



# Study and mitigation of calibration factor instabilities in a water vapor Raman lidar

Leslie David<sup>1</sup>, Olivier Bock<sup>1</sup>, Christian Thom<sup>2</sup>, Pierre Bosser<sup>3</sup>, and Jacques Pelon<sup>4</sup>

<sup>1</sup>IGN LAREG, Univ Paris Diderot, Sorbonne Paris Cité, Paris, France

<sup>2</sup>IGN LOEMI, Saint-Mandé, France

<sup>3</sup>Lab-STICC, ENSTA Bretagne, Brest, France

<sup>4</sup>LATMOS, CNRS/UPMC/IPSL, Paris, France

Correspondence to: Leslie David (leslie.david@mines-ales.org)

Received: 30 May 2016 – Discussion started: 5 July 2016

Revised: 23 May 2017 – Accepted: 7 June 2017 – Published: 31 July 2017

**Abstract.** We have investigated calibration variations in the Rameau water vapor Raman lidar. This lidar system was developed by the Institut National de l'Information Géographique et Forestière (IGN) together with the Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS). It aims at calibrating Global Navigation Satellite System (GNSS) measurements for tropospheric wet delays and sounding the water vapor variability in the lower troposphere. The Rameau system demonstrated good capacity in retrieving water vapor mixing ratio (WVMR) profiles accurately in several campaigns. However, systematic short-term and long-term variations in the lidar calibration factor pointed to persistent instabilities. A careful testing of each subsystem independently revealed that these instabilities are mainly induced by mode fluctuations in the optic fiber used to couple the telescope to the detection subsystem and by the spatial nonuniformity of the photomultiplier photocathodes. Laboratory tests that replicate and quantify these instability sources are presented. A redesign of the detection subsystem is presented, which, combined with careful alignment procedures, is shown to significantly reduce the instabilities. Outdoor measurements were performed over a period of 5 months to check the stability of the modified lidar system. The calibration changes in the detection subsystem were monitored with lidar profile measurements using a common nitrogen filter in both Raman channels. A short-term stability of 2–3 % and a long-term drift of 2–3 % per month are demonstrated. Compared to the earlier Development of Methodologies for Water Vapour Measurement (DEMEVAP) campaign, this is a 3-fold improvement in the long-term sta-

bility of the detection subsystem. The overall water vapor calibration factors were determined and monitored with capacitive humidity sensor measurements and with GPS zenith wet delay (ZWD) data. The changes in the water vapor calibration factors are shown to be fairly consistent with the changes in the nitrogen calibration factors. The nitrogen calibration results can be used to correct the overall calibration factors without the need for additional water vapor measurements to within 1 % per month.

## 1 Introduction

Water vapor plays an active role in many atmospheric processes involved in climate change and variability (Soden et al., 2002). Accurate monitoring of water vapor profiles is thus of paramount importance to better understand these processes and improve climate models. Raman lidar has become, after several decades of research and development, a privileged device for the measurement of atmospheric water vapor profiles (Whiteman et al., 1992; Wandinger, 2005). According to the GCOS-112 report, the ideal climate requirements for monitoring water vapor in the troposphere are 2 % precision, 2 % absolute accuracy, and 0.3 % per decade stability (GCOS-112, 2007). This ultimate goal motivated significant international efforts to improve the accuracy and stability of current Raman lidar systems. In a different application field, but with similar stringent accuracy constraints, Raman lidars are expected to provide a means to calibrate wet path delays of Global Navigation Satellite

System (GNSS) measurements (Bossler et al., 2010). The Institut National de l'Information Géographique et Forestière (IGN) and the Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS) have jointly developed a multipurpose mobile water vapor Raman lidar system called Rameau (Bock et al., 2001; Tarniewicz et al., 2002). The accuracy of water vapor mixing ratio (WVMR) profiles measured with Rameau has been demonstrated during several campaigns by comparison with other humidity sensors.

To achieve lidar data with high accuracy and stability a careful and periodic calibration of the system is necessary. Calibration gets rid of systematic errors due to uncertainties in Raman cross sections and in instrumental factors (transmittance of optical elements, detectors of quantum efficiency and gain, mean atmospheric transmittance, etc.) and converts the lidar signals to absolute WVMR measurements (Whiteman et al., 1992). Two different calibration techniques have been usually considered: one independent and the other dependent on external reference water vapor sensor measurements. The former has been first described by Vaughan et al. (1988), who calculated the calibration coefficient as a combination of the H<sub>2</sub>O and N<sub>2</sub> Raman backscatter cross sections from Penney and Lapp (1976) and the relative instrumental transmission from experimental measurements and manufacturer data. Later, Sherlock et al. (1999b) suggested a new sensor-independent calibration method, which aimed at monitoring drifts caused by system changes such as aging of optical components. This new method is based on the decomposition of the calibration coefficient in two terms: one is represented by the transmission and detection efficiency of the Raman signals, this being specific to each lidar system, and the other is represented by the Raman effective cross sections. Daytime measurements of diffuse sunlight are performed in order to determine the transmission and detection efficiency of the system. The Raman effective cross sections are estimated from semi-empirical models which take into account the temperature dependency of the Raman cross section. The effective cross sections are then convolved with the instrument function. The absolute accuracy of this method is not better than 10 %, mainly due to theoretical uncertainties in the Raman cross section models (Sherlock et al., 1999b). More recent empirical determination of water vapor Raman cross sections by Avila et al. (2004) finally allowed achieving an absolute accuracy better than 10 %. The use of these improved cross sections in the independent calibration technique of Venable et al. (2011) even resulted in agreement between the sensor-independent calibration and the traditional radiosonde-dependent technique of better than 5 %. The radiosonde-dependent technique is performed by calculating a constant, height-independent, normalization factor from the comparison of the lidar profile with a radiosounding (Ferrare et al., 1995; Leblanc et al., 2012). Other sensor-dependent calibration techniques also use integrated water vapor (IWV) or zenith wet delay (ZWD) measurements, from GPS or from microwave radiometers (Turner and Gold-

smith, 1999), or ground-based humidity sensor data (Revercomb et al., 2003). The dependent technique does not require us to determine the terms affected by large uncertainties such as the Raman cross sections of H<sub>2</sub>O and N<sub>2</sub> and unknown instrumental factors. It is an easy method for transferring the absolute accuracy of a reference sensor to the lidar data, hence potentially achieving an agreement of 5 % with the best operationally used radiosondes, 1–2 % with ground-based humidity sensors, and 2–5 % with GPS IWV or ZWD data (Bock et al., 2013). The dependent technique suffers two limitations, however. First, equipment changes might impact the homogeneity of the reference sensor measurements and thus the long-term stability of the calibrated lidar data. Second, radiosonde-dependent calibration techniques may not be usable to monitor short-term variations due to thermal tweaks in the optical alignment as it would not be feasible to launch sondes hourly or even more frequently. Hence, a combination of sensor-dependent and independent calibration technique is a good option to achieve better absolute calibration with the dependent technique and a good stability, both in the short term and long term, with the independent technique. Sherlock et al. (1999b) suggested controlling the calibration stability with background measurements (keeping the laser off) during daytime. However, this method is based on the need for independent aerosol data, which limits its implementation. Lately, Leblanc and McDermid (2008) developed a stability control procedure with calibrated spectral lamps, reaching a calibration factor standard deviation of 2 % over more than a year. Venable et al. (2011) proposed an improved independent calibration technique which consists in scanning a known light source over the telescope aperture instead of a stationary lamp system. Advantages and drawbacks of these methods are further discussed in (Whiteman et al., 2011).

The Development of Methodologies for Water Vapour Measurement (DEMEVAP) campaign conducted in 2011 at the Observatoire de Haute-Provence, France, brought together Rameau and various instruments for measuring water vapor (Bock et al., 2013). A comparison of various sensor-dependent lidar calibration methods was carried out. Four methods were tested using (i) point measurements from capacitive humidity sensors, (ii) radiosonde upper-air measurements, (iii) IWV measurements, and (iv) the GPS–lidar coupling technique developed by Bossler et al. (2010). A more detailed description of the methods and the overall outcomes of the campaign have been presented in a previous paper (Bock et al., 2013). Despite the accurate vertical profiling capacity of the Rameau lidar, significant variations of the calibration coefficient were observed during this experiment. An overall drift of 15 % over the 45 days of the campaign was noticed as well as 7 % peak-to-peak variations in the calibration coefficient between methods. In addition, for one given method, short-term fluctuations of 5 % were observed during one night. The long-term variation is partly explained by the change of the optical fiber, which is coupling the tele-

scope receiver to the optical system containing the narrow band filters and photomultiplier tubes (PMTs). The change consisted in the replacement of a 0.8 mm fiber with a 0.4 mm fiber. This instrumental change required a complete realignment of the receiver optics which produced a jump in the calibration coefficient. We quickly hypothesized that a change in the beam size or position on the PMT photocathode produced a different signal strength. Indeed, spatial nonuniformity of PMT photocathode sensitivity is a problem which has been known for some time (Takhar, 1967; Dos Santos et al., 1996; Simeonov et al., 1999; Freudenthaler, 2004). Diverse solutions have been tested, such as inserting a diffuser in front of each PMT window (Simeonov et al., 1999) or transmitting the beam through a field lens combined with a mirror tube (Freudenthaler, 2004). These solutions allowed the spreading and/or focusing of the beam evenly onto the surface photocathode. However, they came along with significant signal losses (e.g., Simeonov et al., 1999, estimated a loss of up to 25 %), which is detrimental to the detection sensitivity of the weak Raman signals. While seeking for the minimization of the PMT effects, we have noticed that beam mode fluctuations at the output of the fiber were adding significant variability to the lidar signals.

This paper is devoted to the investigation and reduction of the major instability sources of the Rameau lidar system. Section 2 describes the instrumentation of the Rameau lidar system and discusses the potential instability sources in each of the subsystems. A redesign of the receiving optical system which allowed us to eliminate unexpected vignetting is described. Two new optical layouts which were introduced and tested in order to reduce the main instability sources, e.g., due to the optic fiber mode fluctuations and PMT photocathode sensitivity are also described. Section 3 presents experimental evidence of these instabilities from indoor measurements performed on the detection subsystem and demonstrates improved performance using the two new optical layouts. Section 4 presents outdoor experimental results of the overall the system stability using one of the new optical layouts. Finally, Sect. 5 discusses possible options for further improvement and concludes.

## 2 Presentation of the IGN-LATMOS Raman lidar and inventory of the sources of instability

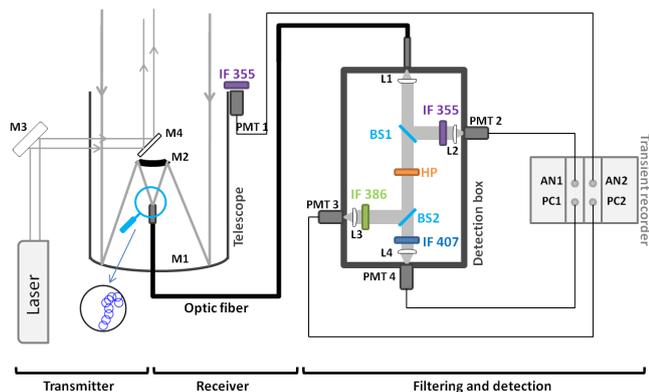
Figure 1 depicts the setup of the Rameau lidar system as it operated at the time of the DEMEVAP campaign (Bock et al., 2013). Below we discuss the potential signal and calibration variation sources in each of the subsystems.

The transmitter subsystem is composed of a Quantel Brilliant frequency-tripled Nd:YAG laser, transmitting pulses of  $\sim 60$ – $70$  mJ at 354.7 nm with a repetition frequency of 20 Hz. The divergence is around 0.5 mrad full-width at half maximum with a pointing stability (short-term jitter measured over 200 shots) of 0.075 mrad (FWHM). Monitoring the

pointing stability over several hours revealed beam wandering of about 0.5 mrad, however. Nevertheless, all these angles are reduced after transmission through an 8.4x refractive beam expander. The laser beam is finally sent out into the atmosphere coaxially with the receiving telescope through two deflecting mirrors. The beam expander also allows ensuring the eye safety of the system. The laser pulse energy is monitored by sampling a small reflected beam by a thin glass plate placed between the beam expander lenses (not shown). Energy is measured with an Ophir laser energy meter and logged into a file on the main PC for further analysis. Energy drops larger than 5 % are corrected by readjusting the alignment of the  $3\text{-}\omega$  crystal. During the DEMEVAP campaign this happened only twice, whereas the short-term laser pulse energy fluctuations remained well below 5 %.

The receiver subsystem is composed of a 30 cm diameter, 72 cm focal length, and a Cassegrain telescope coupled to the filtering and detection subsystem through an optical fiber. During the first stage of DEMEVAP, a 0.8 mm diameter UV grade optical fiber (Sedi Fibre, HCG800) was used. It was later replaced with a 0.4 mm diameter fiber (Sedi Fibre, TCG400). The diameter of the fiber acted as a field stop defining the receiver's field of view as  $\pm 0.28$  mrad (or 0.56 mrad FWHM) in the case of the 0.4 mm fiber. This is large enough to contain the beam movements due to jitter and wandering of the transmitted laser beam. Nevertheless, during DEMEVAP we observed rapid correlated drops in the signal strengths on all three PMTs of the detection subsystem (PMTs nos. 2, 3, and 4), while neither the laser energy nor the signal strength on PMT no. 1 fluctuated. These signal drops can only suggest that the spot of the received laser beam was sweeping off the fiber aperture. Signal samples of two such cases illustrating a signal break and/or fading are presented in the top panels of Fig. 2. In the first case, the signal break amounted to 20 %, while in the second case the signal reduced by a factor of 2.5 within 1 h. Such large movements of the received laser beam can only be explained by thermo-mechanical deformations of the optical bench, which are currently being investigated. During DEMEVAP operations, the drifts in the received signal were not controlled in a systematic manner. Otherwise, the transmitted laser beam would have been steered in such a way as to correct the transmitter and receiver misalignment. The bottom panels of figure 2 show coincident ratios of the Raman signals superposed with water vapor calibration factor values computed from GPS ZWD measurements (Bock et al., 2013). Although there is no one-to-one correspondence between the signal variations in the top panels and the curves in the bottom panels, significant variations in the calibration factors and signal ratios are observed as well. Indeed, one can expect that spot displacements at the input of the fiber impact the shape and position of the beam at the output and thus the calibration factors.

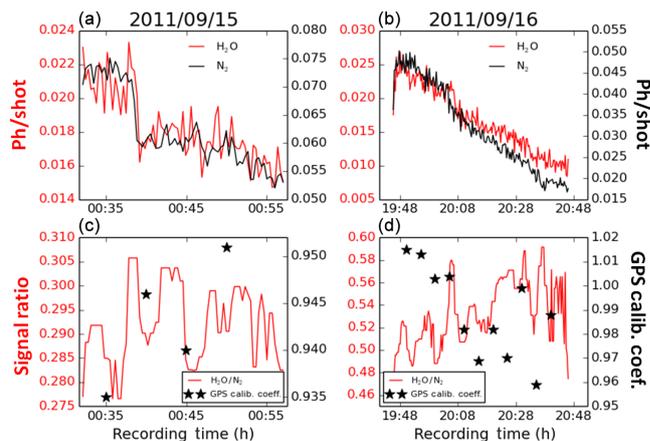
Multimode optical fibers are sometimes thought of as being optical signal scrambling devices, but previous research (Avila, 1998) and our work here indicate that not to be the



**Figure 1.** Schematic view of the Rameau water vapor Raman lidar system

case. Short-length OH-rich fibers are typically used in lidars to reduce signal loss and fluorescence (Sherlock et al., 1999a). The beam at the exit of such a fiber is composed of one or several superposed modes depending on the core diameter and on the numerical aperture (NA) of the fiber (Ghatak and Thyagarajan, 1989). Variations in  $NA_{out}$  due to tilts of the incident beam or even tilts of the fiber core with respect to the cladding have also been reported by Avila (1998). Figure 3 shows pictures of beam spot samples that we observed at the output of our 1 mm diameter fiber for various positions and tilts of the injected light beam (from a 468 nm LED with matched numerical apertures of  $NA_{in} = 0.22$ ). Similar changes in the output beam diameter have been reported by Whiteman et al. (2011). In addition to the changes in size and shape of the output beam, we also observed that the NA of the emerging beam can be larger than specified by the manufacturer. We tested several fibers with diameters of 0.2, 0.4, 0.8, and 1.0 mm, which are all given for an  $NA = 0.22$  (Sedi Fibre). For the 0.2 and 0.4 mm fibers, the  $NA_{out}$  was consistent with the manufacturer specifications. However, for the 0.8 and 1.0 mm fibers we measured up to  $NA_{out} = 0.30$ . At the time of the DEMEVAP campaign, we were using a 0.8 mm fiber in a system designed for a fiber with an NA of 0.22. Using a fiber with a larger than expected NA resulted in vignetting on the apertures of various optical devices of the detection subsystem, which probably added significant sensitivity to beam wandering and thus increased calibration instability.

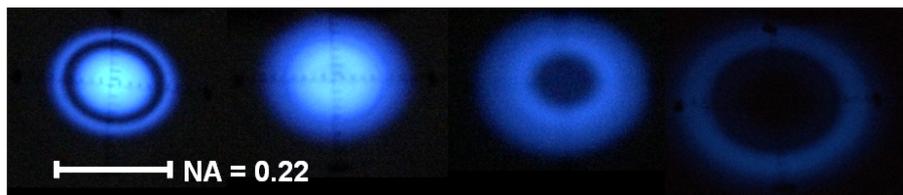
In order to eliminate this vignetting, we used the ZEMAX commercial optical ray tracing software ([www.zemax.com](http://www.zemax.com)) to optimize the optical design of the detector subsystem for different optical fibers. For each fiber, the focal length of the collimating L1 and the distance between the source – i.e., fiber output – and L1 was optimized to obtain a collimated beam throughout the system. The focusing lenses (L2, L3, and L4) were chosen in such a way as to achieve a 5 mm diameter spot centered on the 8 mm diameter PMT photocathodes. We came up with two options depending on the fiber



**Figure 2.** Examples of signal break (a) and signal fading (b) observed in the Rameau lidar measurements during the DEMEVAP campaign in 2011. The top panels show  $H_2O$  (red, left-hand axis) and  $N_2$  (black, right-hand axis) signals (photons per laser shot, ph/shot) averaged over the distance range 317–1317 m. The bottom panels show the ratio of the two Raman signals (red curve, smoothed with a five-point median filter) and the coincident relative calibration coefficients (black stars) computed from 5 min average lidar and ZWD GPS data.

diameter. The first option uses the smaller fibers (either 0.2 or 0.4 mm) and allows keeping the current length of the optical system. The second option uses the larger fibers (0.8 or 1 mm), which have the benefit of being less sensitive to beam wandering at the fiber input. This option requires modifying the layout of the detector subsystem to shorten the optical path between L1 and L3/L4. For now, we have not considered the option of removing the optical fiber to get rid of the mode fluctuations. Indeed, this would imply a complete redesign of the receiver and detector subsystems. With the current system, the fiber allows us to displace the detection subsystem from the transmitter and receiver block which is prone to electromagnetically induced interference.

The filtering and detection subsystem comprises, apart from the collimating lenses, two beam splitters (uncoated and coated UV-glass thin plates), B1 and B2; two high-pass filters (HPs) to ensure a rejection of the 354.7 nm signal to OD 12 ( $10^{-6}$  each); and narrow band interference filters (IFs) to select either the Rayleigh–Mie (354.7 nm), Raman  $N_2$  (386.7 nm), or Raman  $H_2O$  (407.6 nm) signals. The interference filters used during DEMEVAP were from Barr Associates, Inc., with an FWHM of 0.38 nm ( $H_2O$ ), 0.44 nm ( $N_2$ ), and 0.41 nm (Rayleigh–Mie). The filters were not intentionally tilted, so the incidence angles of the beams were assumed to be  $0^\circ$ . The temperature dependence of the Raman cross sections were computed following Whiteman (2003), Avila et al. (1999, 2004), and Adam (2009). The decrease in temperature between the surface and 10 km altitude typically implies a 10% variation in the transfer function of the  $H_2O$  filter. The temporal variation of the temperature profile



**Figure 3.** Pictures of the beam at the exit of a 1 mm diameter fiber illuminated with a 468 nm LED at different incident angles. The beam is observed on a screen located at a distance of 20 mm after the fiber exit. The scale shows a numerical aperture (NA) of 0.22.

during the DEMEVAP campaign (September–October 2011) produced a 3 % variation in the transfer function for altitudes between 1 and 3 km. Corrections were applied to the measured WVMR profiles to mitigate these vertical and temporal variations (Bock et al., 2013). The variations could not be completely removed, however. The reason might be the small unwanted tilts of the interference filters and imperfect collimation of the beams passing through the interference filters. The temperature dependence of the interference filter central wavelength and bandwidth is sometimes invoked, but this dependence is negligible with our filters (Barr Associates, Inc.).

The detectors used in the Rameau lidar system are R7400-03 miniature metal channel dynode PMTs manufactured by Hamamatsu and assembled with a high-voltage (HV) divider by Licel GmbH. The PMTs are biased with individual stabilized high-voltage power supplies. However, it came out that due to thermal variations, the HV bias can fluctuate by a few volts around the nominal value of 850 V. We measured induced gain variations of  $\sim 0.7\%$  per volt. In order to avoid differential gain variations between the  $N_2$  and  $H_2O$  PMTs, both PMTs are henceforth connected to a single HV power supply. The spatial nonuniformity of our PMT photocathodes was not measured directly. However, by comparison with results reported in the literature, we suspect that it is a significant source of instability in our system. Hamamatsu reported variations of  $\pm 10\%$  for the R1387 series (Hamamatsu Photonics K.K., 2007), while Akgun et al. (2005) measured variations of less than 20 % for the R7525HA. Simeonov et al. (1999) realized a sensitivity map of the 8 mm diameter photocathode of a Hamamatsu R5600 PMT (a previous release of Hamamatsu's miniature metal channel dynode PMTs). They reported variations from 0.2 to 2.8 times the mean signal value of the central part of the photocathode. These measurements were achieved by scanning the PMT's photocathode with a  $200\ \mu\text{m}$  light spot. In our system, the spot size is about 5 mm, which should significantly reduce the sensitivity to the spatial nonuniformity compared to the results of Simeonov et al. (1999). We also expect that the photocathode homogeneity of the more recent R7400 series is improved over the older R5600 series.

The signal acquisition in the Rameau lidar is performed with a Licel GmbH transient recorder combining a 200 MHz counter and a 12 bit  $40\ \text{Ms s}^{-1}$  (mega samples per second) analog digitizer allowing for the measurement of Raman

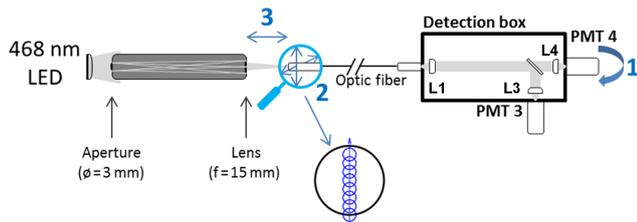
signals in photo-counting mode and the measurement of Rayleigh–Mie signals in analog mode, respectively.

### 3 Quantification of the instability sources from laboratory measurements

In this section, we present the results of laboratory experiments conducted on the detector subsystem to assess the impact of the diverse instability sources identified previously. An experimental device was designed (Fig. 4) to reproduce the beam displacements on one of the PMTs or at the input of the optical fiber. Photo-count signals are measured in the nitrogen and water vapor channels,  $S_3$  and  $S_4$ , and their ratio  $R_{43} = \frac{S_4}{S_3}$  is analyzed as a function of beam displacement. Since the measurements are made at the same wavelength (468 nm from an LED source), the ratio can thus be interpreted as  $R_{43} = \frac{T}{1-T} \cdot \frac{\eta_4}{\eta_3}$ , where  $T$  is the transmission coefficient of the BS2 beam splitter and  $\eta_4$  and  $\eta_3$  are the detection efficiencies of channels 4 and 3, respectively. The ratio is a proxy of the detection subsystem calibration factor. Three instability sources were more specifically investigated with this system. They are identified by numbers on the figure. They simulate the impact of (1) spatial nonuniformity of the PMT when the detected beam is impinging on different zones of the photocathode, (2) the spot movement at the fiber input due to laser beam wandering, and (3) the range-dependent spot size variation of the backscattered signal on the fiber input.

The experimental device presented in Fig. 4 comprises an LED transmitting light through a 3 mm pinhole into a black aluminum tube ending with a 15 mm focusing lens. The output aperture is set to achieve an NA of 0.22, similar to the telescope. The fiber is set at a distance from the output aperture such that the diameter of the light spot impinging on the fiber input is about  $130\ \mu\text{m}$ . This distance is a crucial parameter since an error as small as 1 mm would lead to a spot size 4 to 5 times larger which would rapidly exceed the diameter of the tested fiber (e.g., for 0.2 and 0.4 mm diameter fibers). The fiber is set on a three-axes micrometric translation stage which enables us to adjust this distance and to simulate the spot movements labeled by numbers 2 and 3 in Fig. 4.

The subsections below present the results for three different detection subsystem configurations discussed previously:



**Figure 4.** Schematic of the experimental setup used to test the stability of the detection subsystem. The 468 nm LED source, 3 mm aperture, and 15 mm lens simulate the incident light at the numerical aperture equivalent to the telescope. Three different instability sources are reproduced: (1) rotation of the PMT to test the sensitivity to the position of the beam on the photocathode, (2) lateral translation of the fiber input aperture with respect to the light source to simulate beam wandering at the fiber input, and (3) longitudinal translation of the fiber input to simulate the range-dependent variation of the spot size at the fiber input.

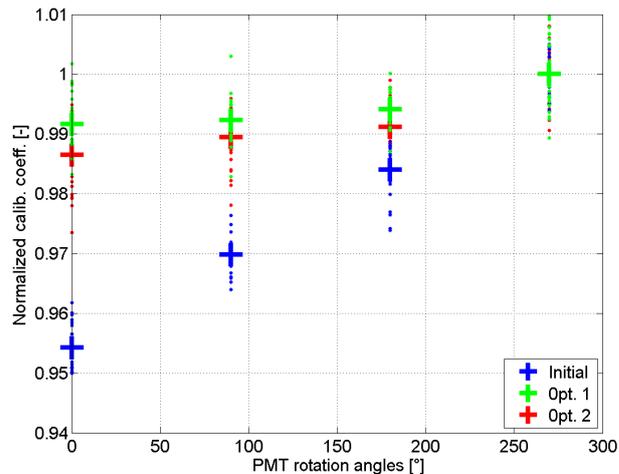
the initial configuration corresponding to the DEMEVAP campaign using a 0.8 mm fiber and subject to vignetting, option 1 using a 0.4 mm fiber and the optimized optical layout, and option 2 using a 1 mm fiber with a shortened and optimized optical layout.

### 3.1 PMT photocathode spatial nonuniformity

In this experiment (no. 1) the PMT is rotated 90 by 90° to explore four different positions of the beam on the photocathode surface. For each position, 5 min of measurements were made and the normalized signal ratios were determined. The experiment was repeated for all three detection subsystem configurations. Figure 5 shows the results. With the initial optical configuration the maximal variation between two PMT positions is 4.6 %, which is on the order of the calibration factor variations observed during DEMEVAP (Bock et al., 2013). With the two optimized configurations, the variations are 1 and 1.3 %, respectively. The complete experiments were repeated several times and led to the same conclusion – that both optimized configurations are significantly more stable than the initial configuration. These results confirm that spatial nonuniformity of the PMT photocathodes combined with beam displacements and vignetting were important factors of instability during DEMEVAP. Interestingly, the impact of spatial nonuniformity of the PMT photocathodes can be significantly reduced when vignetting is eliminated.

### 3.2 Spot movement on the fiber input

In experiment no. 2 we simulate the spot movement on the fiber input which is suspected in the real system to be induced by laser beam wandering and thermo-mechanical deformations of the optical bench. Displacing the fiber head with respect to the light source allows us to scan all spot

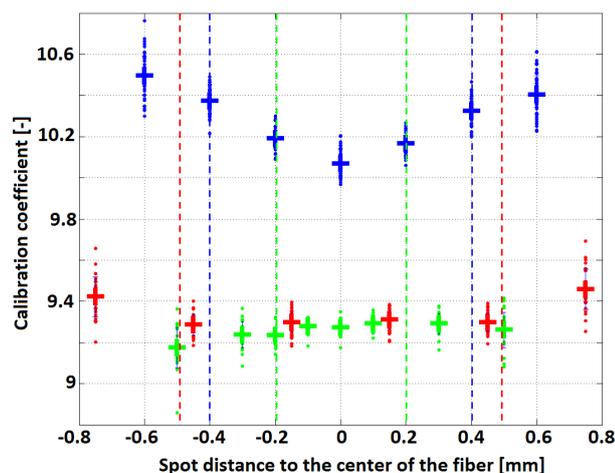


**Figure 5.** Normalized calibration coefficient as a function of PMT rotation angle (test no. 1 in Fig. 4) for three optical configurations of the detection subsystem: (blue) initial configuration (similar to the DEMEVAP campaign), (green) optimized configuration no. 1 (0.4 mm fiber), and (red) optimized configuration no. 2 (1 mm fiber).

positions across the fiber entrance (Fig. 4). As mentioned above, these spot movements produce beam mode fluctuations at the fiber exit and thus beam shape and position variations on both PMT photocathodes. As for experiment no. 1, we repeated measurements for all three detection subsystem configurations. The results are illustrated in Fig. 6. Again we see that with the initial configuration the variations are much larger than with the two optimized configurations. When the results are only considered for fiber displacements over a distance equivalent to one fiber diameter, the variations are about 3.5 % for the initial configuration and about 0.5 % for the two optimized configurations. Note that measurements were also feasible beyond the fiber diameters because of the size and divergence of the incident beam. For positions exceeding the fiber diameters, the signals were much weaker and the results noisier. The signal ratio variations increased up to 1–2 % for the two optimized configurations. These results confirm the previous ones and the hypothesis that displacements of the incident beam at the fiber input can be responsible for significant variations in the instrumentation calibration factor. With an optimized optical layout, the effect can be limited to 1 %.

### 3.3 Range dependence of the spot size

The variation illustrated in this section is not due to alignment and/or instrumental changes but to the systematic range dependence of the spot size at the fiber input. Indeed, when the fiber is placed at or close to the image focus of the telescope, rays coming from far objects will be collected. For objects closer than a critical distance, the image in the plane of the fiber becomes larger than the fiber aperture. The vari-

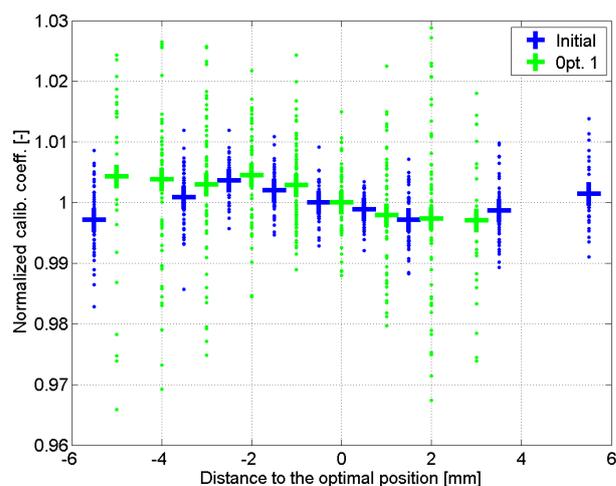


**Figure 6.** Normalized calibration coefficient as a function of lateral spot displacement (test no. 2 in Fig. 4) for three optical configurations of the receiving module (similar to Fig. 5). The vertical dashed lines show the limits of lateral displacements for which less than half of the spot actually enter the fiber.

ations of spot size at the fiber input produce variations of the NA and beam modes at the fiber output which are detrimental to the stability of the measurements because of the PMT spatial nonuniformity. One can argue that there is also an angular variation as the beam propagates away from the lidar in the operational configuration, but it is only of order 1 mrad and can thus be neglected compared to the numerical aperture of our system ( $NA = 0.22$  with a 0.4 mm fiber).

In experiment no. 3 we displace the fiber with respect to the light source to simulate a variation in spot size. This experiment was repeated only for the initial and one of the optimized configurations. Results are shown in Fig. 7. For both configurations, the effect of range dependence of the spot size on the signal ratio is about 0.7%. This effect is generally smaller than the two studied in previous sections. It is also significantly smaller than the variations reported previously by Simeonov et al. (1999) due to range-dependent spot movements on the PMT photocathode, mainly because, in the case of our system, the spot size on the PMT is much larger (5 mm, see Sect. 2) and the beam variations at the fiber output are tempered compared to the variations at the fiber input.

One can notice that the individual results in Fig. 7 are more scattered with the optimized configuration than in the previous experiments. This is due to the fast decrease in the signal-to-noise ratio (SNR) of the measured signals when the spot size exceeds the size of the fiber input. In conclusion, the range dependence of the spot size has thus a rather small impact on the calibration factors. When we repeated the experiment several times, it appeared that the best results (presented here) are only achieved when the beam in the detection subsystem is carefully aligned, however.



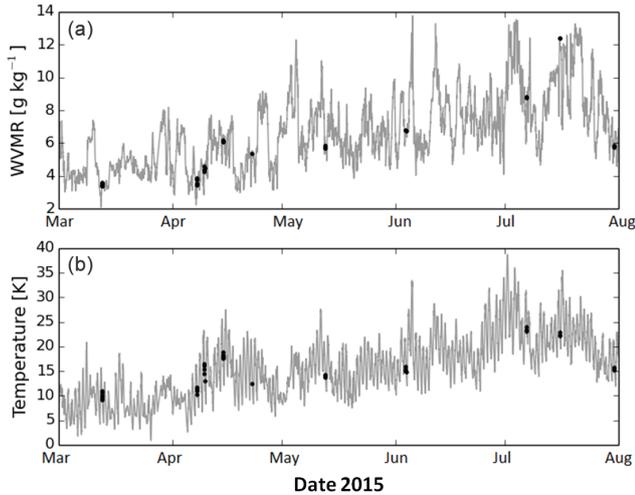
**Figure 7.** Normalized calibration coefficient as a function of distance between the light source and fiber input (test no. 3 in Fig. 4) for two optical configurations of the receiving module: (blue) initial configuration (similar to the DEMEVAP campaign), and (green) optimized configuration no. 1 (0.4 mm fiber).

#### 4 Instrumental stability monitoring during experimental campaign

The validation campaign was conducted at the IGN facility in Saint-Mandé between March and July 2015. The Rameau lidar system was installed in a small van equipped with a rooftop aperture through which the laser beam was transmitted and backscatter signal was collected by the telescope. The lidar measurements were collected during 14 nighttime experiments with clear sky conditions. Each experiment consisted in the acquisition of a number of 5 min sequences (or sessions) of either of two types of measurements: (1)  $N_2$  calibration measurements or (2) water vapor measurements. Table 1 lists the number and type of measurements for all 14 nights. Three GPS receivers (Trimble Net R9) equipped with PTU sensors (Vaisala PTU200) measured ZWD, pressure, temperature, and humidity continuously during the 5 month period. Two PTU sensors and GPS receivers were mounted on the top of a building, about 15 m above the lidar system. Another PTU sensor was operated from a higher nearby building (25 m). The PTU sensors were inter-calibrated with a common standard before the campaign. This setup allowed us to have well-collocated lidar, PTU, and GPS measurements for the external calibration. Figure 8 presents the time series of WVMR and the temperature measured by one of these PTU sensors. It can be seen that the 14 lidar experiments sampled very different atmospheric conditions in terms of temperature and humidity.

Compared to DEMEVAP, several modifications were brought to the Rameau lidar system:

- The optimized configuration no. 1 was implemented with a 0.4 mm diameter fiber.



**Figure 8.** Time series of WVMR (a) and temperature (b) measured by a PTU sensor during the Saint-Mandé campaign in 2015. Data are plotted with 5 min time resolution. The black dots indicate the nights when the Rameau lidar system was operated.

- A new three-axes micrometer positioning stage was installed to more accurately control the 3D position of the first lens of the beam expander. First, this allowed us to set the distance between the two lenses more precisely ( $Z$  axis). Second, it enabled us to more easily steer the transmitted laser beam ( $X$ – $Y$  axis), especially for correcting the drifts in the transmitter and receiver alignments. These drifts were controlled continuously from the elastic and Raman signal returns at a 1 km distance, and corrections were applied when the signals dropped by more than 10 % within 5 min.
- More effort was put on the adjustment of the optical elements throughout the optical path to avoid vignetting. This adjustment is performed by tracking the beam – formed by the light of the 468 mm LED – with different millimetric patterns.
- As for the electronic part, a checkup of the transient recorder by Licel GmbH led to the replacement of the preamplifiers. As a consequence the pulse amplitudes changed and discrimination levels in the acquisition software had to be modified accordingly. They were determined using the pulse height distribution technique recommended by Licel. The new levels were also validated by checking that the detected photon rates matched properly to the expected Poisson law (mean photo count equal to variance). Deviations from Poisson law can be observed when the discriminator level is too low and ringing in the coaxial cables produces multiple pulses in response to one detected photon.

#### 4.1 Water vapor profile retrieval and calibration

The signals were acquired with a spatial resolution of 7.5 m and temporal resolution of 20 s (average over 400 laser shots). After background and saturation correction, the Raman signals are further averaged in 5 min time bins to improve their SNR and the WVMR is computed following the equation below (Whiteman, 2003b):

$$r(z) = C_{\text{lidar}}(z) \frac{S_{\text{H}_2\text{O}}(z) - B_{\text{H}_2\text{O}}}{S_{\text{N}_2}(z) - B_{\text{N}_2}}, \quad (1)$$

where  $S_X$  is the signal measured in channel  $X$  and  $B_X$  is the background ( $X = \text{H}_2\text{O}$  or  $\text{N}_2$ ). The calibration function  $C_{\text{lidar}}(z)$  can be decomposed as follows :

$$C_{\text{lidar}}(z) = r_{\text{N}_2} \frac{M_{\text{H}_2\text{O}}}{M_{\text{N}_2}} \frac{O_{\text{N}_2}(z)}{O_{\text{H}_2\text{O}}(z)} \frac{\xi_{\text{N}_2}(\lambda_{\text{N}_2})}{\xi_{\text{H}_2\text{O}}(\lambda_{\text{H}_2\text{O}})} \frac{\tau(z, \lambda_{\text{N}_2})}{\tau(z, \lambda_{\text{H}_2\text{O}})} \frac{\frac{d\sigma_{\text{N}_2}(\pi)}{d\Omega}}{F_{\text{N}_2}(T(z))} \frac{F_{\text{H}_2\text{O}}(T(z))}{\frac{d\sigma_{\text{H}_2\text{O}}(\pi)}{d\Omega}}, \quad (2)$$

where  $r_{\text{N}_2}$  is the mass mixing ratio of nitrogen,  $M_X$  the molecular weight of the species  $X$  ( $\text{H}_2\text{O}$  or  $\text{N}_2$ ),  $O_X$  the overlap function,  $\xi_X$  the instrumental transmission and detection efficiency of the optical and electronic elements in channel  $X$ ,  $\tau(z, \lambda_X)$  the atmospheric transmittance from ground to distance  $z$ ,  $\frac{d\sigma_X(\pi)}{d\Omega}$  the Raman backscattering cross section of the species  $X$ , and  $F_X(T(z)) \frac{d\sigma_X(\pi)}{d\Omega}$  may be interpreted as the effective molecular cross section consistent with the use of a monochromatic optical efficiency term (Whiteman, 2003b). Where necessary, the wavelength is explicitly indicated ( $\lambda_{\text{H}_2\text{O}} = 407.6$  nm or  $\lambda_{\text{N}_2} = 386.7$  nm).

The ratio of the overlap functions was determined from the ratio of signals in the  $\text{H}_2\text{O}$  and  $\text{N}_2$  channels made using a common  $\text{N}_2$  filter ( $\text{N}_2$  calibrations, see next subsection). We found that the ratio reaches unity at a distance of about 150 m and varies in the range of  $\pm 3$  % below. The determination of the overlap ratio in this way is not very accurate at the lowest elevations because the measurements are very noisy. We decided thus to avoid correcting for this effect. As a consequence, a small residual bias might affect the PTU calibrations results presented later because they use the short-range lidar measurements. The uncertainties associated with the other terms (ratios of  $C_X$ ,  $T_X$ , and  $\frac{d\sigma_X(\pi)}{d\Omega}$ ) are typically 5–10 % (Tarniewicz et al., 2002). As mentioned earlier, Venable et al. (2011) showed that it is possible to achieve independent calibrations with an accuracy better than 5 %.

Figure 9 shows an example of a WVMR profile measured by the Rameau lidar and by an operational radiosonde launched by Météo-France at Trappes about 30 km from Saint-Mandé. The integration time of the lidar profiles used in this work is 5 min, and the vertical resolution is between 7.5 and 240 m, depending on the altitude in the profile (Bossler et al., 2007). The radiosondes used at Trappes are MODEM M10. Although during the DEMEVAP campaign in 2011 the MODEM M10 sondes showed large biases

**Table 1.** Summary of lidar sessions during the campaign at Saint-Mandé in 2015. The numbers indicate the number of 5 min measurements of both types: H<sub>2</sub>O refers to standard water vapor configuration and N<sub>2</sub> to measurements realized with a common nitrogen filter (see Sect. 4.2 for further information).

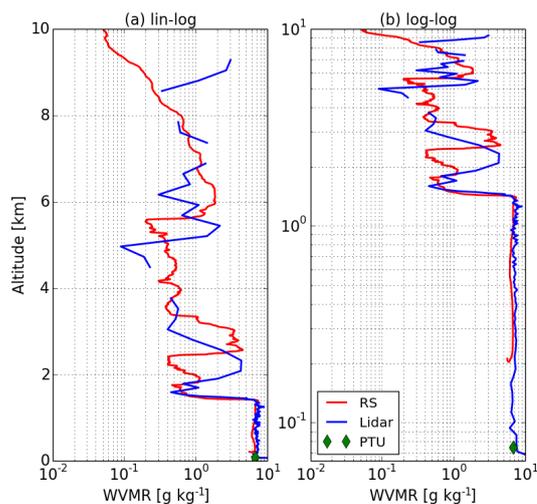
	12 Mar	7 Apr	9 Apr	14 Apr	22 Apr	11 May	12 May	3 Jun	6 Jul	15 Jul	30 Jul
H <sub>2</sub> O	6	5	5	3	1	0	4	4	3	3	3
N <sub>2</sub>	2	1	2	1	0	2	2	2	3	2	3

compared to the other instruments (Bock et al., 2013), a recent study showed that MODEM M10 sondes have improved and now have a very similar performance to the Vaisala RS92 sondes (Ingleby et al., 2016). However, because the radiosonde measurements are made at a distance of 30 km, the profiles do not agree well all the time, especially in the lower levels. This is a main reason why we did not use the radiosonde data as a source for external calibration of the lidar profiles such as done during DEMEVAP.

External calibration of the lidar WVMR profiles was achieved with two types of auxiliary measurements: relative humidity measurements from PTU sensors and ZWD data derived from GPS measurements following the methodology described in Bock et al. (2013). It can be questioned whether the portion of atmosphere between these measurements and the first valid lidar range bin (33.75 m) would introduce a significant bias in the calibration. To evaluate this effect we estimated the mean vertical gradient in WVMR from simultaneous humidity measurements from PTU sensors at two different heights (15 and 25 m). We found a mean difference in WVMR of 0.1 g kg<sup>-1</sup>. Extrapolating the PTU WVMR values linearly to the height of the first valid lidar range bin gives an estimate of the bias of 0.5 g kg<sup>-1</sup> or 5 %, which is quite large. However, this result should be moderated by the fact that the observed vertical gradient in WVMR from our lidar measurements is much closer to zero in the surface layer (Fig. 11b). It is thus expected that this bias is much smaller than 5 %. From this we conclude that neither our PTU measurements nor our ZWD estimates computed from the lidar profile data need be adjusted for the vertical displacement with respect to the lidar data. Regarding the upper layers, the lidar profile data were completed with radiosonde profile data between 5 and 10 km altitude for the ZWD computation. Beyond 10 km the water vapor has negligible contribution to ZWD. Note that our calibration procedure using ZWD data is completely equivalent to the more usual one using IWV data (Bossler et al., 2010).

#### 4.2 Monitoring system stability with the N<sub>2</sub> calibration procedure

The instrumental stability of the Rameau lidar system was monitored during the Saint-Mandé campaign by means of N<sub>2</sub> calibration measurements made with a common N<sub>2</sub> filter placed at the entrance of the receiving module (after L1, see Fig. 1), with the other interference filters being removed. This kind of calibration procedure was first described

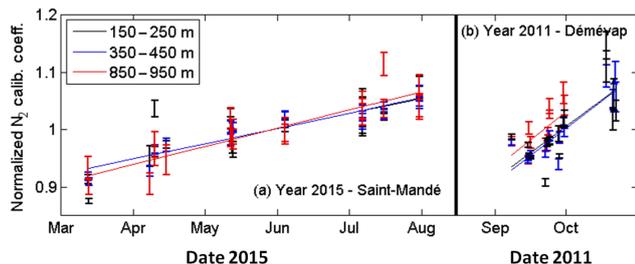


**Figure 9.** Vertical water vapor mixing ratio profiles measured by the Rameau lidar (blue) on 3 June 2015 at 20:42 at Saint-Mandé and a radiosonde (red) on 3 June 2015 at 23:30 at Trappes (30 km from Saint-Mandé): (a) altitude in linear scale, and (b) altitude in log scale. The PTU measurement at 15 m above ground level is also shown (green diamond). The lidar profile was calibrated with the PTU data from 12 March 2015.

by Vaughan et al. (1988) and Whiteman et al. (1992). It should be noted that in operations different filters are used in both channels and measurements are made at different wavelengths (386.7 and 407.6 nm). The calibration factor derived from the N<sub>2</sub> calibration measurements needs to be adjusted to provide absolute WVMR profiles. This adjustment is done by comparison with external WVMR measurements (see next subsection).

Two sequences of N<sub>2</sub> calibration measurements were made each night. Measurements from the N<sub>2</sub> and H<sub>2</sub>O channels were acquired over 5 min. The N<sub>2</sub> calibration factor was computed as a ratio of the mean signals measured in the H<sub>2</sub>O and N<sub>2</sub> channels on a selected layer. Three different layers were tested: 150–250, 350–450, and 850–950 m. The former two had an SNR above 10 on average in both channels, while the latter usually had a lower SNR and was thus more noisy. In general, the results from the three layers were fairly consistent.

The results are shown in Fig. 10 and Table 2. Despite the fact that the new configuration has been implemented in the detector subsystem and optical alignments have been care-



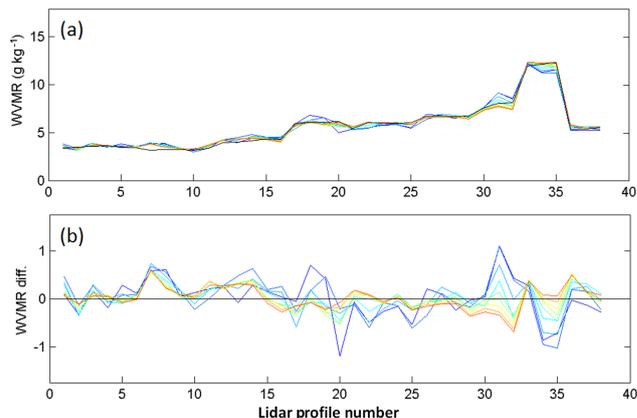
**Figure 10.** Time series of normalized  $N_2$  calibration coefficients from (a) the Saint-Mandé campaign and (b) the DEMEVAP campaign. Three different layers are compared: 150–250, 350–450, and 850–950 m. Normalization is done with respect to the mean value for each layer. The smaller number of points for the higher layer in the case of DEMEVAP is due to removal of measurements with large standard errors.

**Table 2.** Comparison of the stability of the  $N_2$  Calibration coefficients for the Saint-Mandé and DEMEVAP campaigns in three different layers. The slope is estimated by a linear least squares regression. The last column represents the dispersion ( $1-\sigma$ ) obtained after removing the linear trend from the data.

		Slope (% month <sup>-1</sup> )	Dispersion (%)
Saint-Mandé	150–250 m	$2.5 \pm 1.4$	3.1
	350–450 m	$2.6 \pm 0.73$	1.6
	850–950 m	$3.0 \pm 1.1$	2.3
DEMEVAP	150–250 m	$9.2 \pm 2.0$	3.8
	350–450 m	$9.4 \pm 1.2$	2.3
	850–950 m	$10 \pm 3.7$	2.6

fully controlled, a drift of 2.5–3.0 % per month is still observed during the Saint-Mandé campaign. However, this drift is 3 times smaller than the one observed in the DEMEVAP campaign (9–10 % per month). So it can be stated that the modifications significantly improved the stability of the lidar system. The drifts still observed cannot be explained by aging of optical or electronic components but rather by long-range dependence of calibration changes due to realignments of the transmitted laser beam or interventions on the detection subsystem (e.g., change of optical fiber during DEMEVAP or tests with different interference filters during the Saint-Mandé campaign).

The dispersion of the calibration results reported in Table 2 is also improved during the Saint-Mandé campaign compared to DEMEVAP (1.6 % vs. 2.3 % at best). We think that the reduction of the short-term fluctuations is due to the use of a smaller optical fiber (0.4 mm) which exhibits less mode fluctuations and a better centering and spreading of the beam on the PMTs.



**Figure 11.** Comparison of WVMR measurements from lidar and PTU during the Saint-Mandé campaign: (a) WVMR from lidar from nine different layers (color lines) and PTU (black line) and (b) WVMR difference (lidar – PTU). The lidar measurements were corrected beforehand for a linear drift with  $N_2$  calibration measurements and adjusted to absolute WVMR using the PTU measurements from 12 March 2015. The nine different layers start at a height of 33.75 m and have increasing widths from 1 to  $9 \times 7.5$  m (with colors going from blue to red). The  $x$  axis refers to lidar profile numbers in chronological order (see  $H_2O$  row in Table 1 for correspondence with dates).

### 4.3 Water vapor calibration with external measurements

The water vapor calibration factors were determined using two different techniques: humidity measurements from PTU sensors (referred to as the “PTU method” in the following) and ZWD data from GPS measurements (referred to as the “GPS method”). For each method, a set of parameters was examined in order to find the most stable time series of calibration factors.

For the PTU method, different widths and heights of the lidar profile were tested. Figure 11 shows coincident lidar and PTU WVMR measurements for the whole period of the campaign. There is very good agreement between both measurement time series. The minimal root mean square error (RMSE) was  $0.25 \text{ g kg}^{-1}$  absolute or 5 % relative for a layer starting at a height of 33.75 m with a width of 67.5 m.

For the GPS method, we sought the best way to complete the lidar profile in the upper layers testing different starting heights for the radiosonde. The lidar profiles were considered up to 5 km only and completed above with the radiosonde profiles from Trappes up to 10 km. The fraction of the lidar ZWD represents more than 90 % of the total ZWD. Assuming that the accuracy of the sonde measurements in the layer between 5 and 10 km is about 10–20 %, the accuracy of the correction is about 1–2 %. This is the expected accuracy of the GPS ZWD calibration technique.

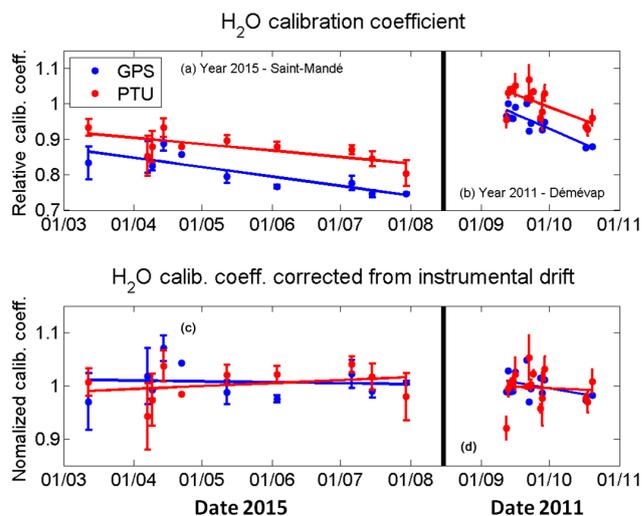
The final results are presented in the top panel of Fig. 12, on the left for the Saint-Mandé campaign and on the right for

the DEMEVAP campaign, where the PTU method is plotted in red and the GPS method in blue. The upper panel shows the H<sub>2</sub>O calibration factors determined from raw lidar measurements, while the lower shows the calibration factors determined after a slope correction based on the N<sub>2</sub> calibration results (Sect. 4.2). Regarding the N<sub>2</sub> calibration, we chose the 350–450 m layer for the correction because of its higher SNR and smaller dispersion (see Table 2). The slope correction was fitted from the N<sub>2</sub> calibration data as a function of time. We did not use the night to night coefficients correction approach described by Whiteman et al. (1992) because our N<sub>2</sub> calibration coefficients are noisy and we did not want to add more dispersion to the H<sub>2</sub>O calibration coefficients. The results during the drier months of the Saint-Mandé campaign (March and April) are more scattered (see Fig. 8). The difference in the mean calibration constants derived from GPS and PTU methods (observed during both campaigns) can be partly explained by the differential overlap function which is not corrected. During the Saint-Mandé campaign, the vertical displacement between the PTU measurements and the lidar profile data discussed in Sect. 4.1 might contribute as well.

Table 3 reports the slopes of linear regression and the dispersion of the water vapor calibration factors before and after the slope correction based on the N<sub>2</sub> calibration results. The slopes estimated from the uncorrected data amount to 2–3 % per month for the Saint-Mandé campaign and 7–9 % per month for DEMEVAP. They are fairly consistent with the slopes determined from the N<sub>2</sub> calibrations (Table 3). As a consequence, the linear correction based on the N<sub>2</sub> calibration results is able to eliminate the overall drift almost completely. Thanks to the N<sub>2</sub> calibration correction, the residual drift is smaller than 1 % per month for the Saint-Mandé campaign and 2.2 % per month for the DEMEVAP campaign. The dispersion of the ZWD calibration coefficients after slope correction is at the 3 % level for both campaigns. The increased dispersion of the water vapor calibration factors compared to the N<sub>2</sub> calibration factors (Table 3) is explained by larger noise in the water vapor lidar measurements. The slightly smaller dispersion for the GPS calibration results of the DEMEVAP campaign is similarly explained by reduced noise in the DEMEVAP water vapor measurements thanks to longer integration time (20 min compared to 5 min in Saint-Mandé). During DEMEVAP, the lidar ZWD estimates also benefited from collocated radiosonde data launched twice a night which were used to complete the profiles in the upper levels, contrary to the Saint-Mandé campaign during which only one non-collocated radiosonde profile was available per night.

## 5 Summary and conclusions

In this paper we have investigated the origin of instrumental calibration instabilities observed with the Rameau water vapor Raman lidar system. Based on the results of an ear-



**Figure 12.** (a, b) Time series of lidar calibration factors determined with the GPS method (blue) and the PTU method (red). (c, d) calibration factors after correction of a linear drift fitted from the N<sub>2</sub> calibration results (350–450 m layer). The results from the Saint-Mandé campaign are plotted on the left part and the results from the DEMEVAP campaign on the right part.

lier campaign, DEMEVAP, we established a list of possible short-term (< 1 h) and long-term (> 1 month) instability sources. We have shown that laser beam wandering at the fiber input causes size and shape variations of the beam at the fiber output which can generate short-term differential variations in the response of PMT signals because of spatial inhomogeneity of the PMT photocathodes. In addition to the spatial inhomogeneity of the PMT photocathodes, high-voltage fluctuations – caused for instance by temperature fluctuations – may also have generated short-term fluctuations in gain.

Although every lidar system has its own instrumental characteristics, the use of PMTs is well generalized and this instability source is certainly a fundamental cause of short-term and long-term instability in many lidars. Because it is impossible to eliminate completely laser beam wandering and jitter, short-term fluctuations are to be expected in all systems, whether they use optic fibers or not. However, mode fluctuations observed with fibers might produce additional instability.

A careful investigation of all optical components revealed that some of the optic fibers used to couple the telescope to the detection subsystem of the Rameau lidar had a larger numerical aperture than specified, which led to unexpected vignetting issues that could have affected measurements of the DEMEVAP campaign. We have shown that the redesign of the optical layout could get rid of the vignetting issue. Combined with a rigorous alignment of the beam in the detection subsystem, these changes allowed us to significantly improve the stability of our measurements.

**Table 3.** Comparison of the stability of the PTU and GPS H<sub>2</sub>O calibration methods during the Saint-Mandé and DEMEVAP campaigns, before (left) and after (right) correction of a linear drift fitted from N<sub>2</sub> calibration results.

	Method	No correction		After N <sub>2</sub> calibration drift correction	
		Slope (% month <sup>-1</sup> )	Dispersion (%)	Slope (% month <sup>-1</sup> )	Dispersion (%)
Saint-Mandé	PTU	-2.11 ± 0.68	4.4	0.59 ± 0.68	3.1
	GPS	-3.35 ± 0.72	6.1	-0.68 ± 0.71	3.3
DEMEVAP	PTU	-7.45 ± 2.22	4.5	-0.58 ± 2.20	3.3
	GPS	-9.12 ± 1.67	4.2	-2.28 ± 1.17	2.3

We conducted a series of laboratory experiments to quantify the short-term calibration instabilities due to beam wandering and PMT photocathode inhomogeneity and to test the potential improvements from two optimized optical configurations of the detection subsystem. We have shown that the optimized optical configurations reduced the calibration fluctuation from 3.5 % to less than 1 %. Auto-alignment of the transmitter and the receiver would further improve the short-term stability. Commercial boresight alignment systems are capable of maintaining the pointing accuracy within 10–20 μm (Whiteman et al., 2012).

We also conducted an outdoor campaign to assess the impacts that the modifications had on the system in terms of its short-term as well as long-term stability (over 5 months). The detection subsystem stability was more specifically monitored with N<sub>2</sub> calibration measurements while the overall system stability was evaluated from H<sub>2</sub>O calibration coefficients determined with the help of PTU humidity measurements and GPS ZWD data. The N<sub>2</sub> calibration factors exhibited drifts in the long term of 2–3 % per month. Drifts of similar amplitude (1 to 5 % per month) have also been reported in other Raman lidar systems (Brocard et al., 2013). These long-term drifts are thought to originate from a combination of the long-term memory of the system due to frequent realignments of the laser to the receiving telescope and occasional abrupt changes due to interventions on the detection subsystem (e.g., tests of interference filters requiring the opening of the detection box and change of filters). Compared to the DEMEVAP results, the long-term stability was nevertheless improved by a factor of 3. These results demonstrate that the optical design of our detection system has been significantly improved and that the overall lidar system has been well stabilized.

The short-term fluctuations in our system amounted to 2–3 % (1 standard deviation). Based on our experience from the earlier laboratory tests, we believe that these fluctuations are not due to random detection noise but to rapid variations in position and/or shape of the optical beams on the PMTs due to primarily the frequent realignment of the laser to the receiving telescope and to natural laser beam wandering.

We have found that the drifts observed in the H<sub>2</sub>O calibration coefficients are fairly consistent with those observed

in the N<sub>2</sub> calibration results. This confirms that they are due to changes in the response of the detection sensitivity of the two Raman channels. Correction of the lidar measurements with the N<sub>2</sub> calibration results is shown to almost completely eliminate the drifts in the lidar measurements. Hence, the combination of N<sub>2</sub> calibration and absolute calibration from PTU or GPS ZWD measurements appears as an interesting approach based on an idea similar to the hybrid calibration technique discussed by Leblanc and McDermid (2008) and Whiteman et al. (2011). The N<sub>2</sub> calibration measurement is quite easy to implement compared to the other techniques sometimes used. Moreover, there is no need for auxiliary aerosol measurements, and, contrary to the daytime measurement technique proposed by Sherlock et al. (1999b), it is not impacted by changes in differential atmospheric transmission. In the case of our system, the implementation of the N<sub>2</sub> calibration measurements could be improved by setting a spectral calibration light source directly at the entrance of the fiber, without going through the telescope. It would allow keeping the N<sub>2</sub> and H<sub>2</sub>O interference filters in place.

A major issue in our system is the need for frequent realignments of the laser to the receiving telescope. We believe this is a result of thermal deformations of the mechanical frame connecting the laser transmitter to the telescope. Strengthening of the mechanical frame and athermalization of the optical bench are considered to solve the problem. In addition, a small-scale version of the optical detection subsystem could also be foreseen. In such a small size system, the fiber would be removed and the receiving module would be placed at the focus the telescope. Such a miniaturized system would be less subject to vignetting and allow using micro-sized PMTs that can be expected to have more homogeneous photocathodes. Such improvements would allow the improved sounding of water vapor profiles and contribute usefully both to atmospheric research and GNSS vertical positioning applications.

*Data availability.* The data can be obtained on request to [olivier.bock@ign.fr](mailto:olivier.bock@ign.fr).

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors gratefully acknowledge Licel GmbH for fixing the instrumental problems on the transient recorder and the IGN-LOEMI staff (Olivier Martin, Yann Le Borgne) for participating in the nighttime measurements during the Saint-Mandé campaign. The authors would like to express their thanks to David Whiteman and a second anonymous reviewer for their constructive comments that helped improve the manuscript. This work was developed in the framework of project VEGA and supported by the French national program LEFE/INSU.

Edited by: Ulla Wandinger

Reviewed by: David Whiteman and one anonymous referee

## References

- Adam, M.: Notes on Temperature-Dependent Lidar Equations, *J. Atmos. Ocean. Technol.*, 26, 1021, <https://doi.org/10.1175/2008JTECHA1206.1>, 2009.
- Akgun, U., Ayan, A. S., Bruecken, P., Duru, F., Gülmez, E., Mestvirishvili, A., Miller, M., Olson, J., Onel, Y., and Schmidt, I.: Complete tests of 2000 Hamamatsu R7525HA phototubes for the CMS-HF Forward Calorimeter, *Nuclear Instruments and Methods in Physics Research A*, 550, 145–156, <https://doi.org/10.1016/j.nima.2005.03.171>, 2005.
- Avila, G.: Results on Fiber Characterization at ESO, in: *Fiber Optics in Astronomy III*, edited by: Arribas, S., Mediavilla, E., and Watson, F., vol. 152 of *Astronomical Society of the Pacific Conference Series*, p. 44, 1998.
- Avila, G., Fernández, J. M., Maté, B., Tejada, G., and Montero, S.: Ro-vibrational Raman Cross Sections of Water Vapor in the OH Stretching Region, *J. Mol. Spectrosc.*, 196, 77–92, <https://doi.org/10.1006/jmsp.1999.7854>, 1999.
- Avila, G., Fernández, J. M., Tejada, G., and Montero, S.: The Raman spectra and cross-sections of H<sub>2</sub>O, D<sub>2</sub>O, and HDO in the OH/OD stretching regions, *J. Mol. Spectrosc.*, 228, 38–65, <https://doi.org/10.1016/j.jms.2004.06.012>, 2004.
- Bock, O., Tarniewicz, J., Thom, C., Pelon, J., and Kasser, M.: Study of external path delay correction techniques for high accuracy height determination with GPS, *Phys. Chem. Earth A*, 26, 165–171, [https://doi.org/10.1016/S1464-1895\(01\)00041-2](https://doi.org/10.1016/S1464-1895(01)00041-2), 2001.
- Bock, O., Bossler, P., Bourcy, T., David, L., Goutail, F., Hoareau, C., Keckhut, P., Legain, D., Pazmino, A., Pelon, J., Pipis, K., Poujol, G., Sarkissian, A., Thom, C., Tournois, G., and Tzanos, D.: Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP, *Atmos. Meas. Tech.*, 6, 2777–2802, <https://doi.org/10.5194/amt-6-2777-2013>, 2013.
- Bosser, P., Bock, O., Thom, C., and Pelon, J.: Study of the statistics of water vapor mixing ratio determined from Raman lidar measurements, *Appl. Optics*, 46, 8170–8180, <https://doi.org/10.1364/AO.46.008170>, 2007.
- Bosser, P., Bock, O., Thom, C., Pelon, J., and Willis, P.: A case study of using Raman lidar measurements in high-accuracy GPS applications, *J. Geodesy*, 84, 251–265, <https://doi.org/10.1007/s00190-009-0362-x>, 2010.
- Brocard, E., Philipona, R., Haeefe, A., Romanens, G., Mueller, A., Ruffieux, D., Simeonov, V., and Calpini, B.: Raman Lidar for Meteorological Observations, RALMO – Part 2: Validation of water vapor measurements, *Atmos. Meas. Tech.*, 6, 1347–1358, <https://doi.org/10.5194/amt-6-1347-2013>, 2013.
- Dos Santos, J. M. F., Veloso, J. F. C. A., Lopes, J. A. M., and Morgado, R. E.: A simple method to improve the spatial uniformity of venetian-blind photomultiplier tubes, *IEEE T. Nucl. Sci.*, 43, 1335–1340, <https://doi.org/10.1109/23.507061>, 1996.
- Ferrare, R. A., Melfi, S. H., Whiteman, D. N., Evans, K. D., Schmidlin, F. J., and Starr, D. O.: A Comparison of Water Vapor Measurements Made by Raman Lidar and Radiosondes, *J. Atmos. Ocean. Tech.*, 12, 1177, [https://doi.org/10.1175/1520-0426\(1995\)012<1177:ACOWVM>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<1177:ACOWVM>2.0.CO;2), 1995.
- Freudenthaler, V.: Effects of Spatially Inhomogeneous Photomultiplier Sensitivity on LIDAR Signals and Remedies, in: *22nd International Laser Radar Conference (ILRC 2004)*, edited by: Pappalardo, G. and Amodeo, A., 37–40, 2004.
- GCOS-112: GCOS Reference Upper-Air Network (GRUAN): Justification, requirements, siting and instrumentation options, WMO Tech. Document, WMO/TD 1379, GCOS-112, 42 pp., 2007.
- Ghatak, A. K. and Thyagarajan, K.: *Optical Electronics*, Cambridge University Press, 1989.
- Hamamatsu Photonics K. K.: *Photomultiplier tubes*, 2007.
- Ingleby, B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., St James, J., Tennant, W., and Weedon, R.: Progress towards high-resolution, real-time radiosonde reports, *B. Am. Meteor. Soc.*, <https://doi.org/10.1175/BAMS-D-15-00169.1> 97(11), 2149–2161, 2016.
- Leblanc, T. and McDermid, I. S.: Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements, *Appl. Optics*, 47, 5592, <https://doi.org/10.1364/AO.47.005592>, 2008.
- Leblanc, T., McDermid, I. S., and Walsh, T. D.: Ground-based water vapor raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring, *Atmos. Meas. Tech.*, 5, 17–36, <https://doi.org/10.5194/amt-5-17-2012>, 2012.
- Penney, C. M. and Lapp, M.: Raman-scattering cross sections for water vapor, *J. Opt. Soc. Am.*, 66, 422–425, 1976.
- Revercomb, H. E., Turner, D. D., Tobin, D. C., Knuteson, R. O., Feltz, W. F., Barnard, J., Bösenberg, J., Clough, S., Cook, D., Ferrare, R., Goldsmith, J., Gutman, S., Halthore, R., Lesht, B., Liljegren, J., Linné, H., Michalsky, J., Morris, V., Porch, W., Richardson, S., Schmid, B., Splitt, M., van Hove, T., Westwater, E., and Whiteman, D.: The Arm Program's Water Vapor Intensive Observation Periods, *B. Am. Meteorol. Soc.*, 84, 217–236, <https://doi.org/10.1175/BAMS-84-2-217>, 2003.
- 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. Optics*, 38, 5838–5850, <https://doi.org/10.1364/AO.38.005838>, 1999a.
- Sherlock, V., Hauchecorne, A., and Lenoble, J.: Methodology for the Independent Calibration of Raman Backscatter Water-Vapor Lidar Systems, *Appl. Optics*, 38, 5816–5837, <https://doi.org/10.1364/AO.38.005816>, 1999b.

- Simeonov, V., Larcheveque, G., Quaglia, P., van den Bergh, H., and Calpini, B.: Influence of the Photomultiplier Tube Spatial Uniformity on Lidar Signals, *Appl. Optics*, 38, 5186–5190, <https://doi.org/10.1364/AO.38.005186>, 1999.
- Soden, B. J., Wetherald, R. T., Stenchikov, G. L., and Robock, A.: Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor, *Science*, 296, 727–730, <https://doi.org/10.1126/science.296.5568.727>, 2002.
- Takhar, P. S.: Resolution and Cathode Uniformity in Scintillation Counters, *IEEE T. Nucl. Sci.*, 14, 438–442, <https://doi.org/10.1109/TNS.1967.4324451>, 1967.
- Tarniewicz, J., Bock, O., Pelon, J., and Thom, C.: Raman lidar for external GPS path delay calibration devoted to high accuracy height determination, *Phys. Chem. Earth*, 27, 329–333, [https://doi.org/10.1016/S1474-7065\(02\)00008-6](https://doi.org/10.1016/S1474-7065(02)00008-6), 2002.
- 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, *J. Atmos. Ocean. Tech.*, 16, 1062, [https://doi.org/10.1175/1520-0426\(1999\)016<1062:TFHRLW>2.0.CO;2](https://doi.org/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. Meteor. Soc.*, 114, 1471–1484, <https://doi.org/10.1002/qj.49711448406>, 1988.
- Venable, D. D., Whiteman, D. N., Calhoun, M. N., Dirisu, A. O., Connell, R. M., and Landulfo, E.: Lamp mapping technique for independent determination of the water vapor mixing ratio calibration factor for a Raman lidar system, *Appl. Optics*, 50, 4622–4632, <https://doi.org/10.1364/AO.50.004622>, 2011.
- Wandinger, U.: Raman Lidar, p. 241, [https://doi.org/10.1007/0-387-25101-4\\_9](https://doi.org/10.1007/0-387-25101-4_9), 2005.
- Whiteman, D. N.: Examination of the traditional Raman lidar technique. I. Evaluating the temperature-dependent lidar equations, *Appl. Optics*, 42, 2571–2592, <https://doi.org/10.1364/AO.42.002571>, 2003.
- Whiteman, D. N.: Examination of the traditional Raman lidar technique. II. Evaluating the ratios for water vapor and aerosols, *Appl. Optics*, 42, 2593–2608, <https://doi.org/10.1364/AO.42.002593>, 2003b.
- Whiteman, D. N., Melfi, S. H., and Ferrare, R. A.: Raman lidar system for the measurement of water vapor and aerosols in the earth's atmosphere, *Appl. Optics*, 31, 3068–3082, <https://doi.org/10.1364/AO.31.003068>, 1992.
- Whiteman, D. N., Venable, D., and Landulfo, E.: Comments on "Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements", *Appl. Optics*, 50, 2170, <https://doi.org/10.1364/AO.50.002170>, 2011.
- Whiteman, D. N., Cadirola, M., Venable, D., Calhoun, M., Miloshevich, L., Vermeesch, K., Twigg, L., Dirisu, A., Hurst, D., Hall, E., Jordan, A., and Vömel, H.: Correction technique for Raman water vapor lidar signal-dependent bias and suitability for water vapor trend monitoring in the upper troposphere, *Atmos. Meas. Tech.*, 5, 2893–2916, <https://doi.org/10.5194/amt-5-2893-2012>, 2012.