieved by the TMW lidar since MOHAVE-2009. Geopotential height is computed from the simultaneous RS92 pressure and temperature data, and then converted to geometric height. Below 14 km (approx. 100 hPa), only the strictly coincident profiles were used to compute the mean (i.e., 12 CFH flights and the 12 corresponding 1-hour integrated lidar profiles). Using the 1-hour integrated profiles prevents the

lidar reaching altitudes above 14 km. In the UTLS, 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 close look at the top-left plot of **figure 15** (UTLS, 15-20 km or 200-50 hPa) reveals that the standard deviation of the lidar measurements above 17 km (approx. 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 to two CFH profiles yields a precision in the UTLS equivalent to integrating the lidar measurements for 5 to 8 full nights (i.e., 40 to 70 hours). That is, one weekly CFH launch yields a precision in the UTLS similar to that of the integrated lidar measurement of 4 full nights per week. Similarly, 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-hour-long measurements per week. These estimations therefore show that, with our current target to run TMW for two hours per night, 4 nights per week, we can achieve a precision in the UTLS equivalent to that achieved when launching one CFH per month

8 Additional Considerations for Long-Term Monitoring

8.1 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 (Whiteman et al., 2006). 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 surface measurements obtained every 5 minutes from a co-located Vaisala Weather Station (MAWS-1). The scaling process is illustrated in **figure 16** with the same coloring conventions as **figure 4** (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 surface measurements by the Vaisala MAWS. 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. 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 have occurred. 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 meters 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. However, it has potentially 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).

8.2 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 solution to this problem is to launch not one, but several radiosondes during the same lidar experiment (typically four radiosondes for a 2-hour long lidar experiment). Though radiosondes are affordable such an intensive launch plan is too expensive and therefore not cost effective 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 minutes 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

605 constant is detected between absolute calibration campaigns, then radiosondes must be launched
606 immediately to quantify the impact of these changes on the absolute calibration constant. The
607 Hybrid method is described in details in Leblanc and McDermid (2008). Though radiosondes are
608 used in the hybrid method described here, any source of accurate measurement may be used for
609 the absolute calibration campaigns.

610 Two lamps (200W and 45W Quartz-Tungsten Halogen) have been used since the hybrid method
611 was introduced in October 2007. A complete review of the signal ratios of the 387 and 407 nm
612 channels obtained during the routine lamp runs is presented in **figure 17**. Three periods are
613 presented, which corresponds to the different receiver configurations already discussed. The top
614 figure shows the signal ratios for all three ranges (high-, mid- and low-intensity) between
615 October 2007 (MOHAVE-II, when the lamp runs started) and summer 2009. This figure is
616 indeed an extension of figure 8 of Leblanc and McDermid (2008), in which the observed jumps
617 and drops in the channel ratios until April 2008 are discussed. These features will not be
618 discussed again here, though the instrumentation changes are noted on the figure. The second
619 plot (middle) highlights the major configuration changes made between June and October 2009:
620 receiver redesign, field stop optimization, and most importantly, lamp change (from 200W to
621 45W) during MOHAVE-2009. The third period (bottom plot), extending from MOHAVE-2009
622 to present time, shows the channel ratios since the 45W lamp has been in operation. No
623 instrumentation change was made throughout this period. Inspection of all three plots leads to
624 several important conclusions. First, and as anticipated in Leblanc and McDermid (2008), the
625 channel ratios obtained during lamp runs remain very stable over time (standard deviation mostly
626 below 1.5% over timescales of a year) unless an instrumentation change occurs. Second, no
627 apparent drift is observed at these timescales with the exception of the mid-range channel ratio
628 during period 3, which shows a 5% drop in late spring 2010 not associated with any
629 instrumentation change. Third and most importantly, all observed standard deviations as well as
630 the spring 2010 mid-range channel ratio drop are closely related to the magnitude of the lamp
631 irradiance with respect to the magnitude of these channels' background noise. During period 1,
632 the lamp is brighter (200 W), and provides channel illumination far above all channels' internal
633 noise. The observed standard deviations in this case reflect only the spectral stability of the lamp
634 irradiance, i.e., about 1.5% (irradiance ratio). During period 3, the lamp is dimmer (45W), and
635 only the standard deviation of the ratio of the high-intensity channels reflects the lamp spectral
636 stability. The standard deviation of the other two ranges reflects a combination of the lamp
637 spectral stability and non-negligible internal background noise. Indeed the mid-range ratio drop
638 in spring 2010 is a direct consequence of the mean background noise reduction in the 407 nm
639 channel. This apparent reduction was associated with the presence (before spring 2010), then
640 absence (after spring 2010) of undesired isolated spikes in the background noise. Due to of its
641 subtle nature, this feature was never investigated and there is no explanation for it today.

642 During MOHAVE-2009, the 200W lamp initially used was replaced by a dimmer one because
643 the signals acquired by the high-intensity channels were potentially affected by saturation. This
644 hypothesis was indeed verified as the change in the channel ratio associated with the lamp
645 replacement turned out to be very different from those observed on the other two ranges (ratio
646 multiplied by 4 for the high intensity range as opposed to a 20% increase for the other two
647 ranges, see middle plot of **figure 17**). This finding therefore leads to an important conclusion
648 regarding the general use of a lamp: neither 200W nor 45W is currently appropriate to obtain an
649 optimized channel ratio for all ranges. In our case, we can (and shall) safely use the Hybrid
650 technique for the mid- and low-intensity ranges only before MOHAVE-2009, and we can (and

651 shall) safely use it for the high-intensity range only after MOHAVE-2009. Ideas for an improved
652 methodology that could be used simultaneously for all ranges are currently being explored.

653

654 ## 9 Conclusion

655 A high capability water vapor Raman lidar was developed at the Jet Propulsion Laboratory Table
656 Mountain Facility (California) with the objective of making accurate, routine, long-term water
657 vapor measurements in the UTLS. The system, referred to as "TMW', was built in 2005 and
658 optimized over the following four years.

659 Because the overall goal of the TMW instrument is to provide long-term monitoring, several
660 additional considerations (compared to traditional water vapor Raman lidar) were undertaken to
661 insure optimal lidar calibration stability with time. Besides radiosonde, an alternative source of
662 calibration is now systematically used, namely the TPW measurements from a co-located GPS
663 receiver. A third calibration source, namely the TPW measurements from a 22-GHz microwave
664 radiometer (Nedoluha et al., 2011), is also available and will be used as another alternate source.
665 Furthermore, a dedicated calibration method combining a laboratory lamp and radiosondes, and
666 referred to as "hybrid calibration" method, is now used to minimize the cost of launching
667 radiosondes and increase the accuracy and stability of the absolute calibration.

668 To support and achieve a full optimization, the MOHAVE-I, MOHAVE-II, and MOHAVE-2009
669 validation campaigns took place in October 2006, 2007 and 2009 respectively. During these
670 campaigns the lidar measurements were extensively compared to operational radiosonde (Vaisala
671 RS92) and Research-grade balloon-borne instruments (CFH)

672 The MOHAVE-I campaign revealed the presence of fluorescence contaminating the lidar signal
673 in its upper part (10-20 km). After a first major receiver configuration change in 2007, the
674 MOHAVE-II campaign confirmed the removal of the contaminating fluorescence, but also
675 revealed a substantial loss of signal resulting from this change. After a second major
676 configuration change (summer 2009), the MOHAVE-2009 campaign finally showed that the
677 TMW lidar was brought up to its initial performance expectations, i.e., capable of measuring
678 water vapor mixing ratio in the UTLS with a precision of 10% or better between 10 km and 15
679 km, and reaching 21-22 km with a degraded precision and vertical resolution.

680 The latest comparisons with CFH measurements (MOHAVE-2009) showed excellent agreement
681 throughout the troposphere and UTLS, with mean biases well below the reported measurements
682 uncertainties. The TMW lidar shows a 2% mean dry bias with CFH in the lower troposphere, a
683 3% mean wet bias in the UTLS, and virtually no bias with the corrected RS92 measurements
684 throughout the troposphere (ground to 18 km).

685 A closer look at the lidar measurement standard deviations above 16 km during MOHAVE-2009
686 shows values two to three times larger than that of the CFH. At these altitudes the lidar signals
687 are noise limited and the standard deviation values reflect the precision of the measurements,
688 which therefore provides a good indication of the number of additional lidar measurements
689 needed to achieve the same precision as CFH. It was estimated that one monthly CFH profile
690 yields a precision in the UTLS equivalent to that of the integrated lidar measurement of one full
691 night per week, or four 2-hour-long measurements per week. These estimations show that, with
692 our present target of routinely running TMW two hours per night, 4 nights per week, we can
693 achieve measurements with a precision in the UTLS equivalent to that achieved if launching one

CFH per month. Though the TMW lidar is now fully optimized, its overall power-aperture product is well below most contemporary optimums. For example, commercial laser powers today can easily exceed 10 W at 355 nm as opposed to 6.5 W produced by our decade-old laser. Therefore it is not unlikely that a future laser upgrade, together with the upgrade of other components, will bring the precision of the TMW lidar closer to that of the CFH, thus requiring less laser running time to achieve similar levels of precision in the UTLS. Nevertheless, 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, to be released 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

Aspey, R. A., McDermid, I. S., Leblanc, T., Howe, J. W., and Walsh, T. D.: LABVIEW graphical user interface for precision multichannel alignment of Raman lidar at Jet Propulsion Laboratory, Table Mountain Facility, Rev. Sci. Instr., 79, 094502, doi:10.1063/1.2976672, 2008.

Cooney, J.: Remote Measurements of Atmospheric Water Vapor Profiles Using the Raman Component of Laser Backscatter, Journal of Applied Meteorology, 9, 182-184, 10.1175/1520-0450(1970)009<0182:rmoawv>2.0.co;2, 1970.

Donovan, D. P., Whiteway, J. A., and Carswell, A. I.: Correction for nonlinear photon-counting effects in lidar systems, Appl. Opt., 32, 6742-6753, 1993.

de F. Forster, P. M., and Shine, K. P.: Stratospheric water vapour changes as a possible contributor to observed stratospheric cooling, Geophys. Res. Lett., 26, 3309-3312, 10.1029/1999gl010487, 1999.

Hinkley, E.: *Topics in Applied Physics*, vol. 14, chap. Laser monitoring of the atmosphere, p. 380pp, Springer, New York, 1976.

Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, J. Geophys. Res., 116, D02306, 10.1029/2010jd015065, 2011.

Kaiser, J. F., and Reed, W. A.: Data Smoothing Using Low-Pass Digital-Filters, Rev. Sci. Instr., 48, 1447-1457, 1977.

Leblanc, T., McDermid, I. S., and Aspey, R. A.: First-Year Operation of a New Water Vapor Raman Lidar at the JPL Table Mountain Facility, California, J. Atmos. Ocean. Tech., 25, 1454-1462, doi:10.1175/2007JTECHA978.1, 2008a.

Leblanc, T., McDermid, I. S., McGee, T. G., Twigg, L., Sumnicht, G., Whiteman, D. N., Rush, K., Cadirola, M., Venable, D., Connell, R., Demoz, B., Vömel, H., and Miloshevich, L.: Measurements of humidity in the atmosphere and validation experiments (MOHAVE, MOHAVE II): Results overview, Reviewed and Revised Papers of The 24th International Laser Radar Conference, Boulder, CO, 23-27 June 2008, 1013-1016, 2008b.

Leblanc, T., and McDermid, I. S.: Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements, Appl. Opt., 47, 5592-5603, 2008.

Leblanc, T., Walsh, T. D., McDermid, I. S., Toon, G. C., Blavier, J. F., Haines, B., Read, W. G., Herman, B., Fetzer, E., Sander, S., Pongetti, T., Whiteman, D. N., McGee, T. G., Twigg, L., Sumnicht, G., Venable, D., Calhoun, M., Dirisu, A., Hurst, D., Jordan, A., Hall, E., Miloshevich, L., Vömel, H., Straub, C., Kampfer, N., Nedoluha, G. E., Gomez, R. M., Holub, K., Gutman, S., Braun, J., Vanhove, T., Stiller, G., and Hauchecorne, A.: Measurements of Humidity in the Atmosphere and Validation Experiments (MOHAVE)-2009: overview of campaign operations and results, Atmos. Meas. Tech. Discuss., 4, 3277-3336, 10.5194/amtd-4-3277-2011, 2011.

McDermid, I. S., Beyerle, G., Haner, D. A., and Leblanc, T.: Redesign and improved performance of the tropospheric ozone lidar at the Jet Propulsion Laboratory Table Mountain Facility, Appl. Opt., 41, 7550-7555, 2002.

McGee, T. J., Whiteman, D., Ferrare, R., Butler, J. J., and Burris, J. F.: STROZ LITE - STRatospheric OZone Lidar Trailer Experiment, Opt. Engin., 30, 31-39, 1991.

McGee, T. G., Twigg, L., and Sumnicht, G.: STROZ lidar measurements of ozone, temperature and water vapor during the MOHAVE 2009 campaign, Atmos. Meas. Technol. Disc., 4, Manuscript in Preparation, 2011.

Measures, R. M.: *Laser remote sensing: fundamentals and applications*, Krieger Pub. Co., Malabar, FL, 1992.

Melfi, S. H., Lawrence, J. D., and McCormic.Mp: Observation of Raman Scattering by Water Vapor in Atmosphere, Appl. Phys. Lett., 15, 295-&, 10.1063/1.1653005, 1969.

Miloshevich, L. M., Paukkunen, A., Vömel, H., and Oltmans, S. J.: Development and Validation of a Time-Lag Correction for Vaisala Radiosonde Humidity Measurements, J. Atmos. Ocean. Tech., 21, 1305-1327, doi:10.1175/1520-0426(2004)021<1305:DAVOAT>2.0.CO;2, 2004.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., and Leblanc, T.: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements, J. Geophys. Res., 114, D11305, 10.1029/2008jd011565, 2009.

Nedoluha, G. E., Gomez, R. M., Hicks, B. C., Helmboldt, J., Bevilacqua, R. M., and Lambert, A.: Ground-based microwave measurements of water vapor from the midstratosphere to the mesosphere, J. Geophys. Res., 116, D02309, 10.1029/2010jd014728, 2011

Oltmans, S. J., and Hofmann, D. J.: Increase in lower-stratospheric water-vapor at a mid-latitude Northern-Hemisphere site from 1981 to 1994, Nature, 374, 146-149, 1995.

Sherlock, V., Garnier, A., Hauchecorne, A., and Keckhut, P.: Implementation and Validation of a Raman Lidar Measurement of Middle and Upper Tropospheric Water Vapor, Appl. Opt., 38, 5838-5850, 1999a.

Sherlock, V., Hauchecorne, A., and Lenoble, J.: Methodology for the Independent Calibration of Raman Backscatter Water-Vapor Lidar Systems, Appl. Opt., 38, 5816-5837, 1999b.

Stiller, G. P., Kiefer, M., Eckert, E., von Clarmann, T., Kellmann, S., García-Comas, M., Funke, B., Leblanc, T., Fetzer, E., Froidevaux, L., Gomez, M., Hall, E., Hurst, D., Jordan, A., Kämpfer, N., Lambert, A., McDermid, I. S., McGee, T., Miloshevich, L., Nedoluha, G., Read, W., Schneider, M., Schwartz, M., Straub, C., Toon, G., Twigg, L. W., Walker, K., and Whiteman, D. N.: Validation of MIPAS IMK/IAA temperature, water vapor, and ozone profiles with MOHAVE-2009 campaign measurements, Atmos. Meas. Tech. Discuss., 4, 4403-4472, 10.5194/amtd-4-4403-2011, 2011.

Toon, G., et al.: The MkIV Infrared Fourier-Transform Spectrometer measurements during MOHAVE-2009, Atmos. Meas. Technol. Disc., 4, Manuscript in Preparation, 2011.

Turner, D. D., and Goldsmith, J. E. M.: Twenty-Four-Hour Raman Lidar Water Vapor Measurements during the Atmospheric Radiation Measurement Program's 1996 and 1997 Water Vapor Intensive Observation Periods, Journal of Atmospheric and Oceanic Technology, 16, 1062-1076, doi:10.1175/1520-0426(1999)016<1062:TFHRLW>2.0.CO;2, 1999.

Vaughan, G., Wareing, D. P., Thomas, L., and Mitev, V.: Humidity Measurements In The Free Troposphere Using Raman Backscatter, Q. J. Roy. Meteorol. Soc., 114, 1471-1484, 1988.

Vömel, H., David, D. E., and Smith, K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, J. Geophys. Res., 112, D08305, 10.1029/2006jd007224, 2007.

Whiteman, D. N.: Examination of the Traditional Raman Lidar Technique. II. Evaluating the Ratios for Water Vapor and Aerosols, Appl. Opt., 42, 2593-2608, 2003.

Whiteman, D. N., Demoz, B., Rush, K., Schwemmer, G., Gentry, B., Di Girolamo, P., Comer, J., Veselovskii, I., Evans, K., Melfi, S. H., Wang, Z., Cadirola, M., Mielke, B., Venable, D., and Van Hove, T.: Raman Lidar Measurements during the International H2O Project. Part I: Instrumentation and Analysis Techniques, Journal of Atmospheric and Oceanic Technology, 23, 157-169, doi:10.1175/JTECH1838.1, 2006.

Whiteman, D. N., Veselovskii, I., Cadirola, M., Rush, K., Comer, J., Potter, J. R., and Tola, R.: Demonstration measurements of water vapor, cirrus clouds, and carbon dioxide using a high-performance Raman lidar, Journal of Atmospheric and Oceanic Technology, 24, 1377-1388, 10.1175/jtech2058.1, 2007.

Whiteman, D. N., Rush, K., Rabenhorst, S., Welch, W., Cadirola, M., McIntire, G., Russo, F., Adam, M., Venable, D., Connell, R., Veselovskii, I., Forno, R., Mielke, B., Stein, B., Leblanc, T., McDermid, S., and Vomel, H.: Airborne and Ground-Based Measurements Using a High-Performance Raman Lidar, J. Atmos. Oceanic Technol., 27, 1781-1801, 10.1175/2010jtecha1391.1, 2010.

Whiteman, D. N., et al. (2011): Development of a Correction Technique for Raman Water Vapor Lidar Signal Dependent Wet Bias and Application to Upper Tropospheric Retrievals, Atmos. Meas. Technol. Disc., 4, Accepted, 2011.

Tables

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

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

* 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-hr-long samples over the period October 2007-April 2009 using 12 different calibration methods (single radiosonde coincidences). See text for details

Scaling / Matching	Method 1 (Mean)	Method 2 (Median)	Method 3 (Gaussian)
Method 1	16%	12%	10%

(all datasets)			
Method 2 (time-coincident)	8%	7%	8%
Method 3 (min. variability)	12%	11%	10%
Method 4 (time+altitude coincident)	11%	11%	11%

Figure Captions

Figure 1. TMW lidar receiver initial configuration, i.e., 2005-2007 for the five high- and mid-intensity channels, and 2005-present for the three low-intensity channels. A fourth low intensity channel, 355 nm extra-low intensity, was added in 2007.

Figure 2. Effect of the TMW lidar signals vertical filtering based upon two different vertical resolution definitions. Top plot: Definition based on the cut-off frequency of a digital filter (Kaiser filter with 50dB attenuation). Bottom plot: Definition based on the full-width at half-maximum (FWHM) of an Impulse Response to a Dirac Delta Function.

Figure 3. Example of the effect of vertical filtering on the TMW water vapor profiles. Left plot: Mixing ratio profiles for all three ranges (pink: low-intensity, green: mid-intensity, and blue: high-intensity range). Right plot: 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.

Figure 4. 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.

Figure 5. 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 (10/2007 to 4/2009). See text for details on each method, and Table 2 for a compilation of the standard deviations of all 12 methods.

Figure 6. An example of the construction of the final (black) water vapor (left) and Relative Humidity (right) 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.

Figure 7. Comparison of the climatological mixing ratio profiles measured over the period June 2005-October 2007 by the TMW lidar and simultaneous and co-located Vaisala RS92 radiosondes. Left plot: Mean profiles. Right plot: Mean difference, lidar measurement uncertainty and standard deviation.

Figure 8. Comparison of mean water vapor mixing ratio profiles measured simultaneously by the TMW lidar and CFH during MOHAVE-I. Left plot: Lidar signals contaminated by fluorescence (no 355 nm block). Center plot: Lidar signals free of fluorescence (355 nm block

859 installed). Right plot: Same as left plot, but after fluorescence correction is applied to lidar
860 signals.

861 **Figure 9.** TMW lidar receiver second configuration (2007-2009) for the high- and mid-intensity
862 channels.

863 **Figure 10.** Averaged water vapor mixing ratio computed from 10 profiles measured
864 simultaneously during MOHAVE-II by the TMW lidar, two other lidars, CFH, and Vaisala RS92
865 radiosondes.

866 **Figure 11.** TMW lidar receiver final configuration (2009-present) for the high- and mid-
867 intensity channels.

868 **Figure 12.** Example of four water vapor mixing ratio profiles measured simultaneously by
869 TMW, two other lidars, CFH, Aura-MLS, and RS92 radiosondes during MOHAVE-2009.

870 **Figure 13.** Comparison of the MOHAVE-2009 campaign-averaged relative humidity profiles
871 calculated from the 28 simultaneous profiles measured by the TMW lidar and Vaisala RS92
872 radiosondes. Top plots: mean profiles (left) and difference with the uncorrected RS92 data
873 (right). Bottom plots: mean profiles (left) and difference with the corrected (Miloshevich et al.,
874 2004; 2009) RS92 data (right).

875 **Figure 14.** Top: averaged mixing ratio profiles (left) calculated from the 3 coincident profiles
876 measured in the UTLS by the TMW lidar and Aura-MLS during MOHAVE-2009, and their
877 difference (right). Bottom: averaged mixing ratio profiles (left) calculated from the 4
878 simultaneous profiles measured by the TMW lidar and Aqua-AIRS during MOHAVE-2009, and
879 their difference (right).

880 **Figure 15.** Comparison of the MOHAVE-2009 campaign-averaged mixing ratio profiles (left)
881 calculated from all coincident profiles measured by the TMW lidar and the CFH, and their
882 difference (right). Top row (a-b): UTLS on a pressure grid; middle row (c-d): troposphere on a
883 pressure grid; bottom row (e-f): troposphere and lower stratosphere on a geometric height grid.
884 See text and figure for details on the coincidence criteria.

885 **Figure 16.** Schematics of the "stick-and-slide" method used to perform a calibration of the
886 TMW lidar profiles using co-located and simultaneous GPS Total Precipitable Water
887 measurements. Lidar data points are symbolized by red open rectangles, GPS column data by
888 blue vertical bars, and the actual column data pairs used for the normalization by filled green
889 vertical bars. The sliding uncalibrated lidar profile is shown in red, the constrained water vapor
890 ground-measurement in blue, and the "sticky" interpolated layer in green. See text for details.

891 **Figure 17.** Historical evolution of the 407 nm to 387 nm channel ratios (pink: low-intensity,
892 green: mid-intensity, and blue: high-intensity) between October 2007 and present. All
893 instrumentation changes are denoted by arrows and comments, including the change of lamp
894 (from 200W to 45W) in October 2009.

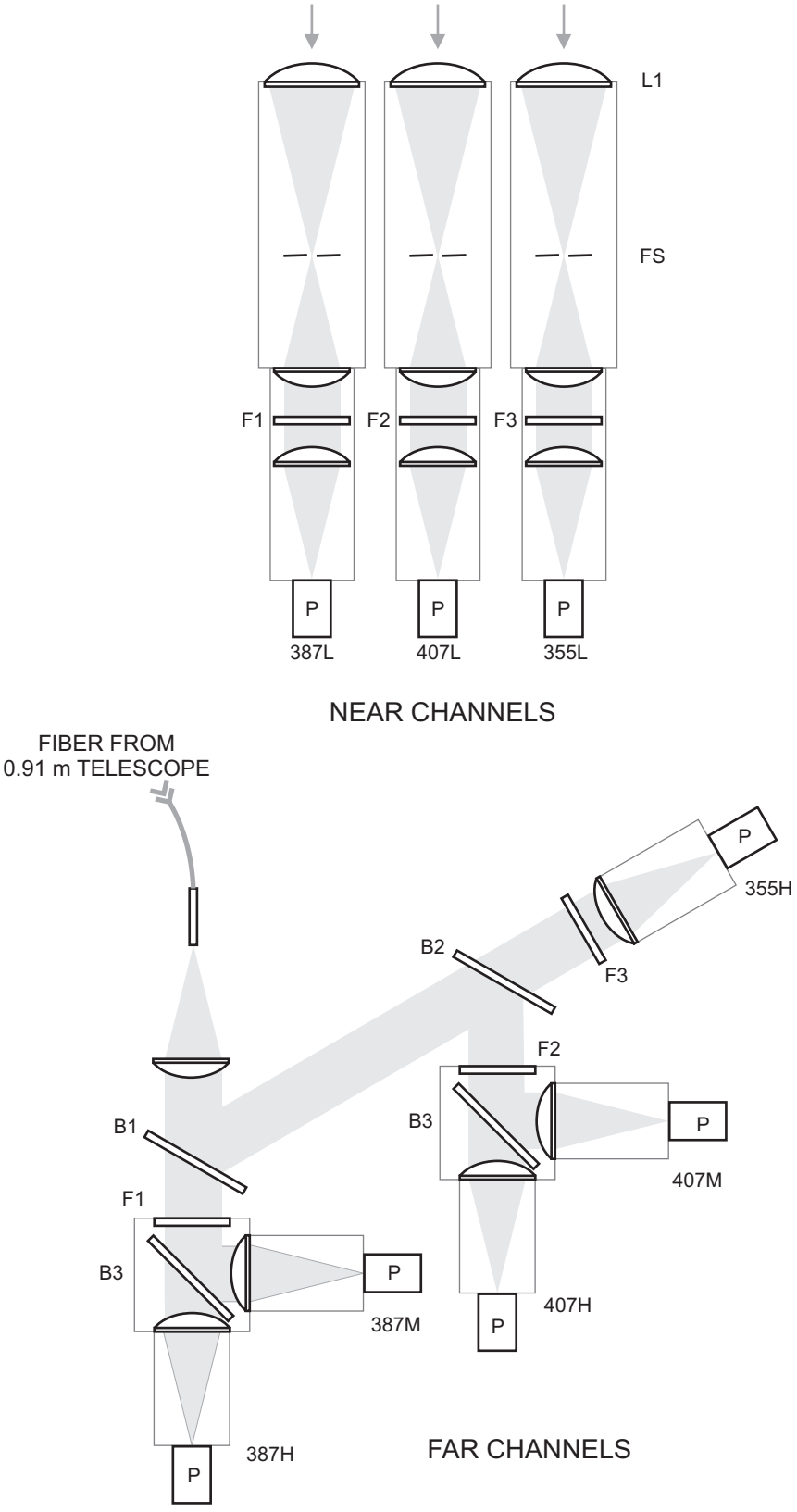

Figure 1. TMW lidar receiver initial configuration, i.e., 2005-2007 for the five high- and mid-intensity channels, and 2005-present for the three low-intensity channels. A fourth low intensity channel, 355 nm extra-low intensity (not shown here), was added in 2007

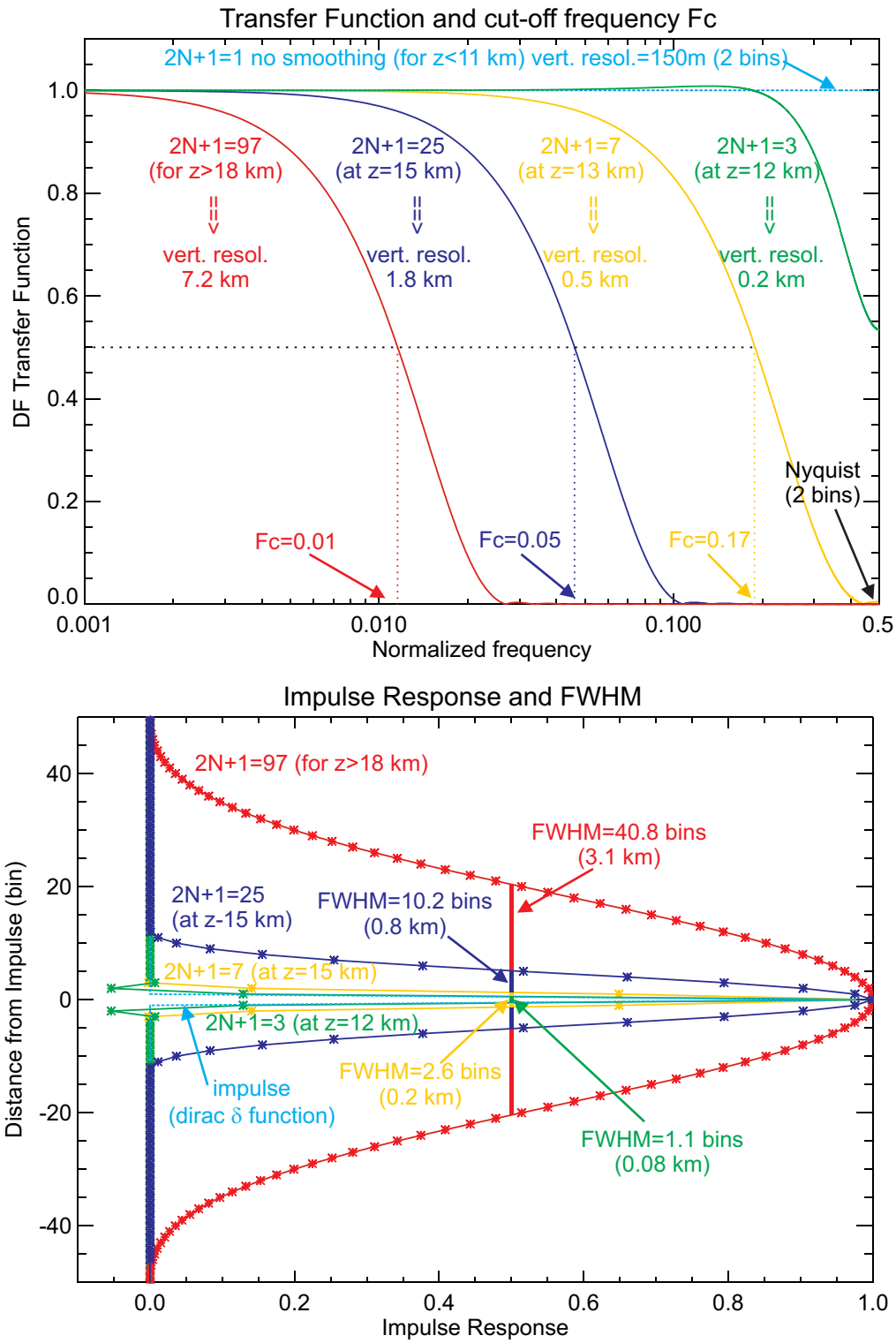

Figure 2. Effect of the TMW lidar signals vertical filtering based upon two different vertical resolution definitions. Top plot: Definition based on the cut-off frequency of a digital filter (Kaiser filter with 50dB attenuation). Bottom plot: Definition based on the full-width at half-maximum (FWHM) of an Impulse Response to a Dirac Delta Function

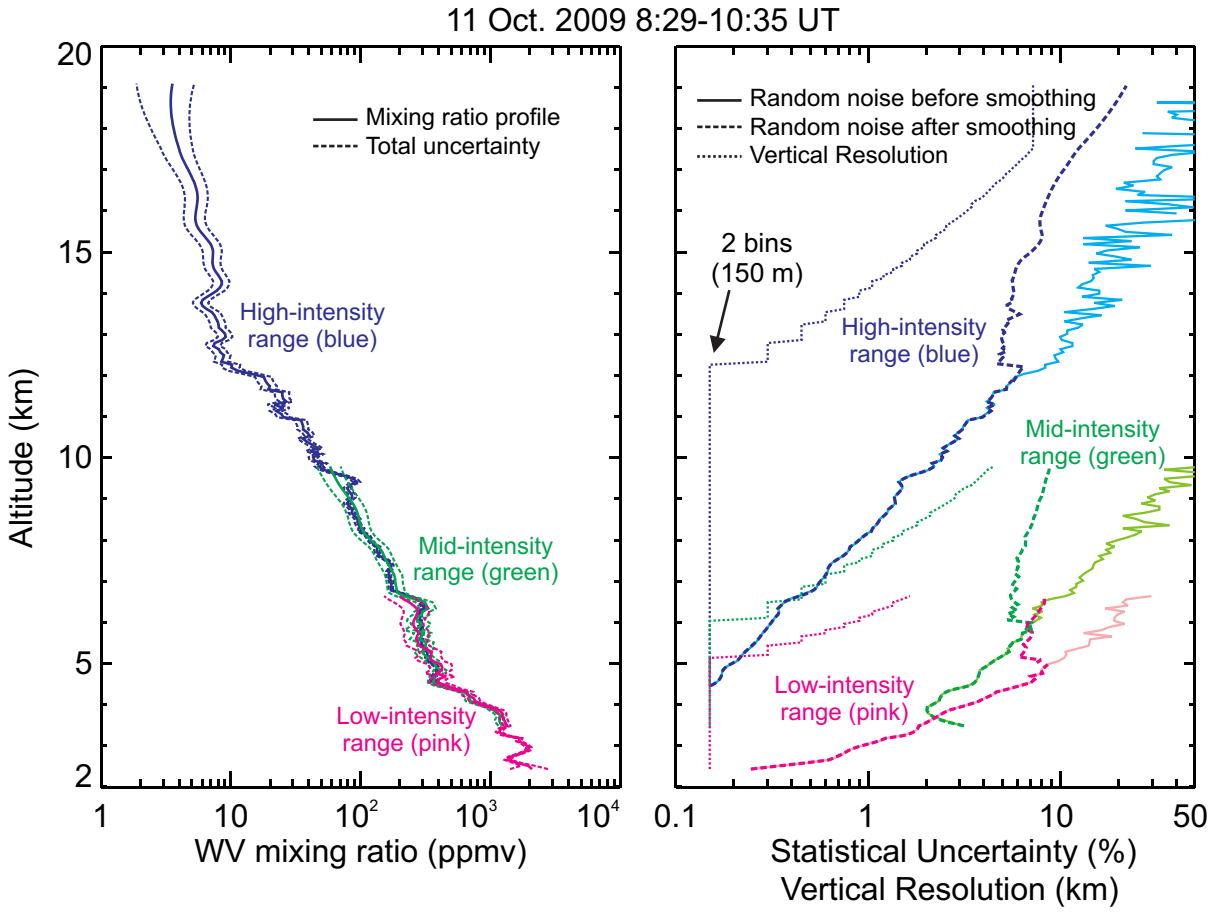

Figure 3. Example of the effect of vertical filtering on the TMW water vapor profiles. Left plot: Mixing ratio profiles for all three ranges (pink: low-intensity, green: mid-intensity, and blue: high-intensity range). Right plot: Statistical uncertainty (%) before (solid curves) and after (dashed curves) filtering, and corresponding vertical resolution (km) (dotted curves) following the definition based on digital filter cutoff frequency

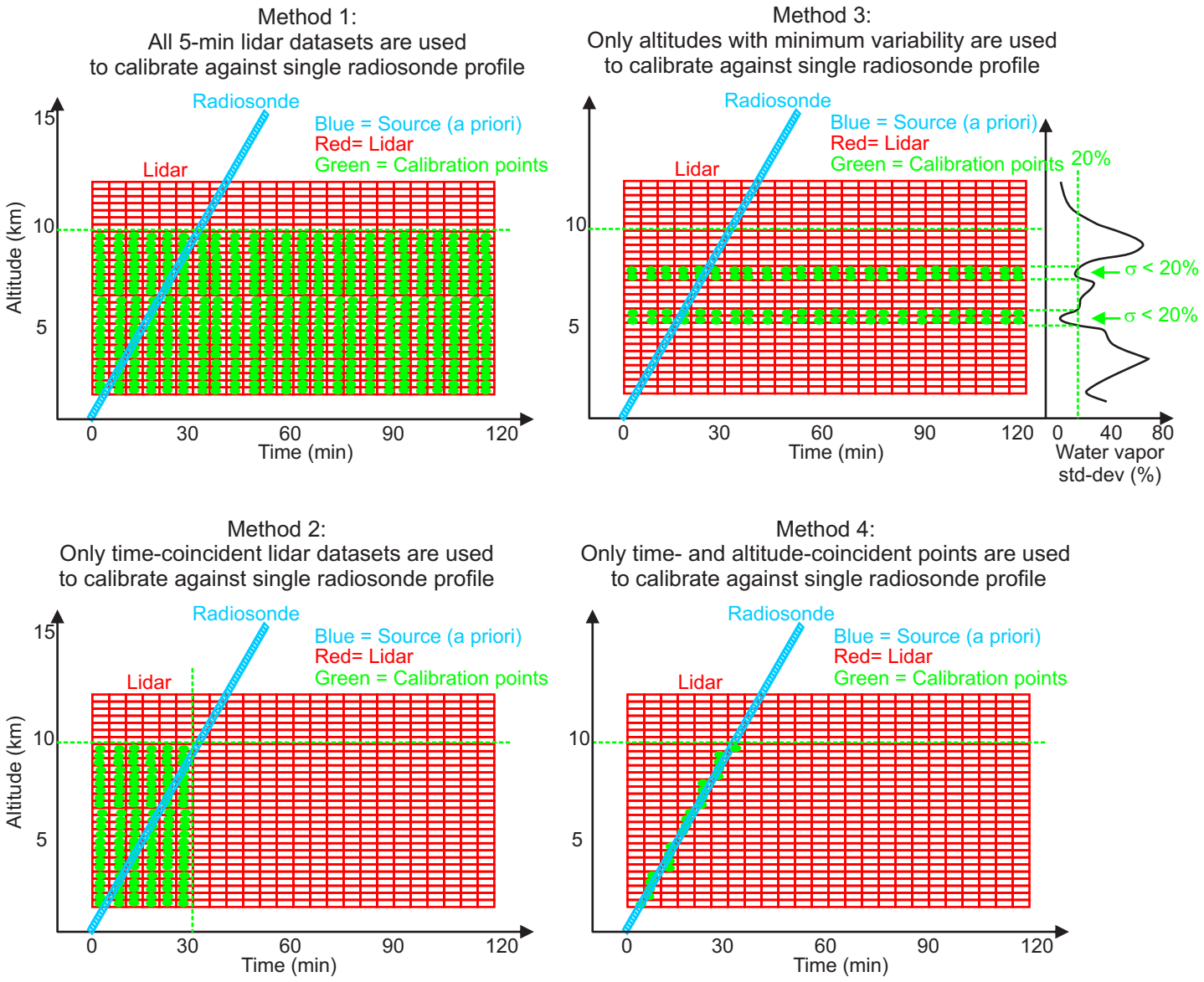

Figure 4. Schematics of the four methods tested to optimize the calibration of the TMW lidar profiles using co-located radiosondes. Lidar data points are symbolized by red open rectangles, radiosonde data points by small blue open squares tilted with time, and the actual data pairs used for the normalization by filled green circles

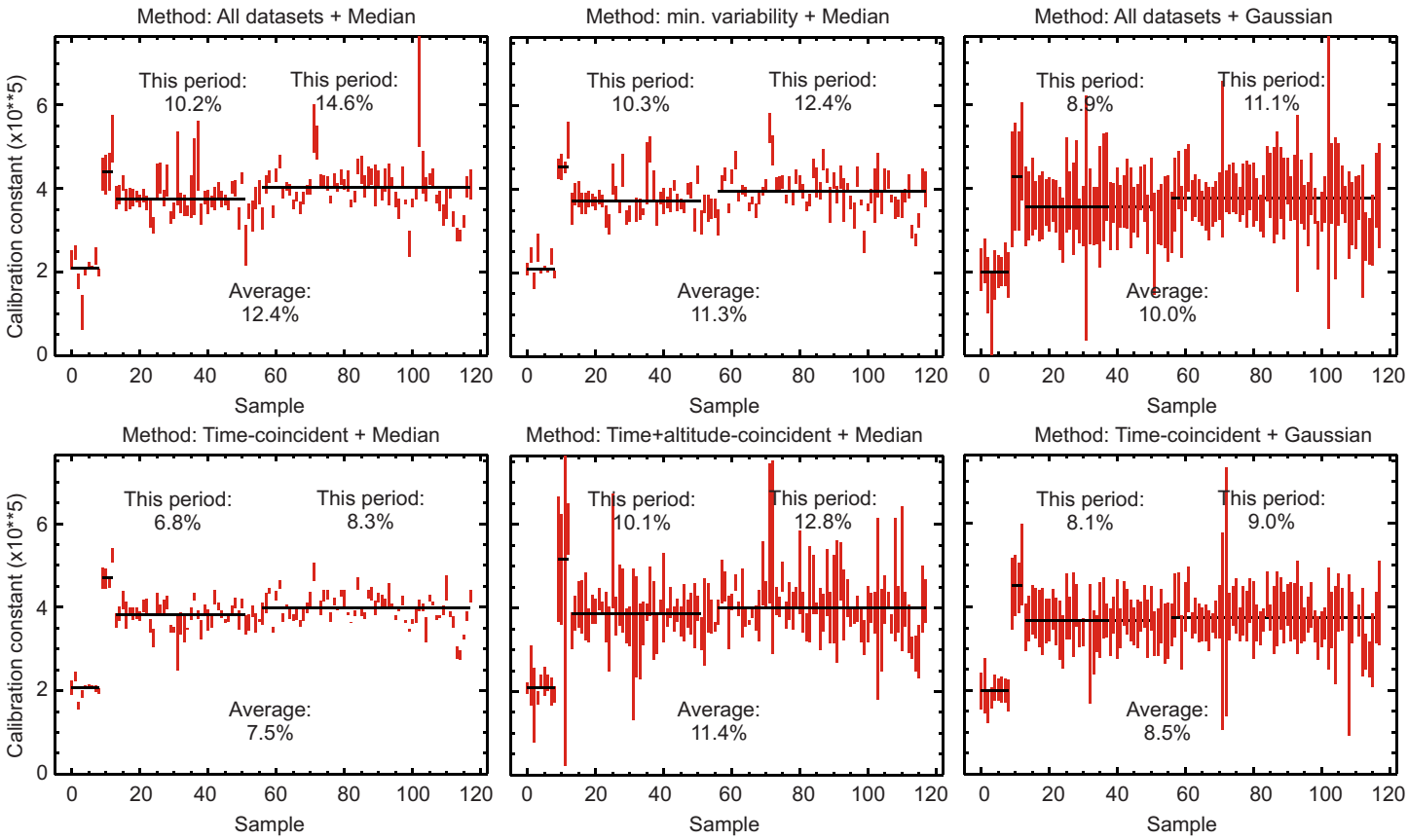

Figure 5. Six examples of the 12 calibration methods tested. The percentages represent the standard deviations of the calculated calibration constants over 2 sub-periods and averaged over the entire 118 samples (10/2007 to 4/2009). See text for details on each method, and Table 2 for a compilation of the standard deviations of all 12 methods

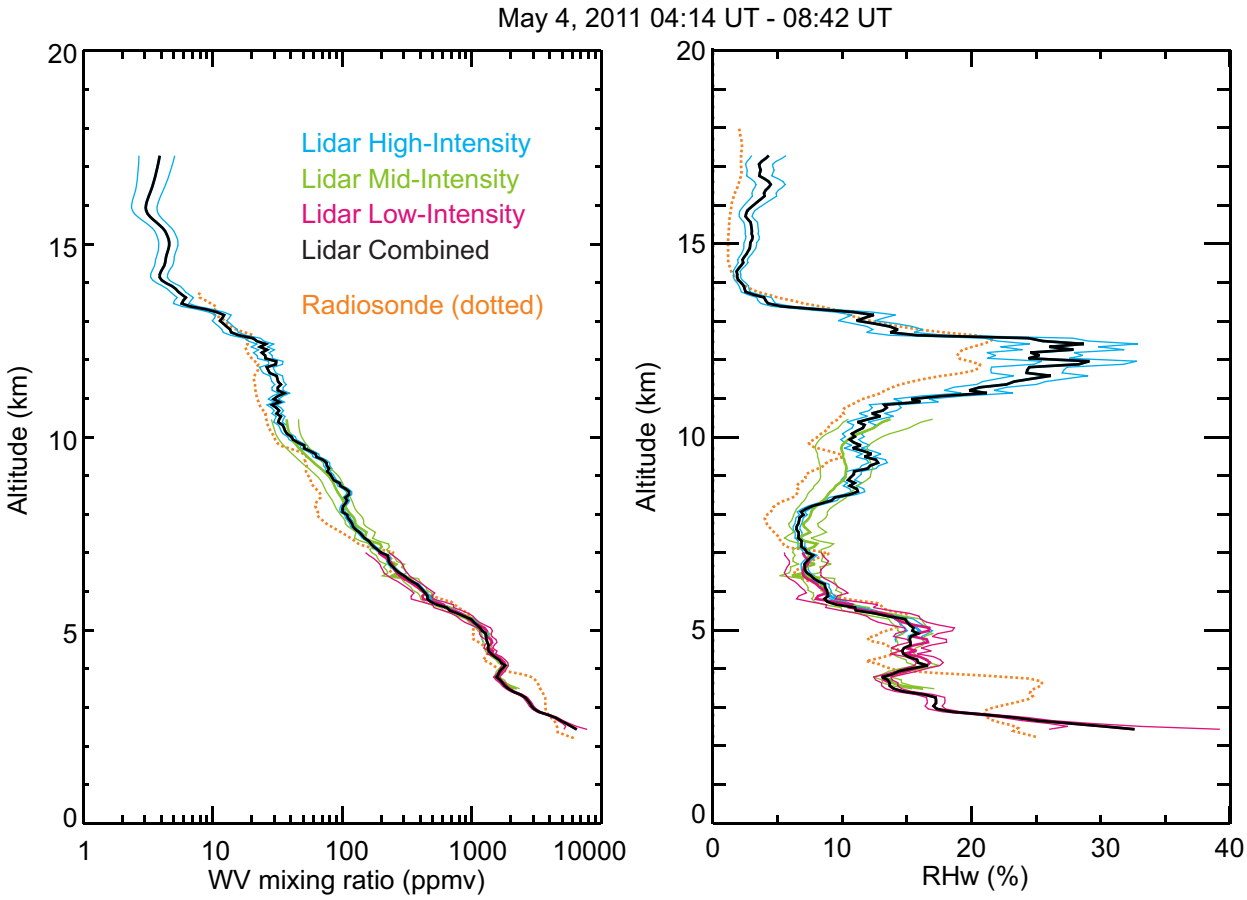

Figure 6. An example of the construction of the final (black) water vapor (left) and Relative Humidity (right) profiles from the three channel ranges (pink: low-intensity, green: mid-intensity, and blue: high intensity). The lidar profiles total uncertainty is over-plotted using thin solid curves, and a co-located radiosonde profile is over-potted using a dotted orange curve

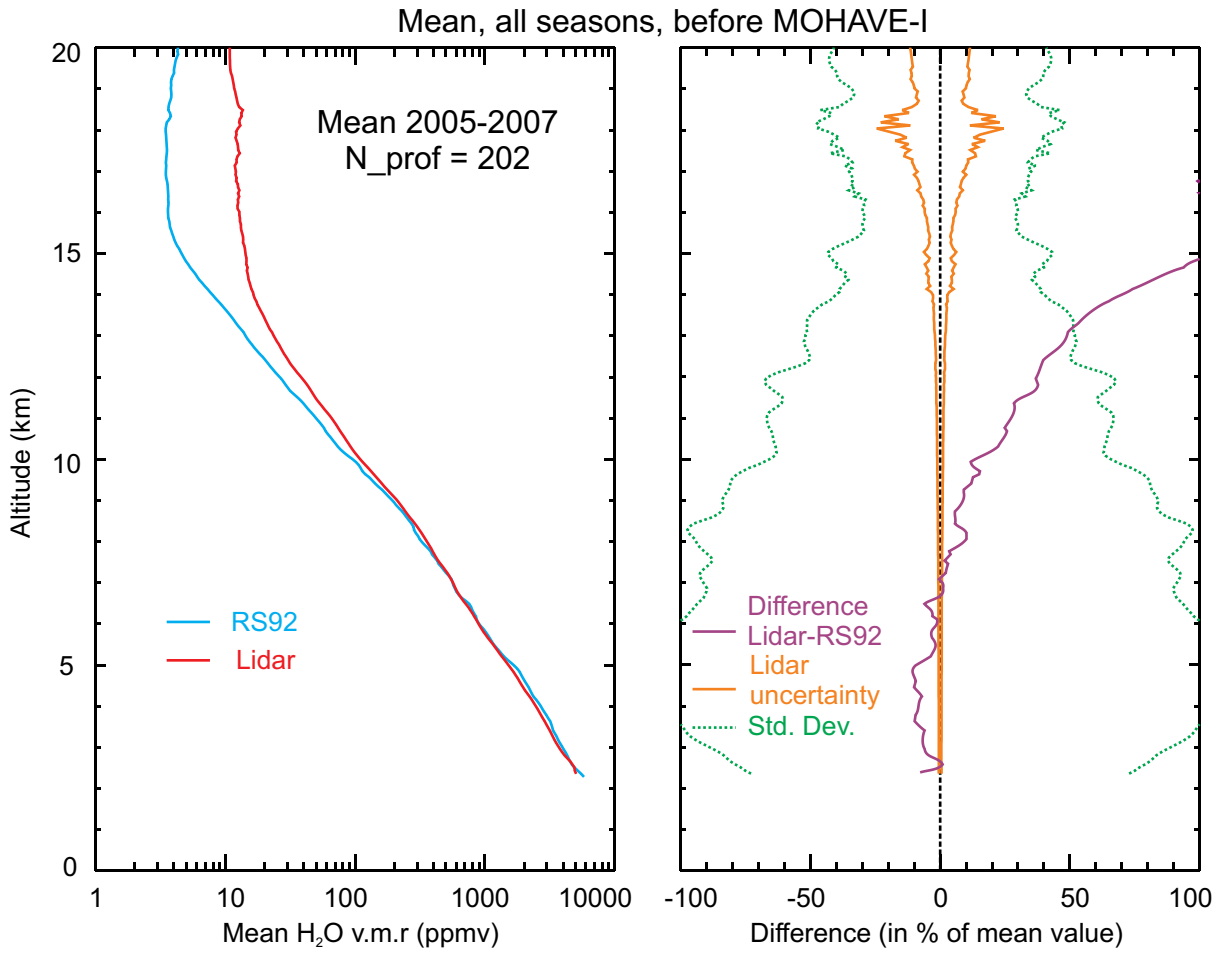

Figure 7. Comparison of the climatological mixing ratio profiles measured over the period June 2005-October 2007 by the TMW lidar and simultaneous and co-located Vaisala RS92 radiosondes. Left plot: Mean profiles. Right plot: Mean difference, lidar measurement uncertainty and standard deviation

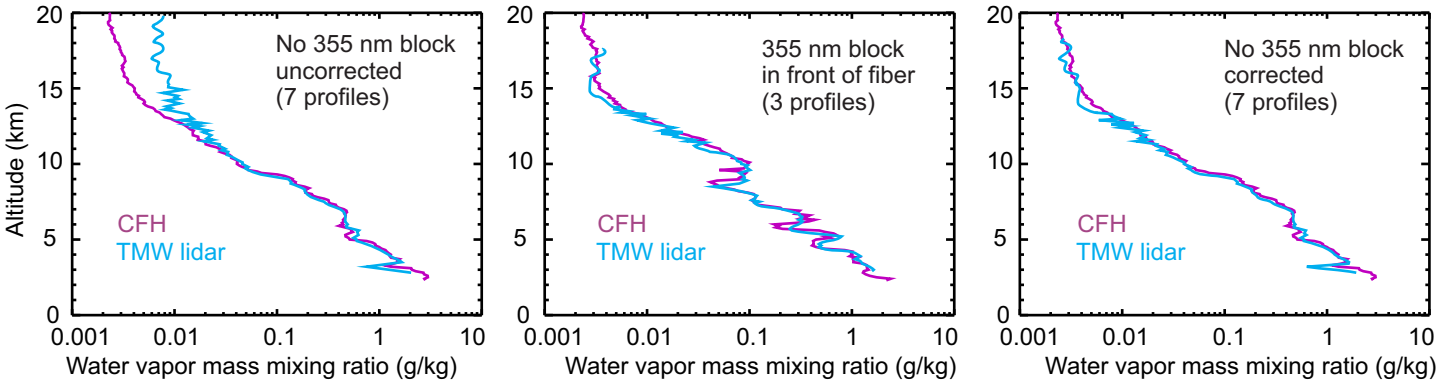

Figure 8. Comparison of mean water vapor mixing ratio profiles measured simultaneously by the TMW lidar and CFH during MOHAVE-I. Left plot: Lidar signals contaminated by fluorescence (no 355 nm block). Center plot: Lidar signals free of fluorescence (355 nm block installed). Right plot: Same as left plot, but after fluorescence correction is applied to lidar signals

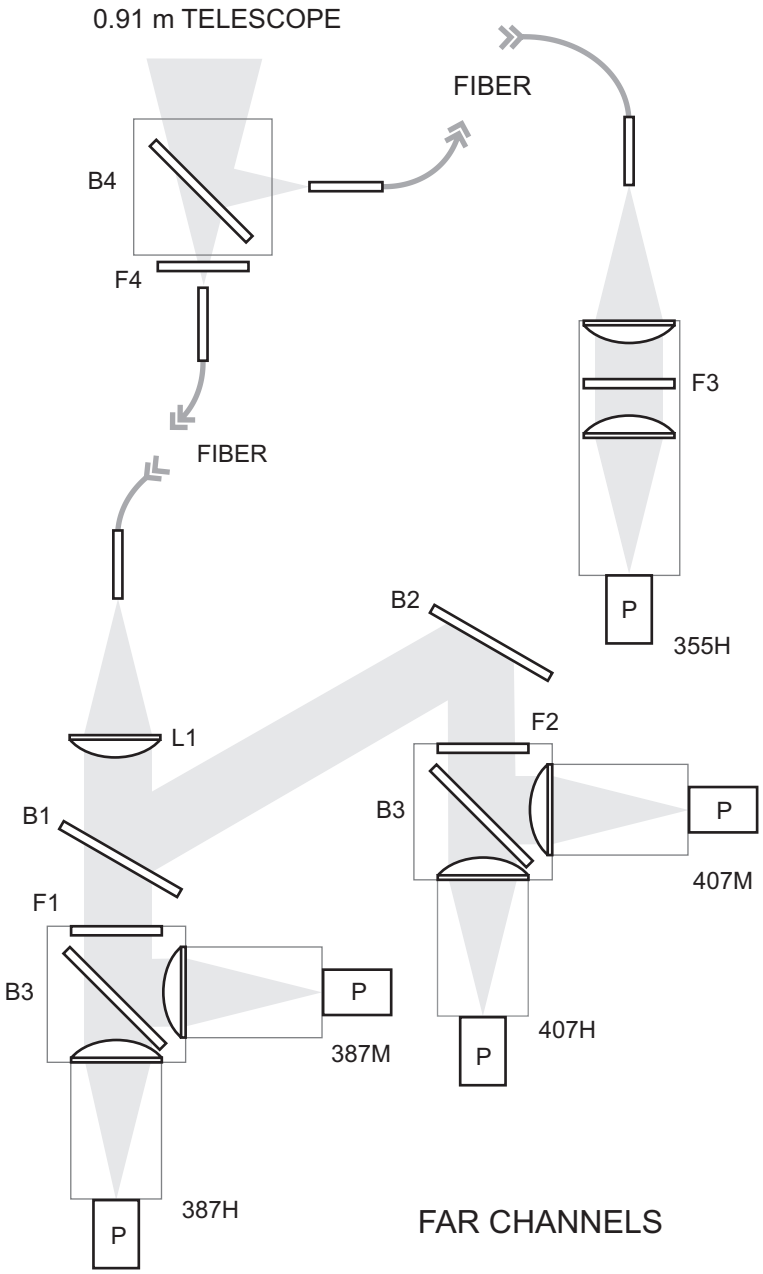

Figure 9. TMW lidar receiver second configuration (2007-2009) for the high- and mid-intensity channels

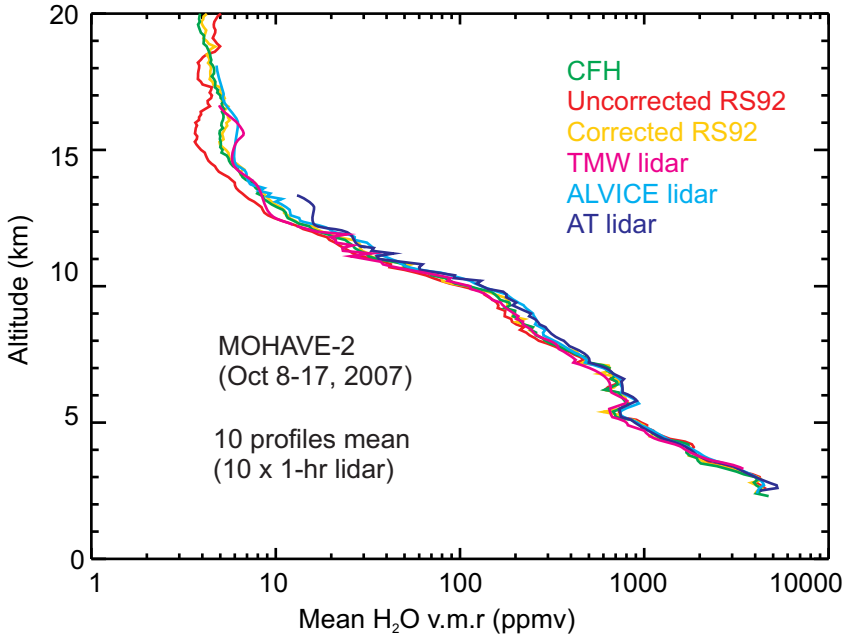

Figure 10. Averaged water vapor mixing ratio computed from 10 profiles measured simultaneously during MOHAVE-II by the TMW lidar, two other lidars, CFH, and Vaisala RS92 radiosondes.

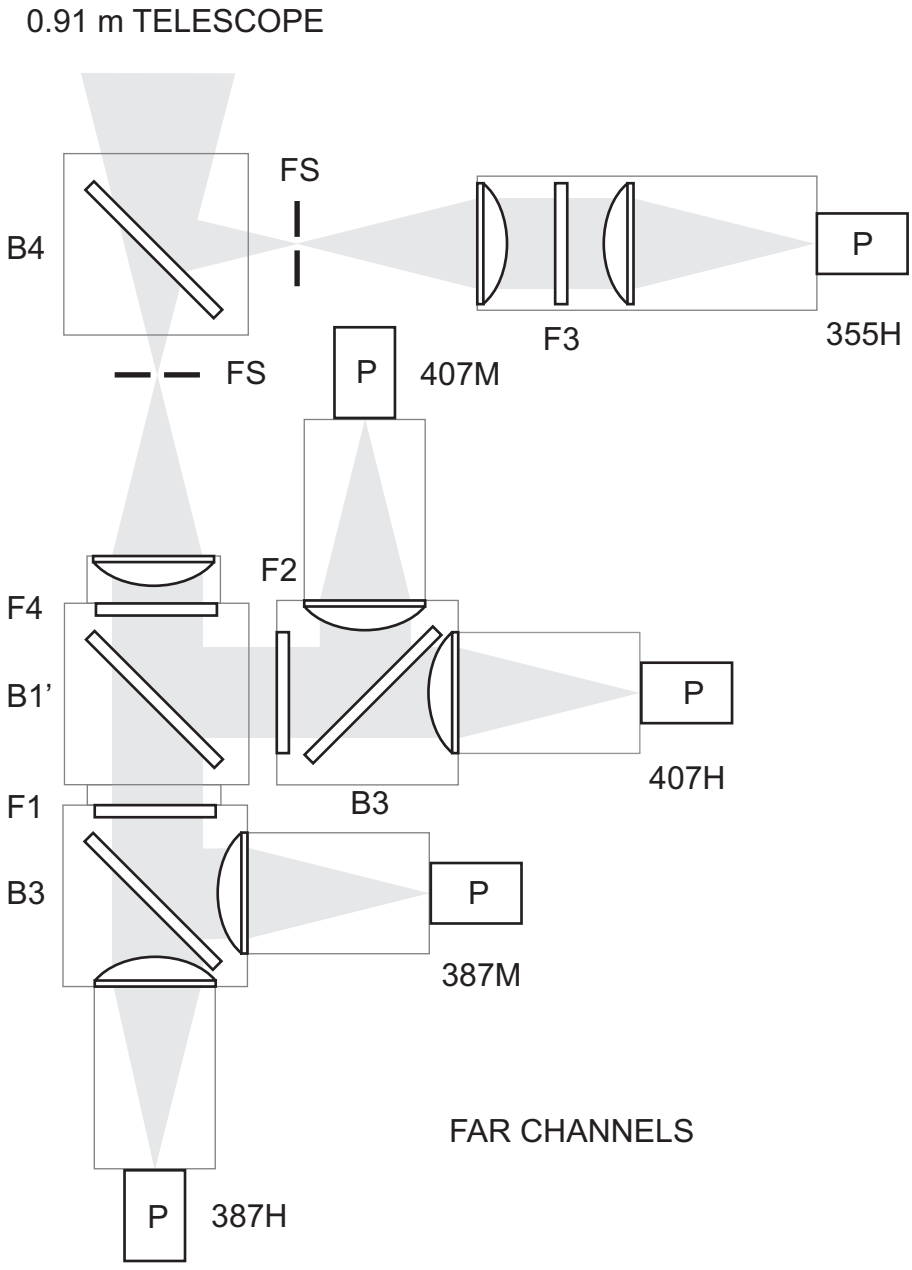

Figure 11. TMW lidar receiver final configuration (2009-present) for the high- and mid-intensity channels

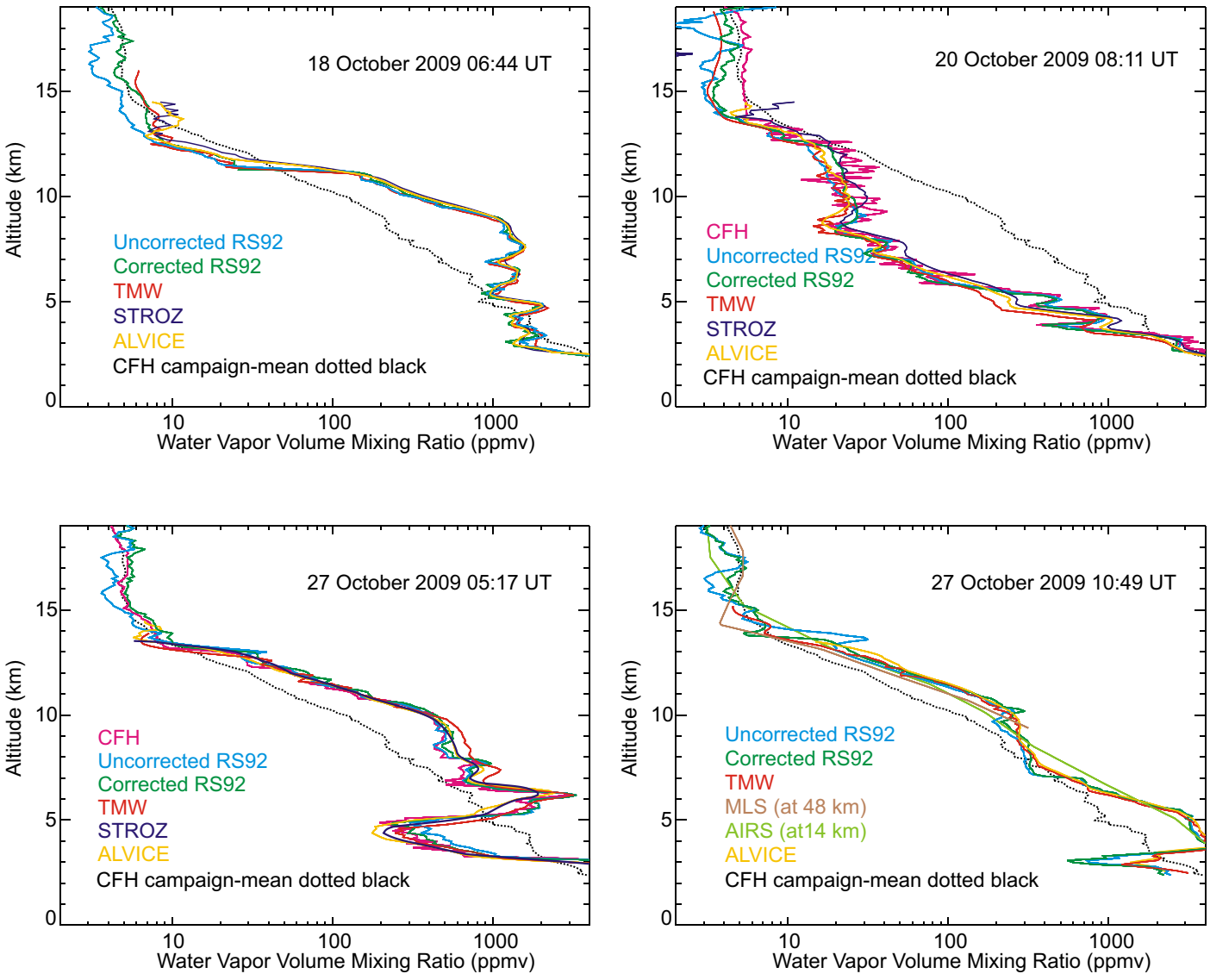

Figure 12. Example of four water vapor mixing ratio profiles measured simultaneously by TMW, two other lidars, CFH, Aura-MLS, and RS92 radiosondes during MOHAVE-2009

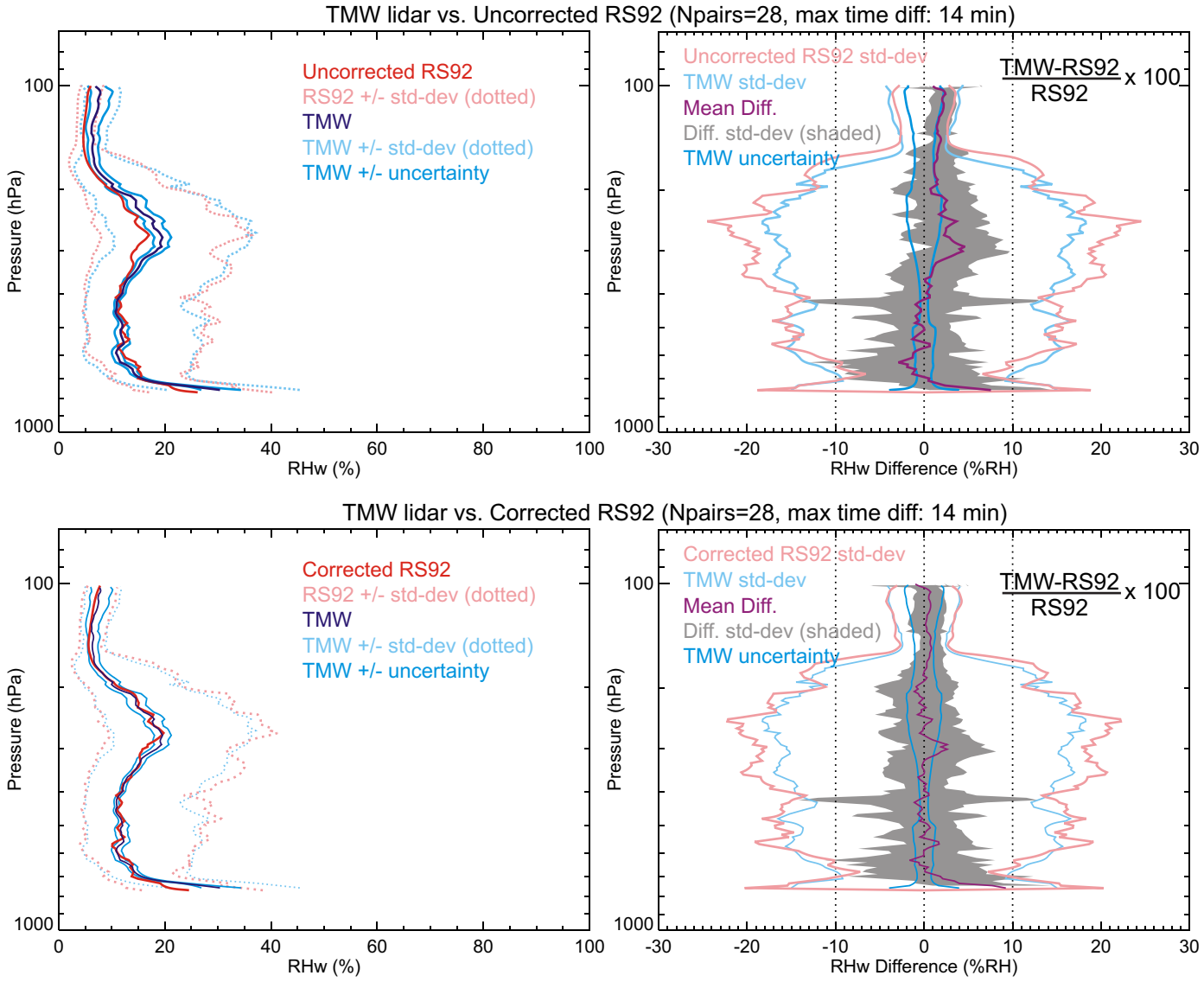

Figure 13. Comparison of the MOHAVE-2009 campaign-averaged relative humidity profiles calculated from the 28 simultaneous profiles measured by the TMW lidar and Vaisala RS92 radiosondes. Top plots: mean profiles (left) and difference with the uncorrected RS92 data (right). Bottom plots: mean profiles (left) and difference with the corrected (Miloshevich et al., 2004; 2009) RS92 data (right)

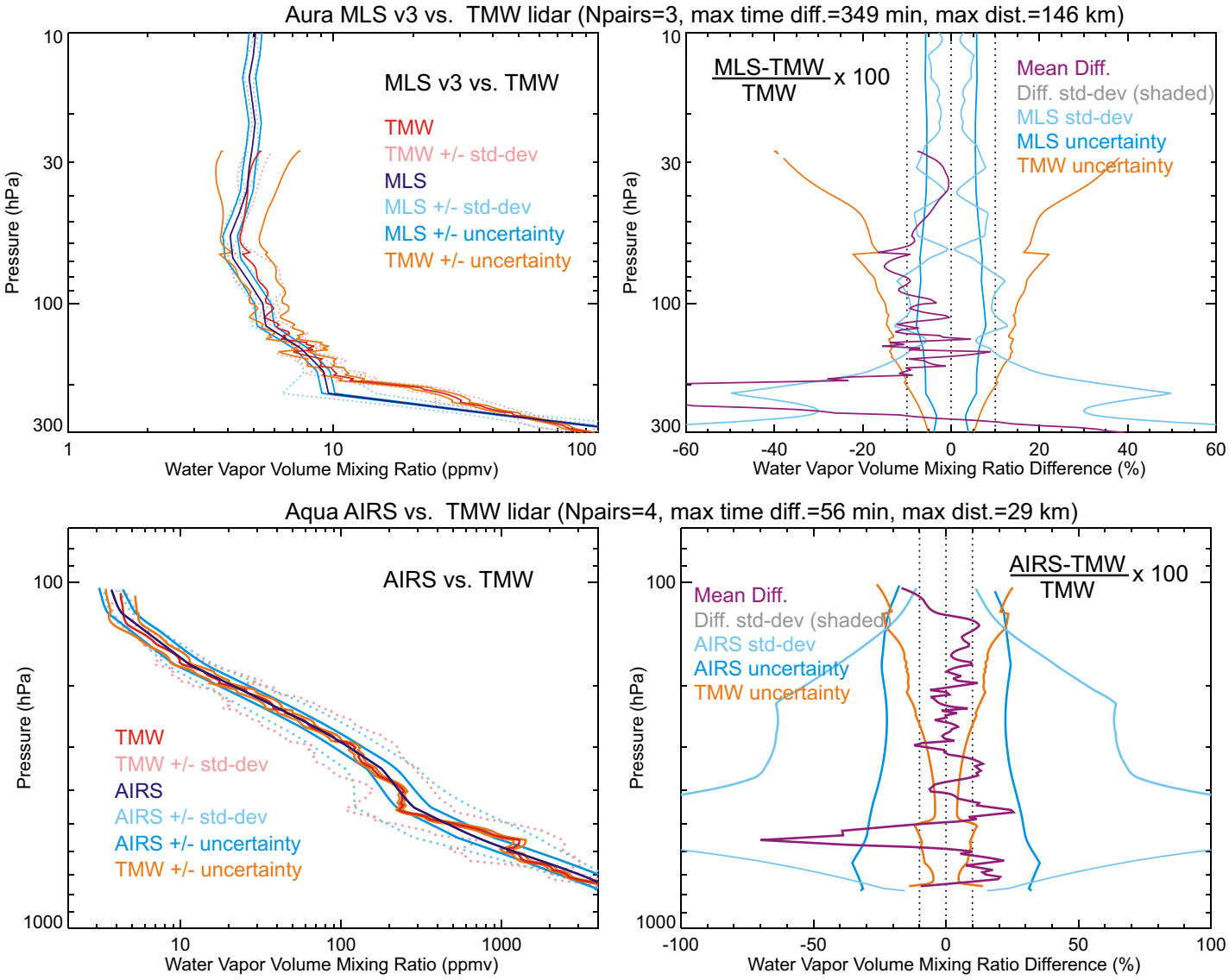

Figure 14. Top: averaged mixing ratio profiles (left) calculated from the 3 coincident profiles measured in the UTLS by the TMW lidar and Aura-MLS during MOHAVE-2009, and their difference (right). Bottom: averaged mixing ratio profiles (left) calculated from the 4 simultaneous profiles measured by the TMW lidar and Aqua-AIRS during MOHAVE-2009, and their difference (right)

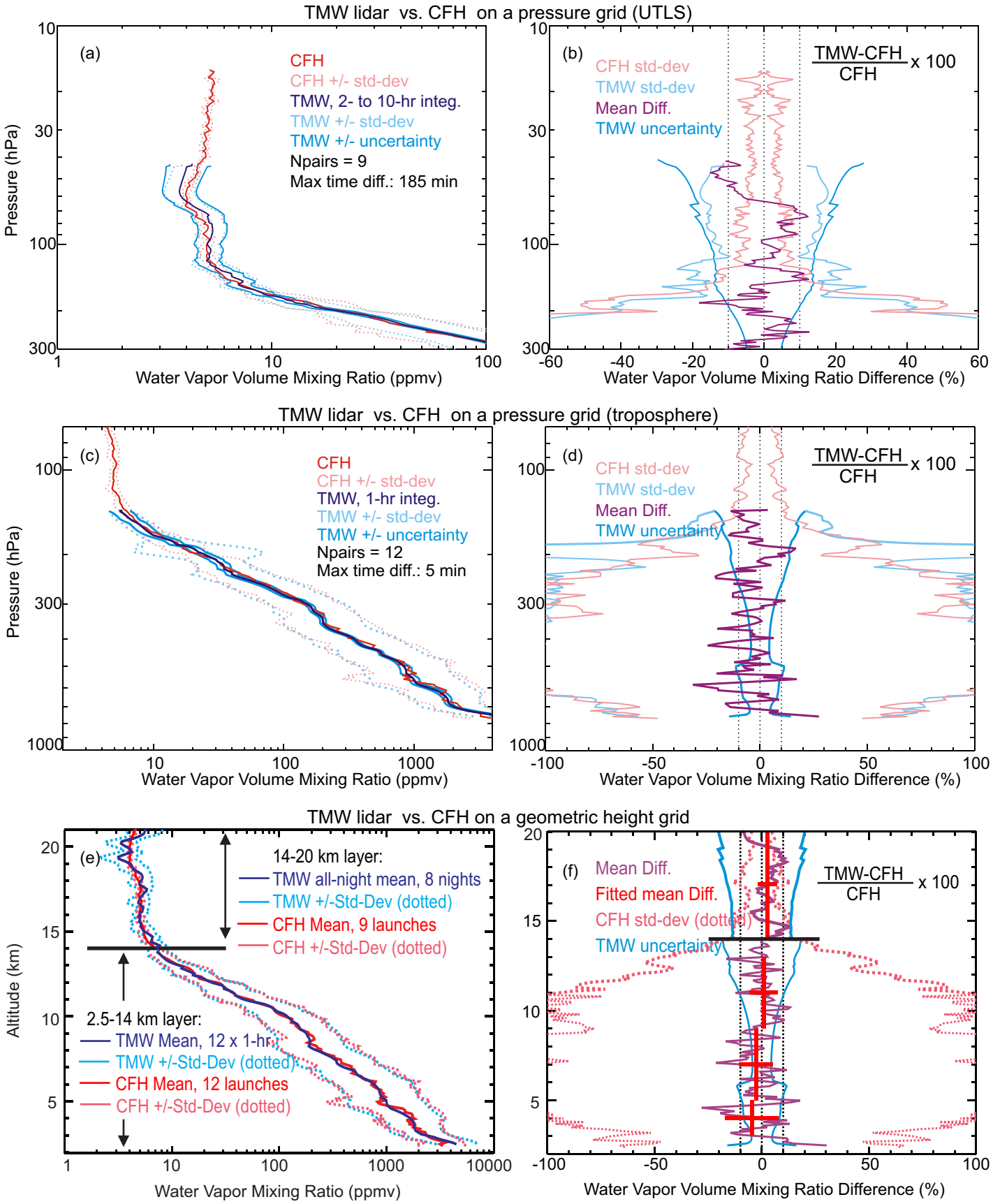

Figure 15. Comparison of the MOHAVE-2009 campaign-averaged mixing ratio profiles (left) calculated from all coincident profiles measured by the TMW lidar and the CFH, and their difference (right). Top row (a-b): UTLS on a pressure grid; middle row (c-d): troposphere on a pressure grid; bottom row (e-f): troposphere and lower stratosphere on a geometric height grid. See text and figure for details on the coincidence criteria

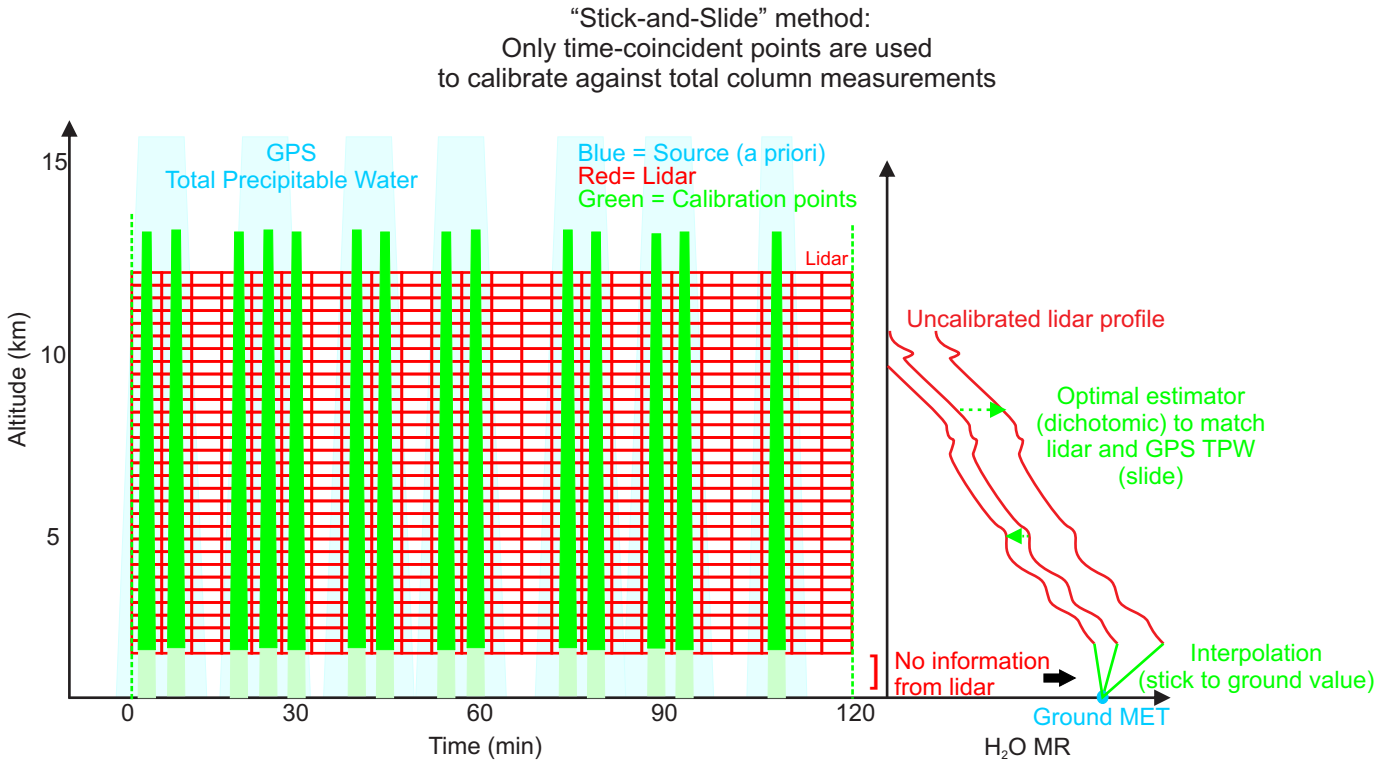

Figure 16. Schematics of the "stick-and-slide" method used to perform a calibration of the TMW lidar profiles using co-located and simultaneous GPS Total Precipitable Water measurements. Lidar data points are symbolized by red open rectangles, GPS column data by blue vertical bars, and the actual column 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

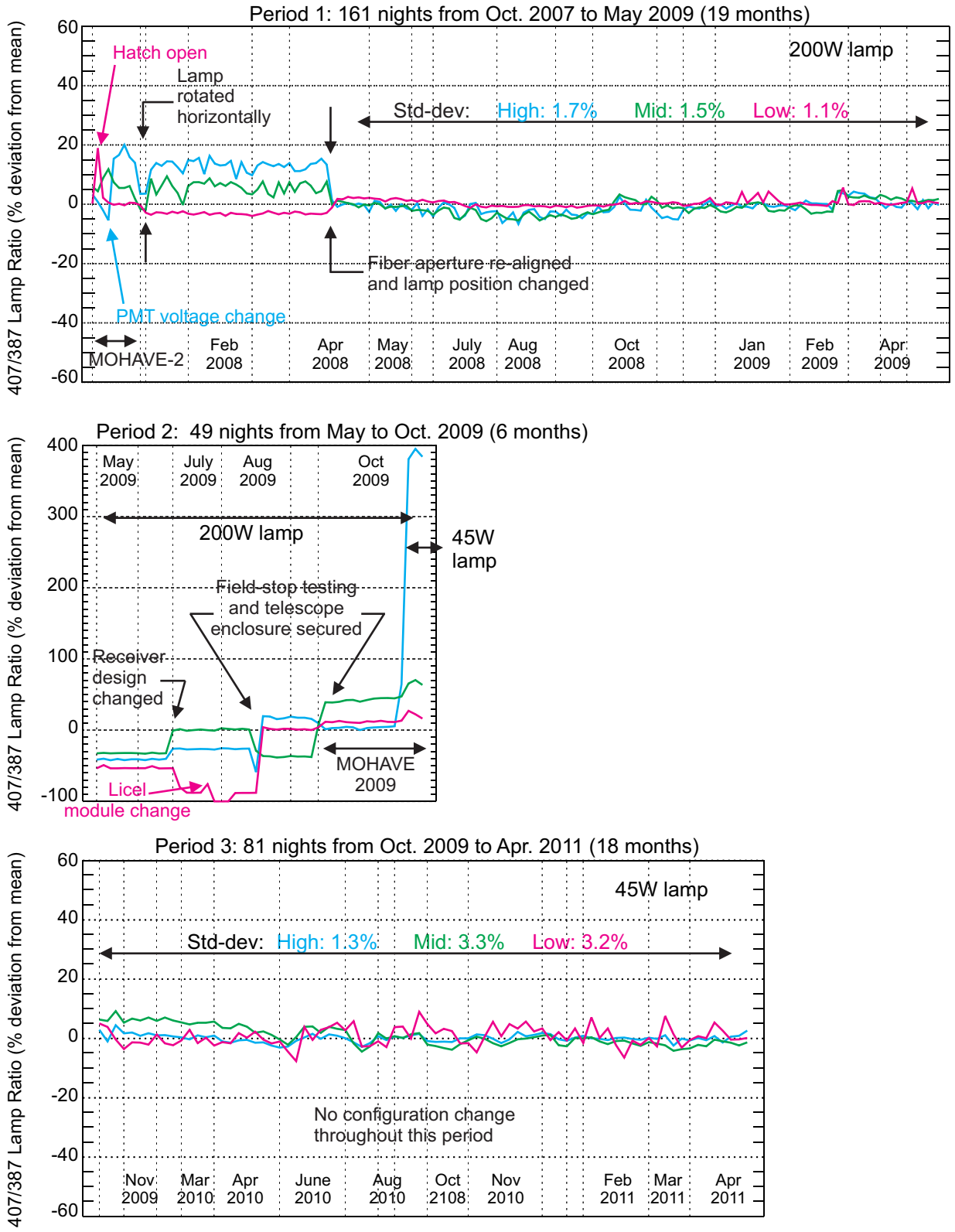

Figure 17. Historical evolution of the 407 nm to 387 nm channel ratios (pink: low-intensity, green: mid-intensity, and blue: high-intensity) between October 2007 and present. All instrumentation changes are denoted by arrows and comments, including the change of lamp (from 200W to 45W) in October 2009