# Long-term evolution of the calibration constant on a mobile/field campaign water vapour Raman lidar

5   Patrick Chazette, Julien Totems, Frédéric Laly

LSCE/IPSL, CNRS-CEA-UVSQ, University Paris-Saclay, CEA Saclay, 91191 Gif sur Yvette, France

Correspondence to: Patrick Chazette (patrick.chazette@lsce.ipsl.fr)

**Abstract.** Numerous field campaigns have been carried out to quantify the water vapour content of the atmosphere using vibrational Raman lidar technology. Each of them raises the question of calibration methods, in particular the reliability of this calibration over time. We present a study on the stability of the calibration of the WALI lidar developed at Laboratoire des Sciences du Climat et de l'Environnement in France (LSCE), over a period of 7 years (2016-2022) and across several field campaigns. A calibration method is applied mainly using radiosondes and, in a few cases, airborne meteorological probes. Complementing the previous approaches, we show that ground-based meteorological measurements can be of great value for lidar calibration under conditions of vertical stability in the lower troposphere and of good knowledge of the lidar overlap function, with full overlap within the planetary boundary layer. Using statistical criteria, we emphasize that these three calibration approaches should remain consistent over time. The observation periods considered here allow us to sample a wide range of water vapour contents in the lower troposphere, from 0.5 g kg$^{-1}$ to more than 10 g kg$^{-1}$ characteristic of the variabilities expected over the mid-latitudes and even over the Arctic. From comparisons between lidar and in situ measurements (radiosondes and/or flights), we observe a variability of more than 10% in the calibration constant between field experiments conducted with and without laser injection seeding. The root mean square error between the lidar and in situ reference measurements is between 0.23 and 0.6 g kg$^{-1}$, mainly due to the atmospheric variability during the calibration. The bias is small, less than 0.08 g kg$^{-1}$. For all the situations studied, the correlation coefficient remains high, above 0.75. The instrumental error is comparable to the 0.4 g kg$^{-1}$ recommended by the World Meteorological Organization (WMO). Such a precision requires the use of a significant number of reference profiles and the remaining limitation is due to the uncertainties associated with in situ weather sensors. We note that the use of ground-based measurements does not introduce any more uncertainty in the lidar

calibration coefficient than vertical profiles obtained by radiosondes or airborne means. Furthermore, the use of re-analyses can be an interesting option for calibration if the lidar profiles are not used in the models themselves, e.g. by means of assimilation.

Keywords: Calibration, Water vapour, Lidar, Raman, radiosonde, airborne

## 1   Introduction

The measurement of the water vapour mixing ratio (WVMR) using active lidar remote sensing based on inelastic Raman scattering of nitrogen and water vapour has been employed since the 1960s (Melfi and Whiteman, 1985; Whiteman, 2003; Ansmann et al., 1992; Cooney, 1970; Melfi et al., 1969; Vaughan et al., 1988). It has often been used as part of coordinated field campaigns to monitor atmospheric processes involving water vapour, both for meteorological purposes (Turner and Goldsmith, 1999; Reichardt et al., 2012) and for process studies (Lange et al., 2019; De Tomasi and Perrone, 2003; Labzovskii et al., 2018). Only recently the need for better coverage of the lower troposphere has emerged to improve constraints on the new generation of mesoscale models dedicated to weather forecasting (Flamant et al., 2021; Wulfmeyer et al., 2015). This need arises from the fact that water vapour mixing ratio (WVMR) is a crucial parameter to study for the energy balance of the troposphere (Held and Soden, 2000; IPCC, 2022).

The necessity to calibrate water vapour Raman lidar, with reference profiles that are unbiased and coincident in time and space with the lidar measurements, strongly limits their use. There have been few inter-comparison campaigns, which are generally based on a limited number of reference profiles, often obtained from radiosondes (Bock and Nuret, 2009; Agusti-Panareda et al., 2009; Chazette et al., 2014b; Di Girolamo et al., 2020). The number of reference profiles is limited by the available resources and the weather conditions encountered during the field experiments. The use of a tungsten calibration lamp has also been proposed (Leblanc and McDermid, 2008), but it has been shown that this method does not to characterize the transmittance of the full lidar system and may introduces biases (Whiteman et al., 2011). More recently, other authors have tested other ways to obtain reference profiles, such as using a kite (Totems and Chazette, 2016) or aircraft (Totems et al., 2019; Chazette et al., 2014b), but these are often more difficult to implement than radiosondes.

About the temporal stability of the calibration, Ferrare et al., (1995) showed that the NASA/GSFC Raman water vapour lidar remained stable within better than 5% over a 2 year period, despite several modifications to the instrument. Bock et al. (2013) and their colleagues

David et al. (2017) have also published the drift of their lidar calibration coefficient over time during the Development of Methodologies for Water Vapour Measurement field campaign (DEMEVAP), in the former work, and investigated the causes of this drift or proposed techniques to remediate it, in the latter work. They explained that the 10 to 15% linear drift observed over a few months was mainly due to beam wandering in their reception optical system, as evidenced by the efficiency of the so-called "$N_2$-calibration technique" (Vaughan et al., 1988) to correct it. They showed that an optimized system, with a larger optical fibre (1 mm) and no vignetting in the wavelength separation apparatus, was much more stable (<1.5% in a laboratory demonstrator).

Lately, water vapour Raman lidars have been deployed during field campaigns carried out for the Water Vapour Lidar Network Assimilation (WaLiNeAs) research program. During the WaLiNeAs campaign of 2022-2023, various European partners collaborated to install a significant number of ground-based lidar stations to form a coherent database of lidar profiles useful for assimilation into operational weather models (Flamant et al., 2021; Laly et al., 2024). Data assimilation requires measured profiles to be as unbiased as possible, as biases are not corrected in the weather forecast models. Data must therefore be debiased from the model analyses before any assimilation (Fourrié et al., 2019). Intrinsically, Raman lidar measurements have few bias. The bias on the WVMR is generally due to the in-situ measurements used for calibration (Totems et al., 2021). It is thus reduced by the availability of multiple reference profiles, or even better, obtained through various experimental approaches.

Operating a Raman lidar network dedicated to measure the vertical profile of the WVMR in the lower troposphere requires proven calibration methods and instruments with low drift over time. In response, the WALI (Water vapour and Aerosol Lidar, now renamed Weather and Aerosol LIdar) Raman lidar (Chazette et al., 2014b), has been maintained in a consistent configuration since 2019, enabling us to assess the drift on the calibration coefficient. A complete error assessment of the instrument has already been conducted using various methods, including end-to-end modelling (Totems et al., 2021), but it needed to be supplemented by monitoring the temporal evolution of the calibration coefficient. This article presents such a study, made possible through field campaigns and dedicated measurements conducted between 2016 and 2022. Some of the campaigns have already been documented in the scientific literature as part of international research programs. Others remain unpublished as they stem from opportunities that have not yet been exploited.

The first campaign, Pollution in the ARCtic System (PARCS), took place in northern Norway in May 2016 (Totems et al., 2019). The second corresponds to the calibration/validation campaign for the ADM-AEOLUS space mission, which took place in April 2019 in the south of Paris. In June 2019, it was followed by the Lacustrine-Water vApour Isotope inVentory Experiment (L-WAIVE) campaign, which aimed to studying the water cycle over an Alpine valley (Chazette et al., 2021). A fourth field campaign took place from June to August 2020 following the health crisis associated with COVID. As part of the European LEMON project (Hamperl et al., 2021), the fifth field campaign took place in the south of France in September 2021. The sixth and final campaign was part of the Water Vapour Lidar Network Assimilation (WaLiNeAs) project, from November 2022 to January 2023 (Laly et al., 2024).

Section 2 briefly describes the WALI instrument, and the method used to derive the WVMR, followed by an introduction to the calibration method and the statistical parameters used to validate it. The other instruments and the field campaigns are presented in Section 3, with an overview of the time-series of the WVMR profile compared to model reanalyses. Section 4 presents the main study of the calibration coefficient statistics across the field campaigns. Section 5 discusses the results and concludes.

## 2 Water vapour Raman lidar and theory

### 2.1 Instrument

For atmospheric research activities, the ground-based Mobile Atmospheric Station (MAS, Raut and Chazette (2009)) van has been equipped with the 354.7 nm water vapour Raman lidar WALI since 2012 (Chazette et al., 2014b). Its emitter is a pulsed Nd:YAG laser (Lumibird, formerly Quantel, Q-Smart 450), tripled in frequency to 354.7 nm, and expanded to fulfil eye safety requirements just 1.5 m from the output window. The lidar system evolved between 2012 and 2017 (Totems et al., 2016, 2021) with the introduction of a laser injection seeder, and a fibre-coupled 150 mm diameter telescope on the Raman channels instead of a 150 mm diameter refractive optical system.

The main characteristics of WALI are given in Table 1. The ultraviolet (UV) pulse energy is ~100 mJ and the pulse repetition frequency is 20 Hz. Its field-of-view of $0.67 \times 2$ mrad (refractor) or 1.6 mrad (reflector with 1 mm diameter fibre) ensures full-overlap of the transmit and receive paths beyond ~150 m or ~200 m, while limiting the sky background noise. The gain of the H10721-210 photomultiplier tube (PMT) detectors can be adjusted (by a factor of ~75, as function of the dynode high voltage) to best accommodate sky background levels between daytime and nighttime within the detector's dynamic range. The gain function ratios

can be estimated within ±1.3% RMS (Totems et al., 2021). The fibre is a solarization resistant fused silica 1-mm diameter step index fiber, 2 m in length. The wavelength separation optical system was upgraded with lenses putting detector surfaces in the pupil plane. Since 2019, its configuration has remained unchanged, with only the geometric factors being adjusted for each

5 campaign following revisions to the laser. Also note that the photomultiplier tube (PMT) detectors have remained on their dedicated channels since 2016. The main limitations of this lidar remain the reliability of the laser and its injection seeder.

Table 1. Main characteristics of the Weather and Aerosol Lidar (WALI) from its initial configuration before 2016 to its final configuration after 2019.

| WALI lidar | Initial (before 2016) | Final (from 2019) |
|---|---|---|
| Laser injection seeder | None | NP Photonics fiber laser, 10 mW @ 1064 nm |
| Laser emitter | Lumibird (ex-Quantel) Q-Smart 450 100-120 mJ, 20 Hz @ 354.72 nm (in vacuum) | |
| Beam expander | 10x (-30-mm/300-mm focal length) UVFS afocal | |
| Receiver optics | Corrected refractor 150-mm diameter, 300-mm focal length | UV-enhanced Al-coated parabolic mirror, 152-mm diameter, 610-mm focal length |
| Fibre | None | Avantes solarization-resistant multimode fiber patch cable, 1000 µm diameter, 2 m length |
| Polychromator | Edmund Optics NUV 50-mm focal length achromat #65976 + Semrock 25x36 mm dichroic plates cutting at 365 and 395 nm | |
| Interference filters | Materion 1''-diameter filters, OD>4 out-of-band blocking $N_2$: 386.76 nm CWL, 0.27 nm FWHM, 62% max T $H_2O$: 407.59 nm CWL, 0.34 nm FWHM, 46 % max T | |
| Last lens | 1''-diameter, 25-mm focal length asphere, spot size on PMT ~2.5 mm | Afocal beam reducer, spot size on PMT ~6 mm |
| Detector | Hamamatsu Photo Multiplier Tubes (PMT) H10721-210 | |
| Signal acquisition | National Instruments PXI 5124 Oscilloscopes, 200 MHz, 12 bits | |

## 2.2   Lidar-derived water vapour mixing ratio

### 2.2.1   Basic equation

From the water vapour ($H_2O$, $H$) and nitrogen ($N_2$, $N$) channels, the lidar-derived WVMR ($r_l$, $l$ is the subscript for the lidar profiles) is calculated after calibration, in g kg$^{-1}$ against the altitude

15 $z$, as follows:

$$r_l(z) = K^V \cdot OR(z) \cdot \frac{S_H^V(z)}{S_N^V(z)} \cdot M_c(z) \cdot A_c(z) \tag{1}$$

With:

- $S_i^V$ the lidar signal of channel $i$ at the high voltage $V$ after down-sampling, accumulation and possible merging of the analog and desaturated photon-counting channels. It is corrected from the sky background, the detector dark count, and the electronic baseline.

- $K^V$ the calibration constant determined by comparison with vertically resolved atmospheric soundings before or/and after the experiment. It is a function of the hight voltage (V) of the PMT, proportional to the ratio of the detection gains $G_i^V$, normalized at the calibration voltage $V_0$:

$$K^V = K_0 \cdot \frac{g_N^V}{g_H^V} \text{ with } g_N^V = \frac{G_N^V}{G_N^{V_0}} \text{ and } g_H^V = \frac{G_H^V}{G_H^{V_0}} \tag{2}$$

$K_0$ is the calibration constant related to reference voltages $V_0$. It can be determined once the $g_i^V$ functions have been defined under the condition $g_i^{V_o} = 1$ (see subsection 2.3.3).

- $O_R(z) = \frac{O_N(z)}{O_H(z)}$ with $O_H(z)$ and $O_N(z)$ the lidar overlap functions for the $H_2O$ or $N_2$ channels, respectively. They are determined from previous horizontal shots in a homogeneous atmosphere or from a coincident low-level radiosonde between the ground and ~1.5 km above the ground level (a.g.l.).

- $M_c(z)$ and $A_c(z)$ are the atmospheric corrections from the molecular and aerosol contributions, respectively.

On a time-average ($\langle \ \rangle$) of $M$ profiles, the WVMR is then expressed by:

$$r_l(z) = K_0 \cdot O_R(z) \cdot \frac{\langle S_H^V(z) / g_H^V \rangle_M}{\langle S_N^V(z) / g_N^V \rangle_M} \tag{3}$$

### 2.2.2  Atmospheric transmission corrections

The Raman lidar-derived WVMR requires a correction of the atmospheric transmittance variation between the Raman wavelengths used. Molecular transmittance is a function of air density and therefore of temperature and pressure profiles, which are either measured by the radiosonde used for calibration or extracted from model reanalyses with a relative error of less than 1%. Atmospheric transmission related to aerosols can be obtained via the aerosol optical thickness (AOT) derived from the $N_2$-channel (Royer et al., 2011; Chazette et al., 2014b):

$$AOT(z, z_0) = \frac{1}{1 + \eta_{aN}} \cdot \left[ \log\left( \frac{D_{atm}(z_0)}{D_{atm}(z)} \cdot \frac{S_N(z)}{S_N(z_0)} \right) - (1 + \eta_{mN}) \cdot MOT(z, z_0) \right] \quad (4)$$

*MOT* is the molecular optical thickness and $z_0$ the reference altitude for the calculation.

The atmospheric transmission corrective multiplicative terms are given by (Chazette et al., 2014b):

$$M_c(z) = exp\left( [\eta_{mH} - \eta_{mN}] \cdot \int_{z_G}^{z} \alpha_m(z') \cdot dz' \right)$$

$$A_c(z) = exp\left( \int_{z_G}^{z} [\eta_{aH}(z') - \eta_{aN}(z')] \cdot \alpha_a(z') \cdot dz' \right) \quad (5)$$

with:

$$\eta_{mN} = \left( \frac{387}{354.67} \right)^{-4.09} \text{ and } \eta_{mH} = \left( \frac{407}{354.67} \right)^{-4.09}$$

$$\eta_{aN}(z) = \left( \frac{387}{354.67} \right)^{-A(z)} \text{ and } \eta_{aH}(z) = \left( \frac{407}{354.67} \right)^{-A(z)} \quad (6)$$

$z_G$ is the altitude of the lidar station. *A* is the Ångström exponent associated with aerosols. We can take $A \sim 1$ with an impact on the WVMR error under 1% for $AOT = 0.2$. For aerosols with low spectral extinction dependence (sea salt, dust) the aerosol correction term can be neglected, and the same applies when the aerosol optical thickness is low ($< 0.1$-$0.2$ at 355 nm) (Chazette et al., 2014b).

### 2.3    Calibration process

Here we describe the method for estimating the calibration constant $K^V$ and the overlap ratio $O_R(z)$. We also deal with the optional variation of the PMT gains on the detection channels, which is necessary to maintain sufficient signal to noise ratio (SNR) over the day/night cycle. This variation entails a second calibration, i.e. the regression of the channel gain ratio $g_N^V / g_H^V$ as a function of the control high voltages.

### 2.3.1    Calibration coefficient

During the calibration process, lidar measurements are performed in parallel with a reference profile $r_{ref}$ obtained from an on-board meteorological sensor. Thus, $K_0$ is derived as:

$$K_0 = \frac{1}{O_R(z)} \cdot r_{ref}(z) \cdot \frac{\langle S_N^V(z) / g_N^V \rangle_M}{\langle S_H^V(z) / g_H^V \rangle_M} \quad (7)$$

To determine $K_0$, the lidar signals are averaged over a maximum of ±30 minutes around the radiosonde launch time, which is compatible with the temporal resolution of the outputs of the global numerical weather prediction models. To improve the SNR, the vertical resolution is set to 50 m to maintain a fine sampling in the atmospheric column. This results in a calibration to an average profile of $M$ individual profiles. To define the altitude range where the calibration is performed, the $SNR_{S_H/S_N}$ of the ratio $S_H/S_N$ is calculated as:

$$SNR_{S_H/S_N}(z) = \left( \frac{\text{var}(S_H(z))}{M \cdot \langle S_H(z) \rangle^2} + \frac{\text{var}(S_N(z))}{M \cdot \langle S_N(z) \rangle^2} \right)^{-1/2} \qquad (8)$$

Values for which $SNR_{S_H/S_N}$ is below 10 are rejected from the subsequent calculations. This operation will exclude noisy or cloudy parts of the profiles. Finally, $K_0$ is estimated on the $N_p$ remaining statistically independent data points $i$ along the selected altitude, as:

$$K_0 = \frac{1}{N_p} \sum_{i=1}^{N_p} \rho_i \text{ with } \rho_i = \frac{r_{ref}(z_i)}{O_R(z_i)} \cdot \frac{\langle S_N^V(z_i) / g_N^V \rangle_M}{\langle S_H^V(z_i) / g_H^V \rangle_M} \qquad (9)$$

Note that $i$ can be taken at different altitudes and/or at different times depending on the meteorological reference data. Considering that the error sources are independent, the total error on $K_0$ noted $\sigma_{K_0}$ is given by the variance law:

$$\sigma_{K_0} = \left( \frac{1}{N_p^2} \sum_{i=1}^{N_p} \frac{\rho_i^2}{SNR_{S_H/S_N i}^2} + \frac{1}{N_p^2} \sum_{i=1}^{N_p} \rho_i^2 \sigma_{O_{R_i}}^2 + \frac{1}{N_p^2} \sum_{i=1}^{N_p} \frac{\rho_i^2 \sigma_{ref,i}^2}{r_{ref_i}^2} \right)^{1/2} \qquad (10)$$

The first term mainly represents the shot noise contribution, and the second term is due to the error in the estimated overlap ratio (standard deviation $\sigma_{O_{R_i}}$). The last term is related to the reference used for calibration (standard deviation $\sigma_{ref}$).

### 2.3.2 Overlap factors

Biases due to the different optical paths between the $H_2O$- and $N_2$-channels, including the geometric overlap, chromaticity, filter angular acceptance, sensitivity variations over the photomultiplier surfaces, should be evaluated. They are all included in the overlap ratio $OR = O_N/O_H$. However, these distortions remain approximately stable in time and only affect the lower part of the lidar profile. Assuming that $O_N$ and $O_H$ tend to 1 after a distance dependent on the lidar system, they can be evaluated using the basic lidar equation applied to the horizontal line of sight (Chazette et al., 2020). The overlap factors for WALI are shown in Figure 1.
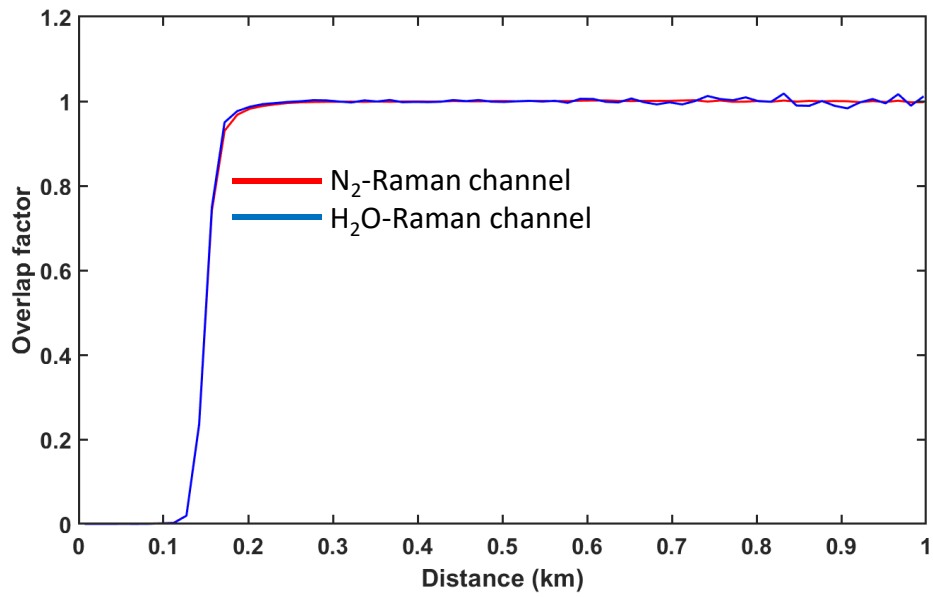
Figure 1. Overlap factor as a function of the distance for both the $N_2$- ($O_N$) and $H_2O$-Raman ($O_H$) channels of WALI.

### 2.3.3 Variation of photomultiplier tube internal gain

For lidars that are limited by the saturation of the photomultipliers due to daytime sky radiance, it is necessary to optimise the SNR by maximising the night-time gain and decreasing the daytime gain. In that case, a separate calibration of the PMT gain functions should be performed. The aim is to achieve a residual uncertainty $\varepsilon_{HV}$ of ~1% on the channel gain ratios. The photomultiplier gain for each channel $k$ typically varies following an exponential law of the high voltage $V$ as

$$g_k(V) = exp\big(\mathcal{P}_k(V)\big) \tag{11}$$

where $\mathcal{P}_k(V)$ can be approximated by a 2nd order polynomial (Chazette et al., 2014b; Totems et al., 2021):

$$\mathcal{P}_k(V) = p_{k,0} + p_{k,1}V + p_{k,2}V^2 \tag{12}$$

The root of the polynomial $\mathcal{P}_k(V)$ is the reference voltage $V_0$ at which the calibration was performed, so that the gain value at this voltage is 1. Note that the 2nd order term accounts for ~20% of the change in the gain value compared to a straight exponential, whereas an additional 3rd order term would only account for ~1%, which is very small. The $p_{k,i}$ coefficients are fairly stable over the duration of a campaign, but can vary slightly over the multi-year lifetime of a PMT detector. If both control voltages on the $N_2$- and $H_2O$-channels change simultaneously, the channel gain ratio $\mathcal{G}$ follows the same type of law:

$$\mathcal{G}(V) = \frac{g_N}{g_H}(V) = exp(a + bV + cV^2) \tag{13}$$

with the coefficients $a$, $b$, $c$ that are the difference between those of the $N_2$ and $H_2O$ photomultipliers and are therefore much smaller. The gain ratio is also easier to calibrate as the ratio is insensitive to variations in the laser energy and the atmospheric aerosol load. We have found empirically that $g_N^V/g_H^V$ varies by 12 to 16% around unity and needs to be corrected.

### 2.3.4 Error calculation

The uncertainties affecting the lidar-derived WVMR are assumed to be independent and the standard deviation $\sigma_l$ on $r_l$ is given by:

$$\sigma_l \approx \sqrt{\underbrace{\frac{1}{SNR_{S_H/S_{N\,i}}^2}}_{\text{Measurement}} + \underbrace{\sigma_{K_0}{}^2 + \sigma_{O_R}{}^2 + \varepsilon_{HV}^2}_{\text{Calibration}} + \underbrace{\varepsilon_m^2 + \varepsilon_a^2}_{\text{Atmosphere}}} \tag{14}$$

It can be divided in three parts: the measurement error mainly due to shot noise, the errors associated with the calibrations and the residual error after correction for atmospheric transmission ($\varepsilon_m$ and $\varepsilon_a$ represent the bias due to the molecular and aerosol transmissions, respectively). The last block is generally negligible compared to the other two. In the calibration block, it is mainly the error on $K_0$ that contributes. This error is not solely statistical, as it also depends on the biases in the measurements used as a reference. The latter, together with the shot noise, is the major contributor to the total error on the WVMR.

## 2.4 Statistical parameters used for the calibration

In this paper we examine the temporal evolution of the calibration coefficient via the statistical deviations between the lidar-derived WVMR profiles and the reference profiles. This will give us a more representative assessment of $\sigma_{K_0}$ and its stability from one campaign to the next.

Three independent statistical indicators are calculated against altitude to assess the consistency between the lidar-derived WVMR profiles and the reference profile data sets. They are the centred root mean square error (RMSE), the mean bias (MB) and the Pearson correlation coefficient (COR). These parameters are often used to evaluate model performance (Tombette et al., 2008; Boylan and Russell, 2006; Kim et al., 2013). Their mathematical expressions are given for each altitude $z$ by:

$$RMSE(z) = \sqrt{\langle\left(\left(r_l(\text{t}, z) - \langle r_l(\text{t}, z)\rangle\right) - \left(r_x(\text{t}, z) - \langle r_x(\text{t}, z)\rangle\right)\right)^2\rangle} \tag{15}$$

$$COR(z) = \frac{\langle (r_l(t,z) - \langle r_l(t,z) \rangle) \cdot (r_x(t,z) - \langle r_x(t,z) \rangle) \rangle}{\sqrt{\langle (r_l(t,z) - \langle r_l(t,z) \rangle)^2 \rangle \cdot \langle (r_x(t,z) - \langle r_x(t,z) \rangle)^2 \rangle}} \tag{16}$$

$$MB(z) = \langle r_l(t,z) \rangle - \langle r_x(t,z) \rangle \tag{17}$$

where $\langle \ \rangle$ is the temporal averaging. WVMR reference profiles using instruments other than lidar are indicated by the subscript $x$. In the following, WVMR profiles from all data sets are re-gridded on a 50 m vertical sampling grid.

Moreover, since MB and RMSE are related to the root mean square difference (RMSD) by the simple relationship:

$$RMSD^2 = RMSE^2 + MB^2 \tag{18}$$

Minimising the RMSD will give the optimum calibration coefficient $K_0$. In order to properly evaluate the calibration error over the altitude range $[z_a, z_b]$ of the lidar profiles, we will also consider the altitude averaged MB ($\widetilde{MB}$) and RMSE ($\widetilde{RMSE}$), defined by

$$\widetilde{MB} = \int_{z_a}^{z_b} MB(z) \cdot dz$$
$$\widetilde{RMSE} = \sqrt{\int_{z_a}^{z_b} RMSE(z)^2 \cdot dz} \tag{19}$$

The cross-correlation coefficient $C$ is also used to assess the respective contribution of the instrument specific statistical noise and the natural variability of the atmosphere:

$$C(\tau, z) = \langle (r_l(t,z) - \langle r_l(t,z) \rangle) \cdot (r_x(t+\tau,z) - \langle r_x(t+\tau,z) \rangle) \rangle \tag{20}$$

where $\tau$ is the time lag.

## 3    Measurements and field campaigns

The reference datasets are derived from field campaigns carried out between 2016 and 2023. They are used to monitor changes in the calibration of the WALI lidar, taking into account that post-telescope modifications to the instrument are minor. In this case, these changes are related to the use or non-use of a seeding laser and to the use of calibrated optical densities on the detection channels. In addition to the WALI lidar, the field campaigns included either airborne measurements or radiosondes, and in some cases both.

### 3.1    In situ meteorological probes

A Tanarg 912 XS ultralight aircraft built by Air Création (https://www.aircreation.com/, last access: 19 June 2024) was used for *in situ* measurements during three field campaigns. With an instrument payload of up to 120 kg, flight times can reach ~2–3 hours depending on the air conditions, with a cruise speed of around 85-90 km.h$^{-1}$. The aircraft climb rate is of the order of 5 ms$^{-1}$. Part of the aircraft payload was a VAISALA PTU-300 meteorological probe for temperature, pressure and relative humidity measurements (https://docs.vaisala.com/v/u/B210954EN-J/en-US, last access: 21 June 2024). This probe was used to perform the lidar calibration as close as possible to the laser beam. It measures the atmospheric pressure, averaged over a 1-minute sampling time, with an absolute uncertainty of 0.45 hPa, the thermodynamic temperature with an uncertainty of 0.2 °C, and the relative humidity with an uncertainty of 2.5%. The final absolute error on the WVMR is thus ~0.30 g kg$^{-1}$ in the planetary boundary layer (PBL) by considering the equation that relates WVMR to relative humidity *RH* and atmospheric pressure *P*:

$$r_x = 0.622 \cdot \frac{RH_x \cdot e_w}{P_x - RH_x \cdot e_w} \tag{21}$$

where $e_w$ is the partial pressure of water vapour at saturation given by the relationship proposed by Buck (1981), which is related to the air temperature *T*:

$$e_w = 6.1121 \cdot exp\left[\left(18.678 - \frac{T_x}{234.5}\right) \cdot \left(\frac{T_x}{(257.14 + T_x)}\right)\right] \tag{22}$$

### 3.2   Radiosondes

Radiosondes are often considered the standard for meteorological profiles, but they cannot be launched frequently and are subject to greater errors than calibrated meteorological probes on the ground. These systematic errors are not necessarily accounted for in the manufacturers' data sheets. Inconsistent humidity biases have been reported when using probes at different locations and times of day (Bock and Nuret, 2009; Serreze et al., 2012). Inaccurate humidity measurements and different reporting methods at low temperatures and humidity are also reported, as well as inconsistencies between the probes themselves. Apart from the manufacturers' data sheets, we have little information on which to base error bars for radiosonde measurements of the WVMR. The two main types of radiosondes used during our field experiments were the MODEM M10 (http://leaflet.meteomodem.com/M10%20EN.pdf, last access: 24 October 2024) and the VAISALA RS41-SG (https://docs.vaisala.com/v/u/B211321EN-K/en-US, last access: 24 October 2024). According to both datasheets, the systematic error in radiosonde measurements is far from negligible:

commonly accepted values are ~1 hPa for pressure, 0.3°C for temperature, and 3-4% for RH, which together amount to $\sigma_{ref}$ ~ 0.4–0.5 g.kg$^{-1}$. Note that profiles measured by operational radiosondes (Ingleby et al., 2016) show even larger discrepancies when compared to model reanalyses, e.g. up to 1°C and 15% for the temperature and RH, respectively, in the troposphere. This leads to an uncertainty of ~1.8 g kg$^{-1}$ in the WVMR when considering Eq. 21 and 22. All these error estimates underline the importance of being able to perform a calibration using several vertical reference profiles, preferably obtained by different measurement methods, and at least several atmospheric radiosonde samples.

### 3.3    Field campaigns

Six field campaigns have been carried out with WALI between 2016 and 2023. They provide the opportunity to compare lidar profiles with reference profiles obtained from radiosonde or airborne measurements. The campaigns and lidar datasets used for this calibration exercise are described below in chronological order. These campaigns were carried out for very contrasting water vapour contents and are therefore highly complementary for lidar calibration. The temporal evolution of the WVMR profiles is given by the WALI measurements, but also by the operational output of the global numerical weather prediction models ECMWF/IFS (European Centre for Medium-range Weather Forecasts/integrated forecast system), here the ECMWF Reanalysis v5 (ERA5) dataset (Hersbach et al., 2020, 2023). Although the reanalyses are mesoscale model outputs, they are constrained by various types of measurements, including radiosondes and spaceborne observations such as the Infrared Atmospheric Sounding Interferometer (IASI) (Collard and McNally, 2009). They are therefore relevant references to consider, although they tend to deviate from observations in the lower layers of the atmosphere (Chazette et al., 2014a; Totems et al., 2019). It should be noted that in the figures showing the temporal evolution of the vertical profiles of the WVMR derived from ERA5, it was chosen to show only the height ranges actually accessible to lidar measurements. This choice makes it easier to highlight the agreements and divergences that will affect the statistical parameters.

### 3.3.1    Pollution in the ARCtic System (PARCS) (May 2016)

The PARCS field campaign took place from 13 to 26 May 2016 in the region of Hammerfest (70°40′N 23°41′E, altitude 90 m a.m.s.l., Norway), 90 km southwest of the North Cape, within the Arctic Circle. It included the ground-based WALI lidar and meteorological measurements taken with two PTU–300 meteorological probes mounted on an ultralight aircraft and a mast at

5 m a.g.l., respectively. The WALI laser was not injection seeded and the N$_2$–Raman channel was equipped with an optical density of 0.43.

The low air temperatures during this field campaign are associated with low WVMR values as shown in Fig. 2 from both lidar and ERA5. The PBL is located at an altitude of approximately 800 m (Chazette et al., 2018) and significant values of WVMR are only be detected in the lower Arctic troposphere. The WVMR generally remained below 4 g kg$^{-1}$ in the PBL except at the end of the period when it reached values of ~7 g kg$^{-1}$ with the arrival of heavy cloud cover, signalling the end of the cloudless sky. The values recorded in the free troposphere were less than 2 g kg$^{-1}$. We conducted a total of 5 flights around the lidar during the field campaign (Chazette et al., 2018) of which only 2 were above 1.4 km above mean sea level (a.m.s.l.). The flights were conducted near the airport, around the Hammerfest peninsula. Each flight involved a slow spiral ascent and are localised in time in Figure 2a.
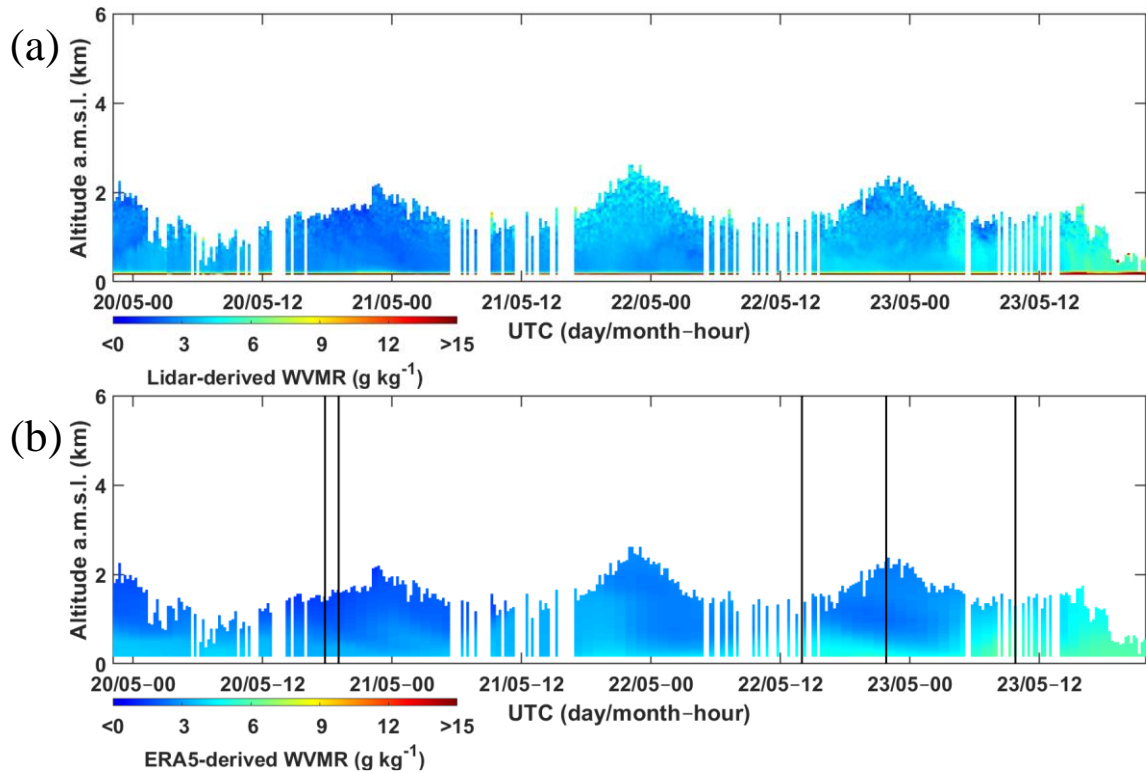
Figure 2. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the PARCS field campaign. Average flight times are indicated by vertical black lines in b).

### 3.3.2   ADM–AEOLUS (April 2019)

The calibration and validation field campaign for the ADM-AEOLUS space mission took place from April to June 2019. During this period, the WALI lidar operated from 17 to 23 June 2019 as a reference for other instruments. It was located at the CEA Orme des Merisiers site (48°42' N 2°9' W, altitude 168 m a.m.s.l.), ~14 km from the Meteo-France's Trappes radiosonde launch

site (48°46' N 2°1' W). The laser was equipped with an injection seeder because temperature measurements were carried out in parallel with the water vapour measurements. It should be note that the seeder modifies the centre wavelength and shape of the laser line and can lead to variations in the effective cross section of the Raman scattering of water molecules when viewed through a narrow spectral filter. The calibration constant may be affected.

Figure 3a shows the temporal evolution of the WMVR as measured by WALI. The values are typical of the humidity conditions found in April in the Paris region. The WVMRs derived from lidar and ERA5 were of the order of 7 g kg$^{-1}$ in the PBL, with values of between 1 and 4 g kg$^{-1}$ in the free troposphere, depending on the air masses advected over the site. On certain nights, dry air masses supported by anticyclonic conditions can be observed. During the lidar measurements, 13 radiosondes were deployed from the Meteo–France station at Trappes, located ~14 km to the northwest of the lidar. The launch times of the radiosondes are also shown in Figure 3a.
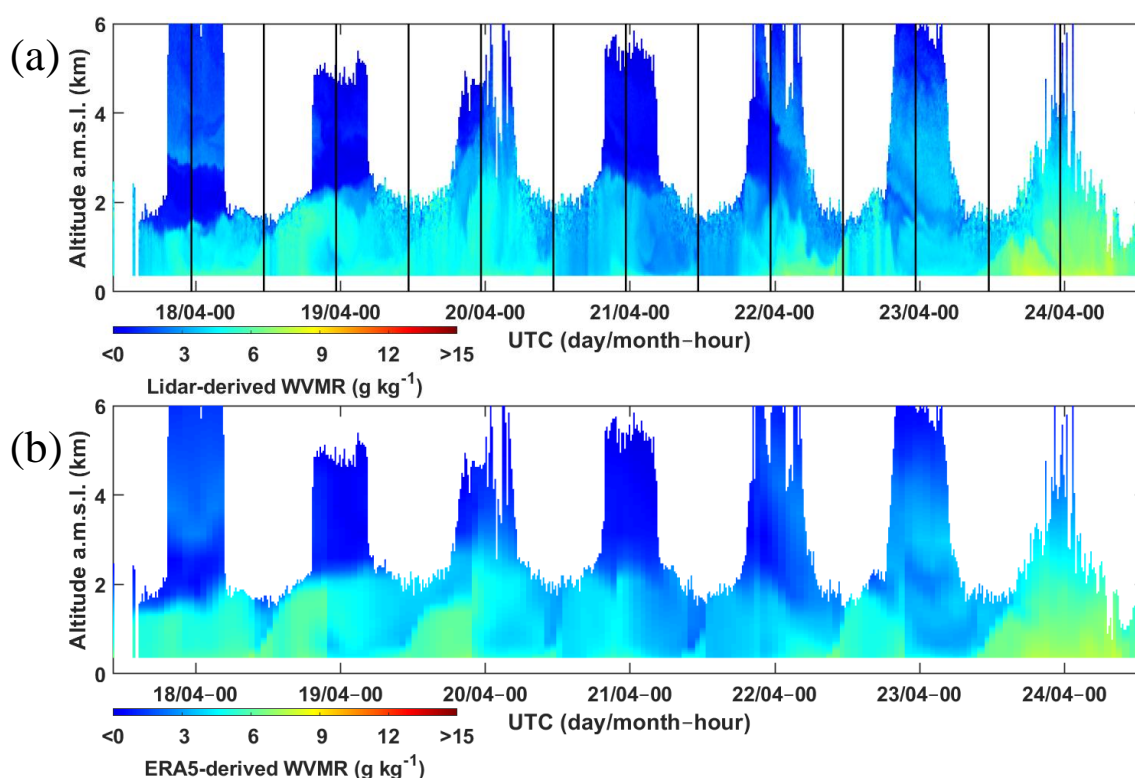
Figure 3. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the ADM–AEOLUS field campaign. Average radiosonde times are indicated by vertical black lines in a).

### 3.3.3 L-WAIVE (June 2019)

The L-WAIVE (Lacustrine-Water vApor Isotope inVentory Experiment) field campaign was carried out in the Annecy valley (45°47' N, 6°12' E, altitude 458 m a.m.s.l., in Haute-Savoie in the French Alps) around Annecy Lake between 12 and 23 June 2019 (Chazette et al., 2021). In order to sample the temporal evolution of the lower troposphere, the ultralight aircraft

performed 19 flights between 13 and 19 June, mainly over Annecy Lake. The aircraft was equipped with a PTU–300 meteorological probe, which carried out in situ sampling in conjunction with the WALI measurements. The laser injection seeder malfunctioned with a bimodal emission. The temporal evolution of the lidar and ERA5 derived WVMR profiles are shown in Fig. 4a and 4b, respectively. The flight time positions are also shown in Fig. 4a. The WVMRs were much higher than in previous cases, with most values above 9 g kg$^{-1}$ in the lower troposphere. Water vapour levels were also high in the free troposphere, exceeding 6 g kg$^{-1}$.
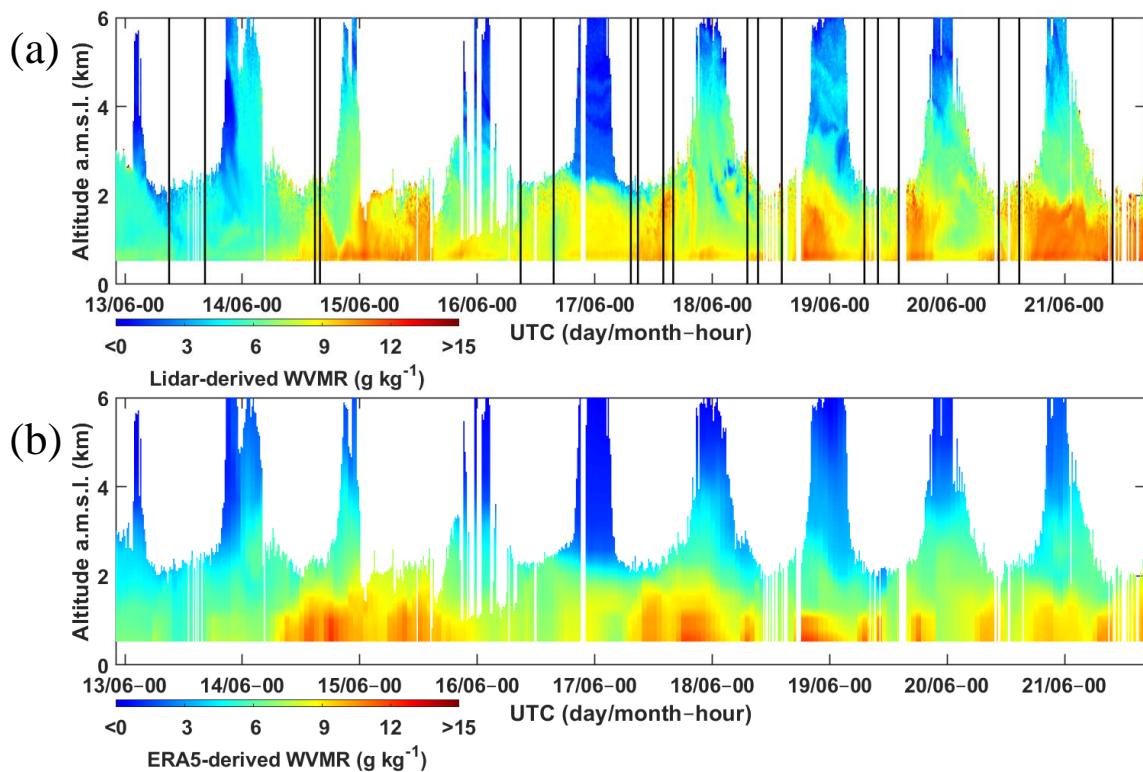
Figure 4. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the L-WAIVE field campaign. Average radiosonde times are indicated by vertical black lines in a).

### 3.3.4  Post COVID (April-August 2020)

Following the COVID crisis (P-COVID), WALI was used to monitor changes in the lower and middle troposphere with the resumption of economic activity between April and August 2020. The lidar was reinstalled at the Orme des Merisiers site. As for the L-WAIVE field campaign, the laser was seeded but centred at a different wavelength due to a seeder malfunction. Throughout the period, 63 nights of radiosondes from the Trappes meteorological site could be used in conjunction with the lidar measurements. This allowed to obtain a wide range of water vapour contents, from about 2 to 12 g kg$^{-1}$ in the low troposphere. The temporal evolution of the WVMR profiles derived from WALI is shown in Fig. 5a while the corresponding ERA5 profiles are shown in Fig.5b. The data gaps correspond to cloudy and/or rainy periods and are obviously not taken into account in the number of radiosondes used.
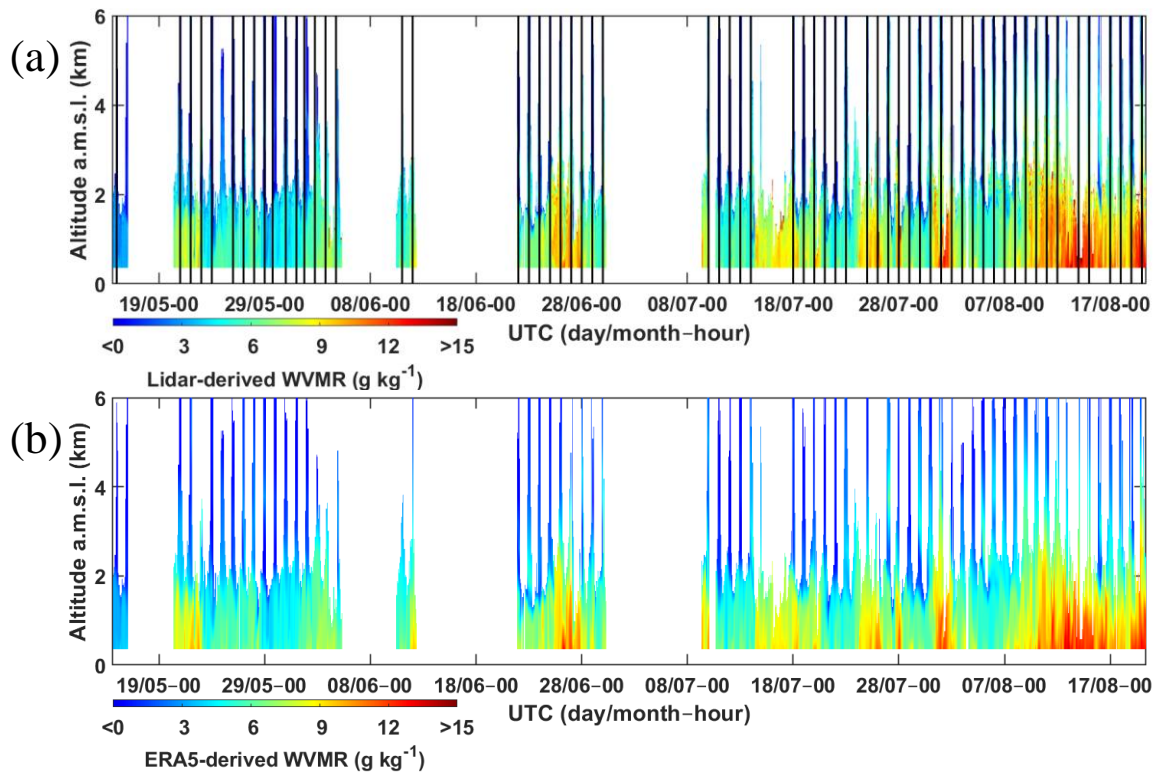
Figure 5. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the Post–COVID field campaign. Average radiosonde times are indicated by vertical black lines in a).

### 3.3.5    LEMON (September 2021)

The field campaign for the European Lidar Emitter and Multispecies greenhouse gases Observation iNstrument project (LEMON, https://lemon-dial-project.eu/, last access: 28 July 2024) provided an opportunity to carry out joint WALI, radiosonde and airborne measurements. As in previous campaigns, we used a PTU-300 meteorological probe from an ultralight aircraft. During the day, 20 radiosondes and 12 flights were carried out to coincide with the lidar measurements. The laser was properly seeded. The measurements were carried out between 16 and 23 September 2021 over the airfield of Aubenas (44°32'N 4°22'E, altitude 281 m a.m.s.l.) with the help of the company Air Creation (https://www.aircreation.com/en, last access: 28 July 2024). The weather conditions were very variable, with the presence of thunderstorms. The lidar was therefore able to sample large variations in the WVMR, ranging from 3 to 14 g kg$^{-1}$ as shown in Fig. 6a. The beginning of the field campaign was a transition between two weather regimes. Humidity dropped sharply with easterly winds over the whole of France. The launch times of the radiosondes are shown in Fig. 6a with the lidar-derived WVMR and the mean flight time in Fig. 6b with the corresponding ERA5 data.
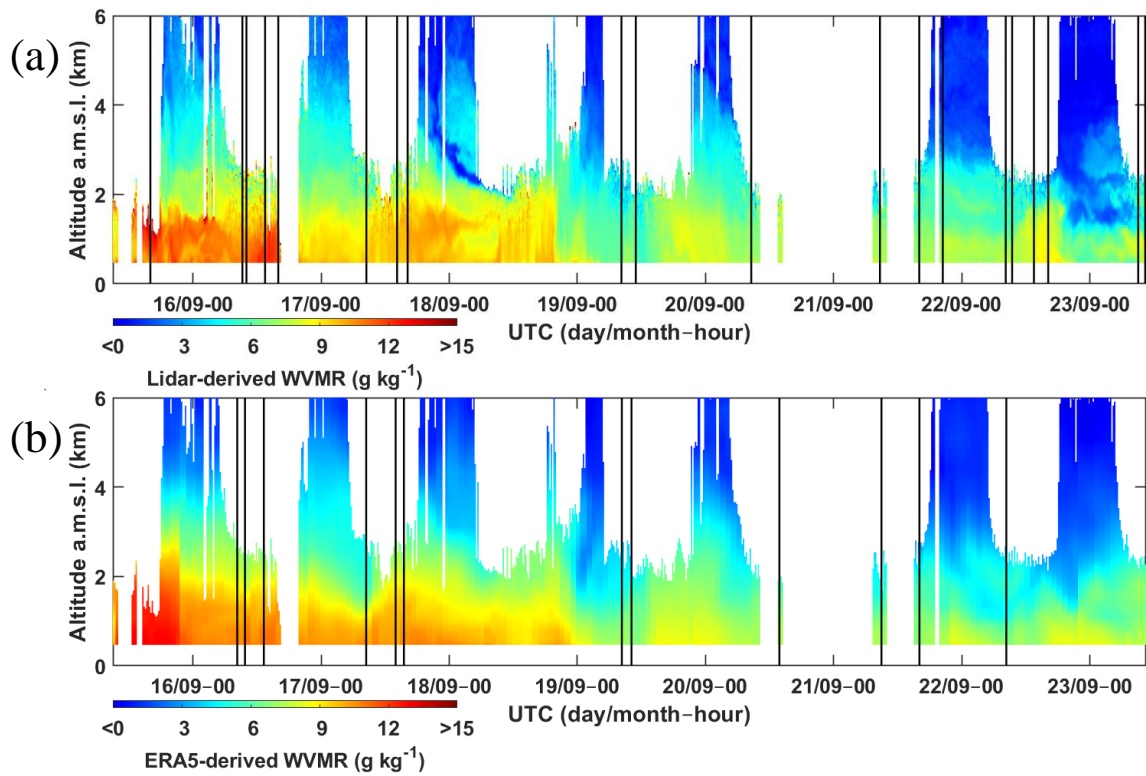
Figure 6. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the LEMON field campaign. Average radiosonde times are indicated by vertical black lines in a) and average flight times are indicated by vertical black lines in b).

### 3.3.6  WaLiNeAs (November 2022-january 2023)

The Water Vapour Lidar Network Assimilation (WaLiNeAs) project was aimed to improve the prediction of extreme precipitation events along the Mediterranean coast. This improvement involves the assimilation of WVMR lidar profiles derived from lidar measurements into mesoscale models (Flamant et al., 2021), as has been done for aerosols (Wang et al., 2014). One of the observation stations set up was located at the Thales Alenia Space site in Cannes (43°32'N 6°57'E, altitude 4 m a.m.s.l.) at the foothills of the Maritime Alps. WALI acquired data continuously without laser injection seeder between 17 November 2022 and 12 January 2023. A PTU 300 probe mounted on a mast at 5 m a.g.l. was used to calibrate the lidar according to a procedure described in Laly et al. (2024). In this procedure, only profiles with a zero-moisture gradient below 400 m a.g.l. were considered and the calibration was performed between the in situ weather sensor and the lidar measurement at 200 m a.g.l., where $O_R = 1$. This altitude is associated with a high signal-to-noise ratio (> 100). Figure 7a and 7b show the temporal evolution of the WVMR profiles over the Cannes site according to the lidar and ERA5, respectively. The period was particularly dry for the season, with WVMR between 2 and 7 g kg$^{-1}$.
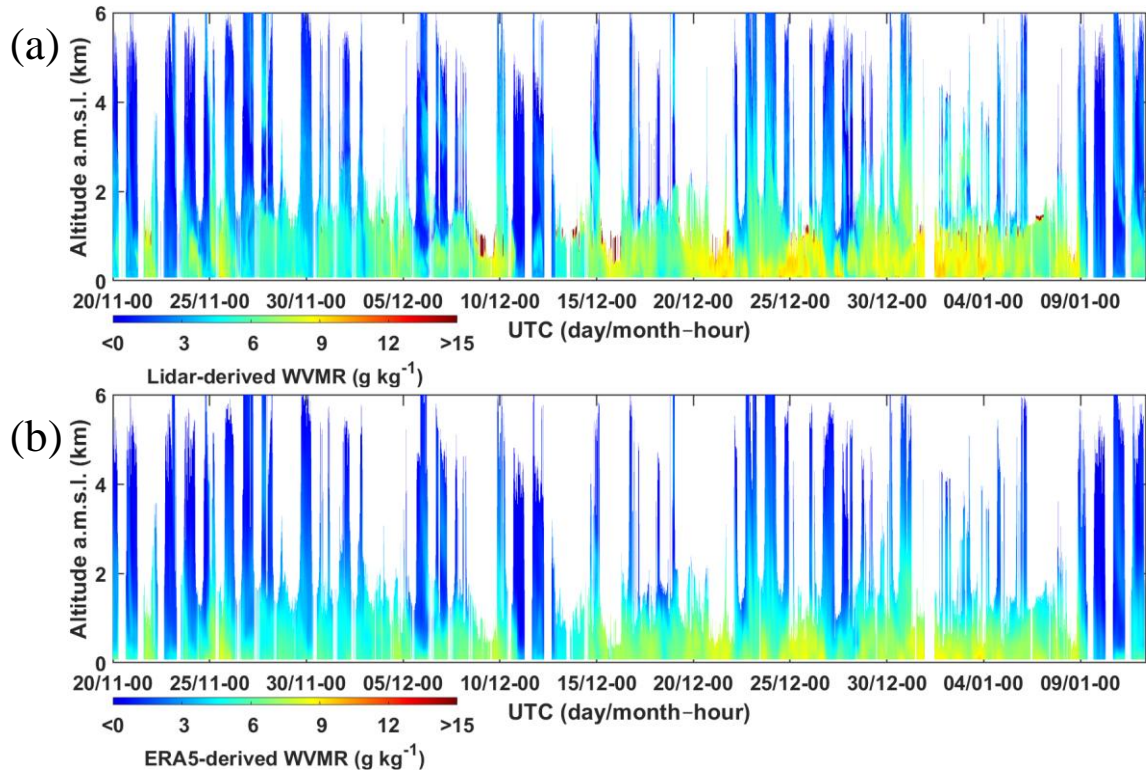
Figure 7. a) Raman lidar- and b) ERA5-derived WVMR vertical profiles during the WaLiNeAs field campaign.

## 4    Statistical studies following calibrations

For each field campaign, the value of $K_0$ is calculated by minimising the RMSD over all the reference WVMR profiles (radiosondes or flights). The effect of the calibration constant on the WVMR profiles is then evaluated using the other statistical variables presented in subsection 2.4. The altitude zones studied depend on both the altitude of the ground-based station and the day/night conditions. Statistical tools are also used to compare the lidar profiles with ERA5 reanalyses at the same altitudes and times as the reference profiles to ensure a relevant comparison.

### 4.1    Calibration during the PARCS field campaign
#### 4.1.1    Calibration from flights

During this field campaign, only a few flights are available for statistical analysis (Figure 2a), and therefore less representative than in subsequent campaigns. This is especially true above 1.3 km a.m.s.l., where the air temperature did not always allow for high altitude excursions, especially below -15°C. Only 2 flights exceeded 1.3 km a.m.s.l.. An example of a vertical profile for 20 May 2016 is shown in Figure 8a. There is good agreement between the lidar and the aircraft measurements. The reanalyses also agree with the measurements. The optimal calibration was performed with $K_0 = 105$ g kg$^{-1}$ after optical density correction. The RMSE remains below 0.2 g kg$^{-1}$ below 1.3 km a.m.s.l. and approaches 0.5 g kg$^{-1}$ above (Figure 8b).

Overall, the $\widehat{RMSE}$ is 0.23 g kg$^{-1}$ and 0.53 g kg$^{-1}$ when comparing the lidar measurements with the airborne measurements and with ERA5, respectively. The absolute MB (Figure 8c) is higher near the surface due to the greater horizontal variability of the lower layers and the seaward departure of the aircraft, resulting in negative values. In the free troposphere, above 0.8 km

5    a.m.s.l. (Chazette et al., 2018), the absolute MB is less than 0.4 g kg$^{-1}$ in the PBL and may be larger above 1.5 km a.m.s.l. due to orographic effects. The value of $\widetilde{MB}$ derived from Fig. 7c is small and equal to 3 10$^{-3}$ g kg$^{-1}$ due to the equilibrium between the lower and upper layers. It should be noted that a small value of ~-0.01 g kg$^{-1}$ is calculated between the lidar and ERA5, which shows a very good average agreement between the reanalyses and the lidar

10   measurements. The correlation coefficient $COR$ between the observations is high below 1.3 km a.m.s.l., generally higher than 0.75 (Figure 8d). Above this height the correlation deteriorates due to the lack of representative flights.
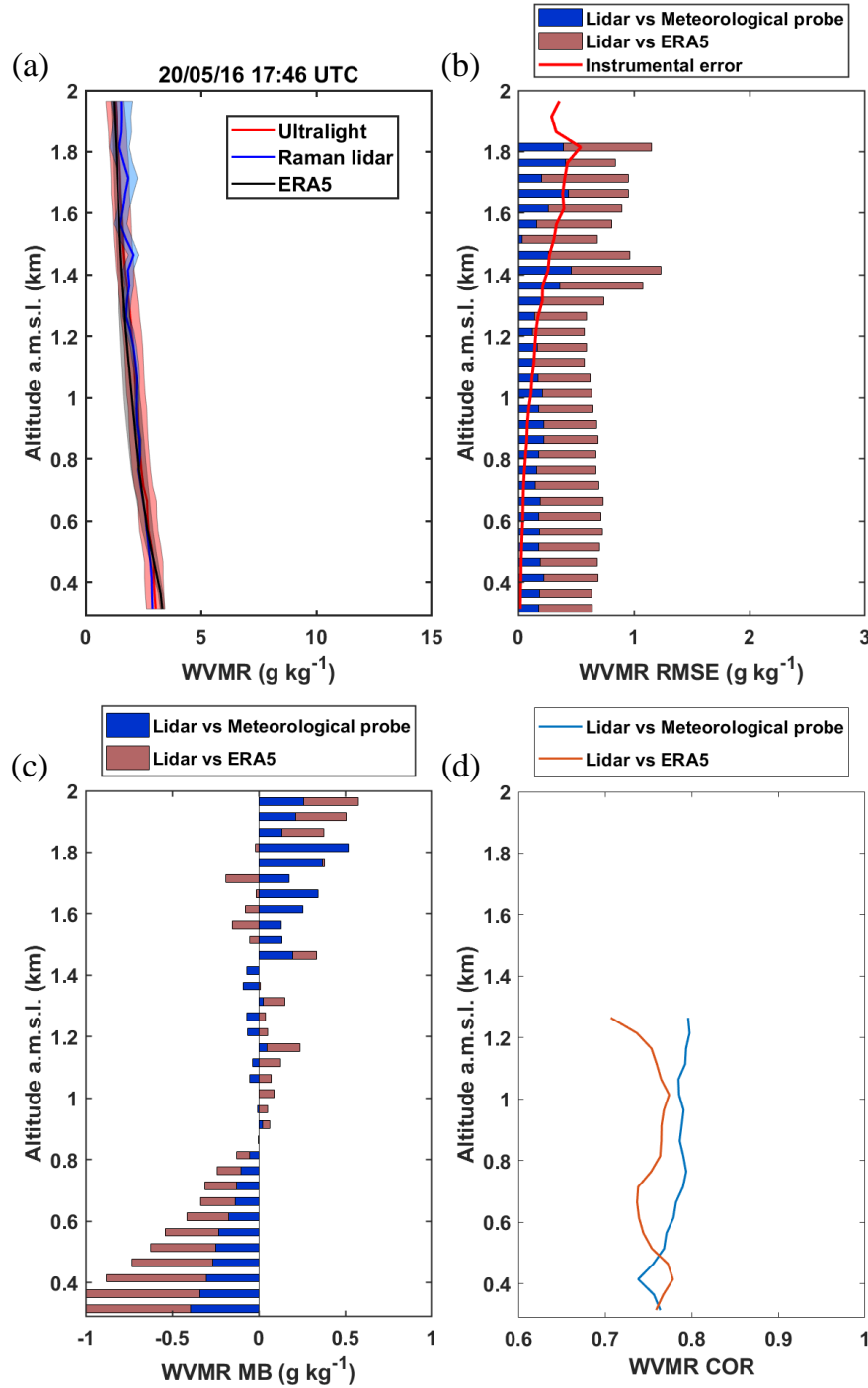
Figure 8. a) Example on 20 May 2016 1746 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the in-situ aircraft measurements during the PARCS field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and d) correlation coefficient (COR) between i) the lidar and the aircraft measurements and ii) the lidar and ERA5.

### 4.1.2 Calibration from the ground-based meteorological probe

The PTU 300 installed at 5 m a.g.l. is also considered as a reference according to the method presented in Laly et al. (2024). In this method, the lidar measurement at 200 m is compared with the meteorological probe measurements at 5 m when the lower atmosphere had a weak

vertical gradient on the WVMR. This corresponds to a slight vertical gradient in the ratio $\frac{S_H^V(z)}{S_N^V(z)}$ and the atmosphere can be considered to be in mixing equilibrium in the layer under consideration. The derived value of $K_0$ is then 108.1 g kg$^{-1}$ very close to the previous value (Figure 9). However, there is an intercept of about 0.27 g kg$^{-1}$, which may be related to the difference in altitude between the ground-based probe and the lowest lidar measurements, and a possible stratification near the ground in the cold conditions encountered. Nevertheless, there seems to be a good consistency between the two calibration approaches with a Pearson correlation coefficient $R^2 = 0.92$.
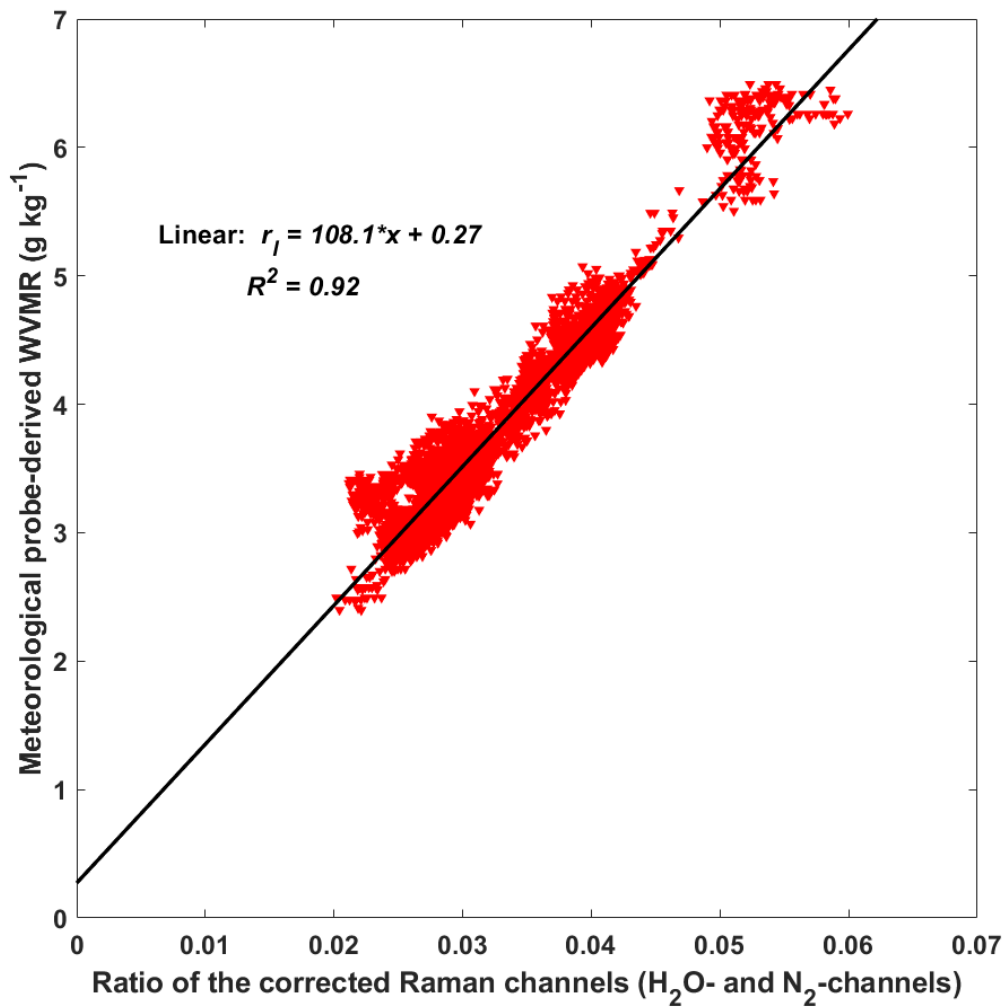
Figure 9. Calibration of the WALI lidar during the PARCS field campaign by comparison of lidar measurements at 200 m a.g.l. and in situ meteorological measurements at 5 m a.g.l. The linear regression is represented by the solid black line and the equation is given in the figure with $r_l$ the WVMR and x the corrected ration of the Raman channels. The Pearson correlation coefficient $R^2$ is also given.

### 4.2 Calibration during the ADM-AEOLUS field campaign

Calibration was carried out using radiosondes (Figure 3a) and the lidar measurements were made using an injector seeder. The optimal calibration was obtained for $K_0 = 117$ g kg$^{-1}$. An example of the vertical profile after calibration is shown in Figure 10a. In the lower atmospheric layers, there is a difference of about 1 g kg$^{-1}$ between the lidar and radiosonde measurements. This difference is not present in the ERA5 reanalyses. As the radiosonde station was not located at the lidar site, differences are expected mainly in the PBL. The RMSE is high (Figure 10b), with $\widetilde{RMSE}$ of 0.55 g kg$^{-1}$. These values include the natural variability during calibration, which may explain its amplitude. Furthermore, there is a significant difference between the RMSE and the estimate of instrumental noise from the lidar measurements alone (solid red line in Figure 10b). This difference shows that the natural variability of water vapour during calibration is a major error source. The increase in instrumental error around 2 km altitude is related to the contribution of daytime data, which is very noisy above this altitude and is therefore not taken into account. The vertically averaged MB derived from Figure 10c is low, less than 0.03 g kg$^{-1}$ when comparing lidar retrievals and radiosondes or ERA5. The correlation coefficient remains high throughout the air column, with values above 0.75 (Figure 10d).
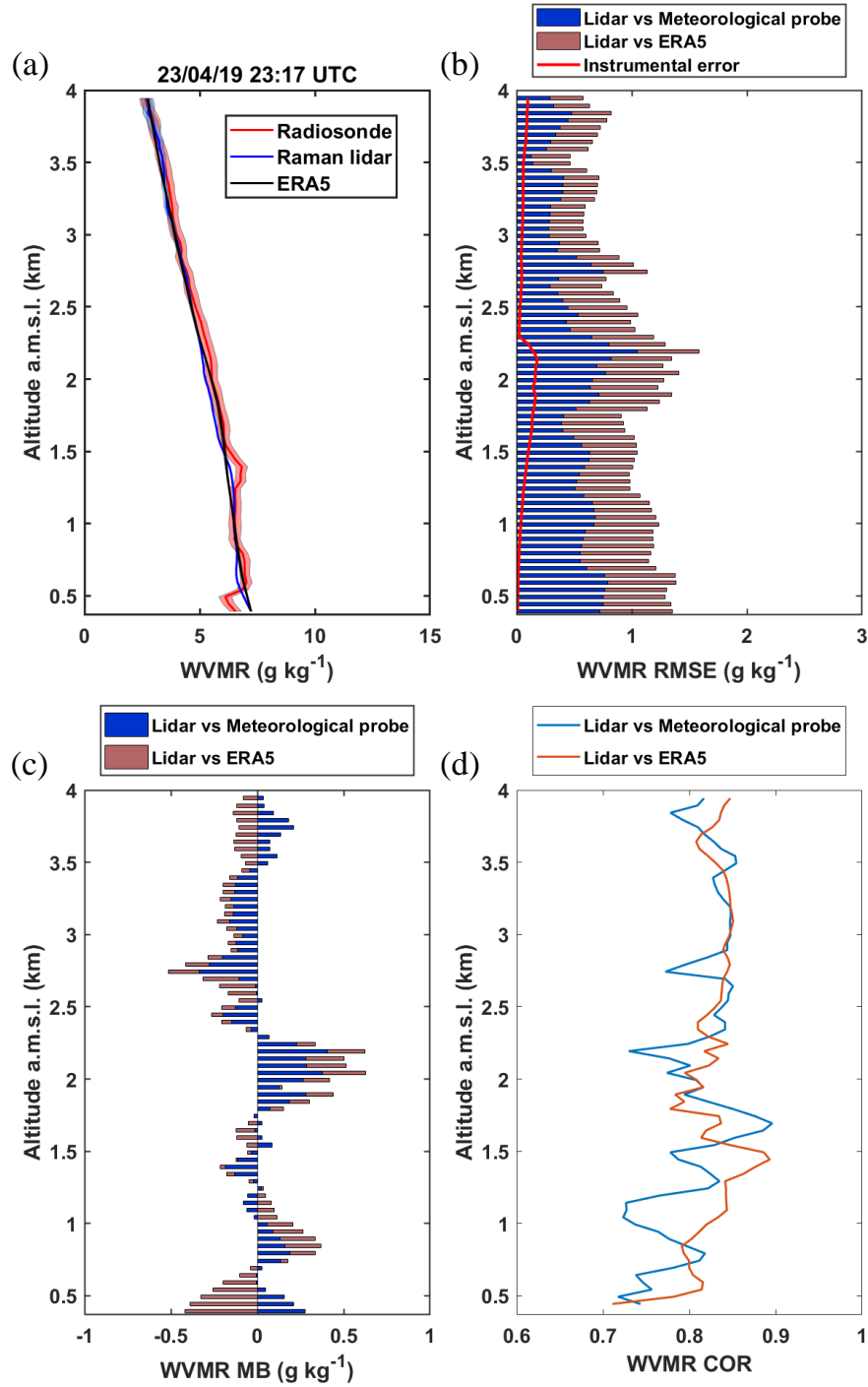
Figure 10. a) Example on 19 April 2023 2317 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the radiosonde measurements during the ADM-AEOLUSS field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and correlation coefficient (COR) between i) the lidar and the radiosonde measurements and ii) the lidar and ERA5.

## 4.3 Calibration during the L-WAIVE field campaign

In contrast to the previous campaign, only flights with meteorological probes were used for calibration (Figure 4a). This campaign was performed two months later the ADM-AEOLUS

field campaign. The value of $K_0$ was found to be 108 g kg$^{-1}$, close to the one retrieved following the PARCS field campaign. Good agreement is shown between the lidar measurements and those from the on-board meteorological sensor (Figure 11a). A higher scatter is observed in the ERA5 data, which is mainly related to the orographic complexity around the lidar site. The

5    RMSE ($\widetilde{RMSE} = 0.60$ g kg$^{-1}$) are similar to those of the ADM-AEOLUS field campaign (Figure 11b), but higher for the comparison to ERA5 ($\widetilde{RMSE} = 0.73$ g kg$^{-1}$) also due to the orography. The instrumental noise is significantly lower, showing again the importance of the natural variability of the WVMR profiles. As expected, the MB between lidar and meteorological measurements remains low (Figure 11c). It reaches values of $\widetilde{MB} = -0.08$ g kg$^{-1}$ in the

10   comparisons between lidar and radiosondes. It compares more favourably with ERA5 ($\widetilde{MB} = -0.03$ g kg$^{-1}$) because there is a compensation between the lower and upper parts of the profile. The correlation coefficient (Figure 11d) remains high, with values above 0.75.
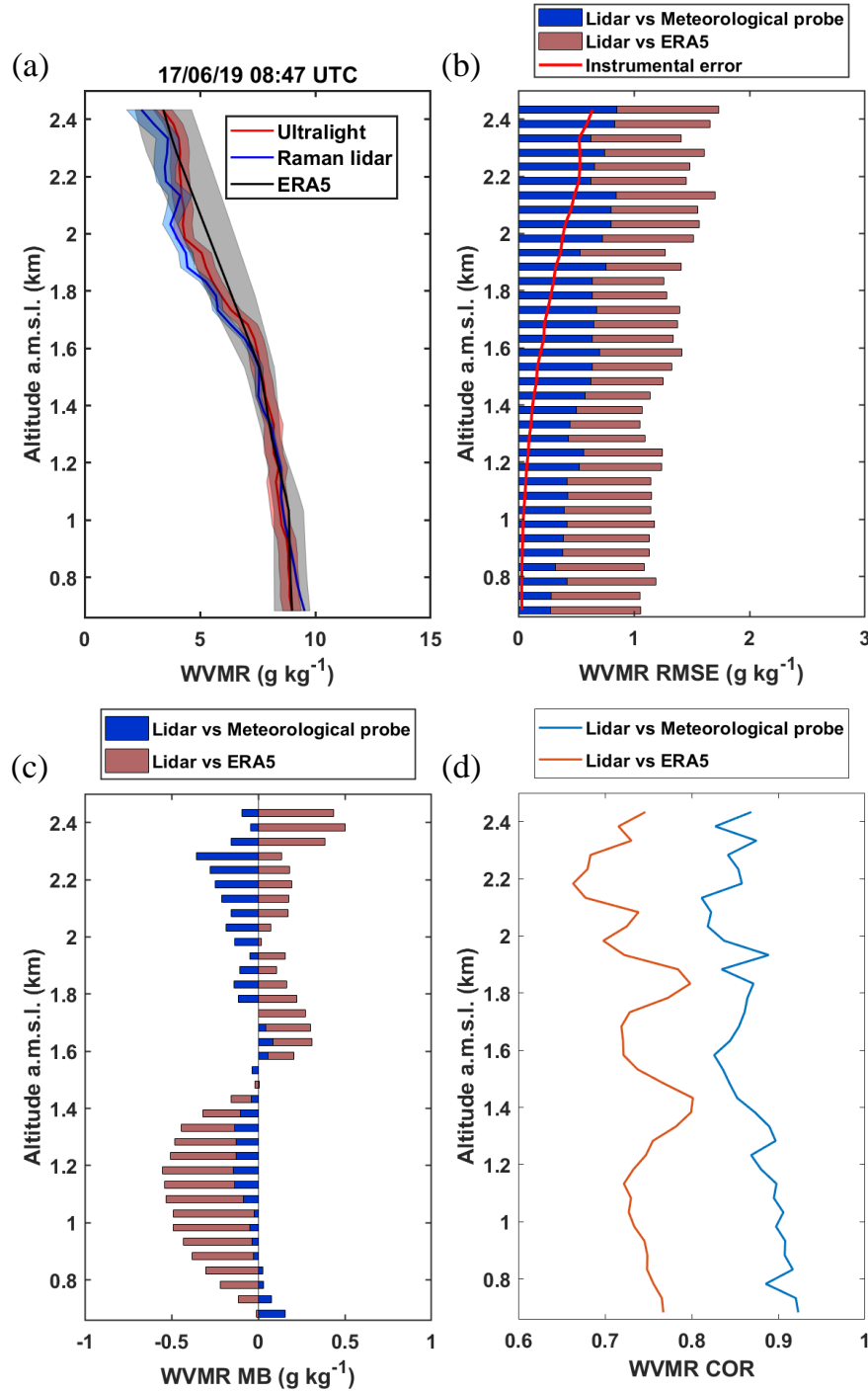
Figure 11. a) Example on 17 June 2019 0847 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the in-situ aircraft measurements during the L-WAIVE field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and correlation (COR) between i) the lidar and the aircraft measurements and ii) the lidar and ERA5.

## 4.4 Calibration during the post COVID field campaign

Lidar measurements were performed at the same site as the ADM-AEOLUS campaign. The measurement period was much longer, allowing calibration with a larger number of radiosondes (Figure 5a). The profiles were selected above the night inversion layer, to limit the low layer

effects associated with the different locations of the lidar and the radiosonde stations. The injection seeder was used, but we found out after the experiment that it was strongly bimodal. Unfortunately, this is a classic malfunction of this type of instrument. Adjusting $K_0$ gives a value of 89 g kg$^{-1}$, which is lower than before. The calibrated lidar profiles fit the radiosonde very well (Figure 12a). However, there are significant discrepancies with ERA5 on this profile ($> 1$ g kg$^{-1}$), which have already been reported for the lower troposphere under certain atmospheric conditions (Chazette et al., 2014a; Totems et al., 2019). The RMSE remains around 0.45 g kg$^{-1}$ (Figure 12b) with $\widehat{RMSE} \sim 0.60$ g kg$^{-1}$, well above the estimated instrumental error (~0.2 g kg$^{-1}$). The MB remains low (Figure 12c) with $\widetilde{MB} = -0.07$ g kg$^{-1}$ and -0.11 g kg$^{-1}$ when comparing lidar with radiosonde and ERA5, respectively. The correlation coefficient is high with $COR > 0.95$ (Figure 12d).
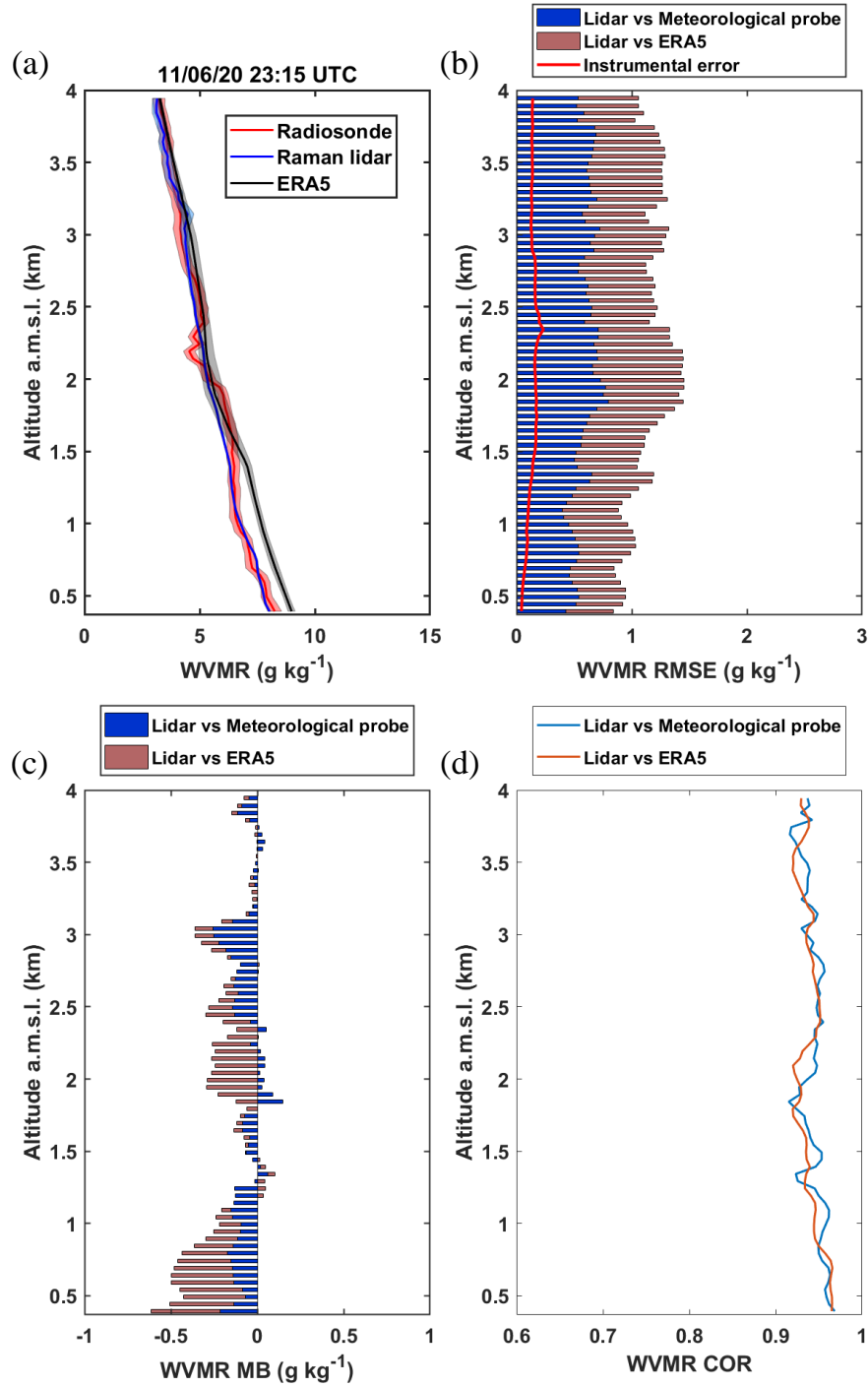
Figure 12. a) Example on 11 June 2020 2315 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the radiosonde measurements during the post COVID field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and correlation coefficient (COR) between i) the lidar and the radiosonde measurements and ii) the lidar and ERA5.

### 4.5 Calibration during LEMON

During the LEMON field campaign, flights and radiosondes were carried out to coincide with the lidar measurements. The laser was injection seeded with a repaired device, without any malfunction.

### 4.5.1  Calibration from radiosondes

Using radiosondes (Figure 6a), the calibration constant was evaluated to be $K_0 = 121.5$ g kg$^{-1}$, which is slightly higher than the value obtained during the ADM-AEOLUS field campaign. The lidar-derived WVMR profile agrees with the radiosonde profile as shown in Figure 13a. The ERA5 profile is also very similar. The RMSE (Figure 13b) takes values lower than 0.40 g kg$^{-1}$, resulting in $\widetilde{RMSE} = 0.48$ g kg$^{-1}$ between lidar and radiosonde measurements. It increases significantly for ERA5, with an equivalent value in the atmospheric column $\widetilde{RMSE} = 0.91$ g kg$^{-1}$, while the instrumental error remains below 0.2 g kg$^{-1}$. The MB remains low between radiosonde and lidar measurements (Figure 13c) with $\widetilde{MB} = 0.03$ g kg$^{-1}$. However, it is significantly higher when compared to ERA5 (0.24 g kg$^{-1}$). Such a value can be explained by the strong instabilities encountered during this field experiment, which remain difficult to describe using mesoscale modelling. Thunderstorms with southerly winds were present during the first two days, gradually shifting to northerly winds before returning to southerly winds at the end of the campaign. The correlation between lidar and radiosonde measurements remains high with values > 0.90 (Figure 13d). It is slightly lower when compared to ERA5 (> 0.70).
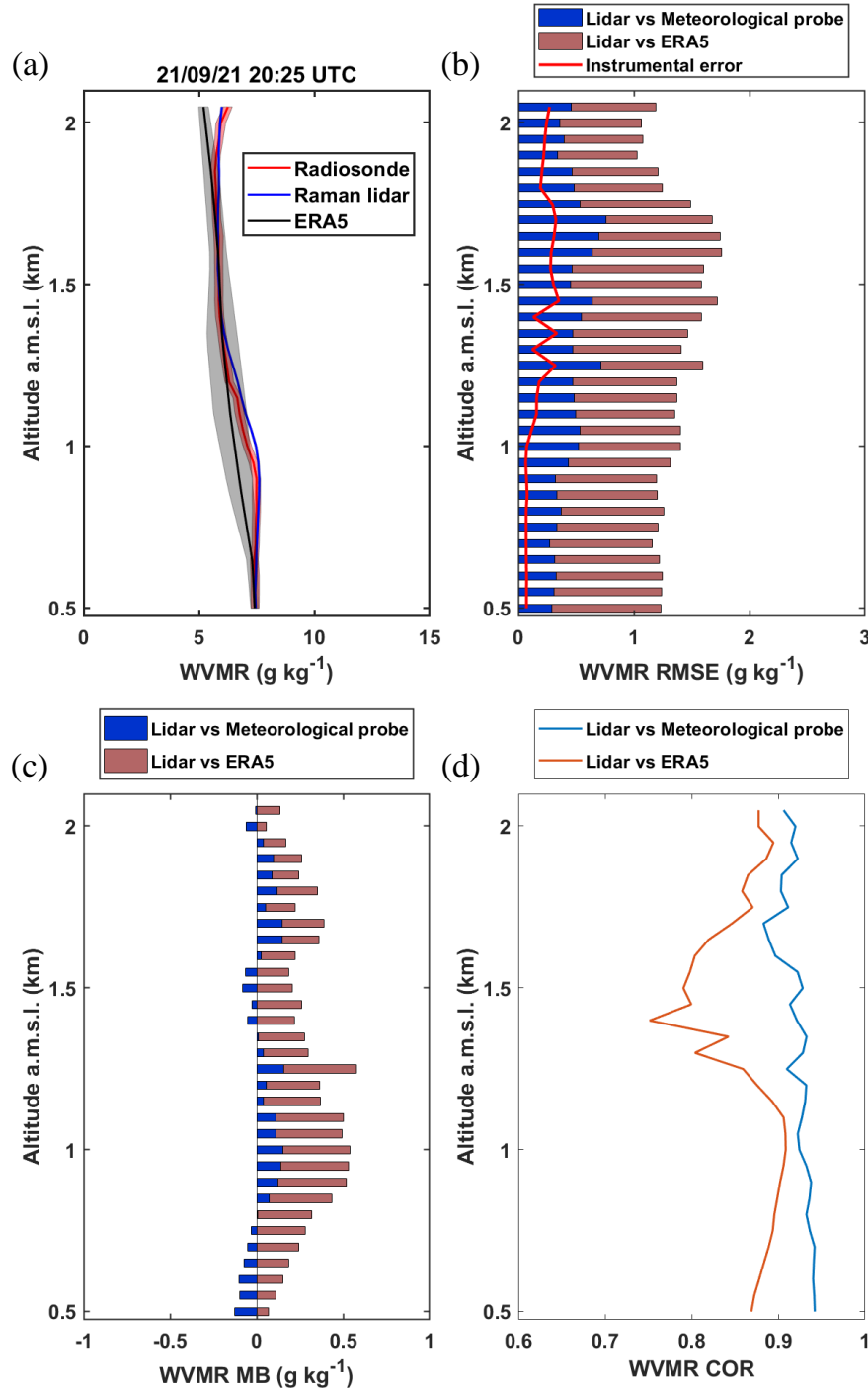
Figure 13. a) Example on 21 September 2021 2025 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the radiosonde measurements during the LEMON field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and correlation coefficient (COR) between i) the lidar and radiosonde measurements and ii) the lidar and ERA5.

### 4.5.2 Calibration from flights

The flights carried out in parallel with the radiosondes (Figure 6b) give $K_0 = 122$ g kg$^{-1}$, a value very close to that obtained with the radiosondes. The advantage of the flights is that they remain close to the vertical of the lidar site, while the radiosondes can drift with the wind, potentially

distorting the reference profiles. The WVMR profiles are in very good agreement as shown in Figure 14a. The RMSE (Figure 14b) is lower than in the radiosonde comparison, with $\widetilde{RMSE}$ around 0.32 g kg$^{-1}$, rising to 0.65 g kg$^{-1}$ when compared to ERA5. The MB (Figure 14c) remains very low between the lidar and the airborne measurements ($\widetilde{MB}$ = -0.02 g kg$^{-1}$). However, it

5 increases significantly when compared to ERA5, with a mean bias value $\widetilde{MB}$ of 0.29 g kg$^{-1}$, probably due to unstable weather that is difficult to represent with the model. The correlation coefficient (Figure 14d) is slightly stronger than with the radiosondes due to the closer proximity of the flights, with values > 0.8 for both the flights and ERA5.
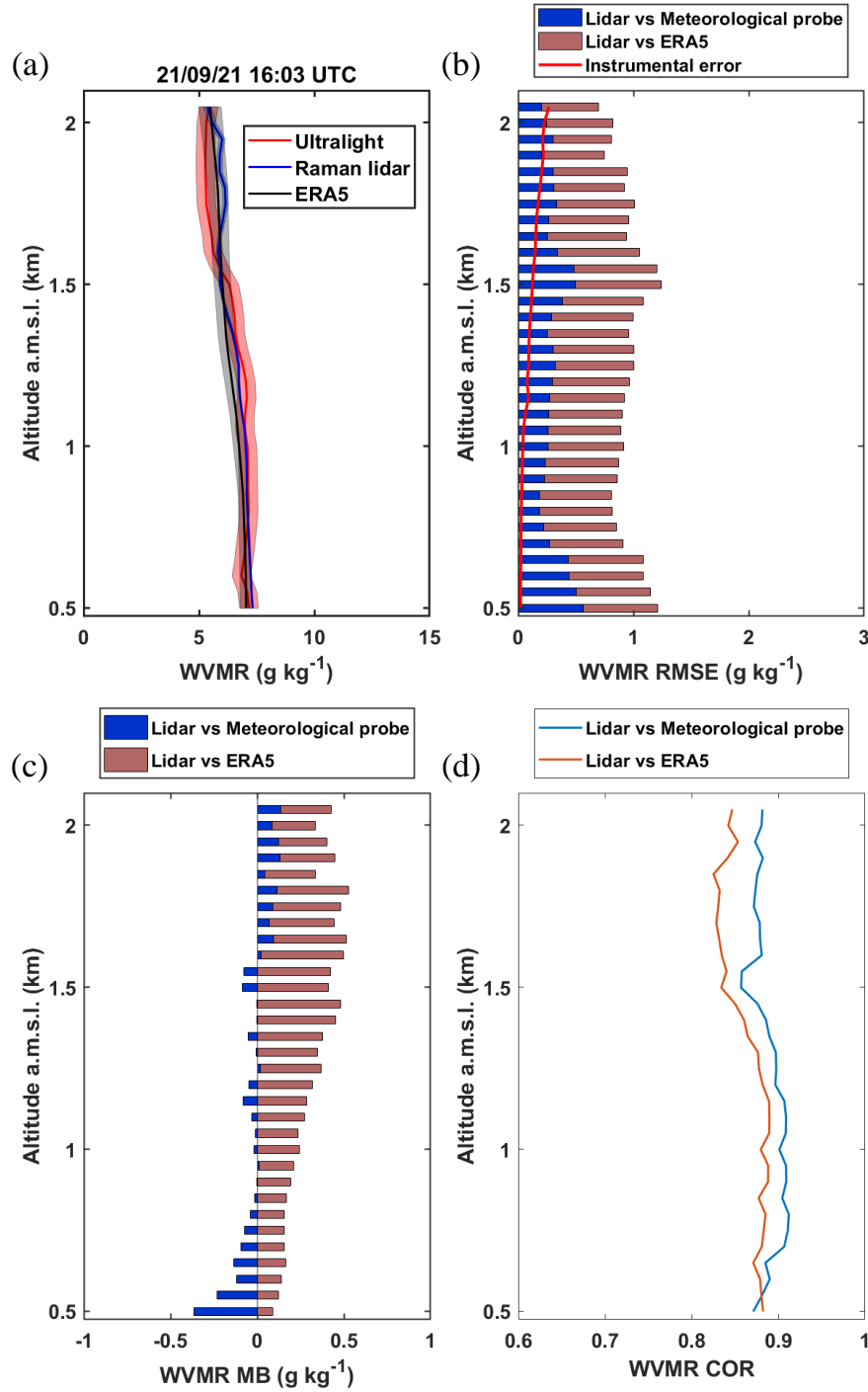
Figure 14. a) Example on 21 September 1603 UTC of the water vapour mixing ratio (WVMR) profile derived from the Raman lidar WALI and the in-situ ultralight measurements during the LEMON field campaign. The corresponding ERA5 profile is also shown. b) Root mean square error (RMSE) and instrumental error (solid red line), c) mean bias (MB) and correlation coefficient (COR) between i) the lidar and the aircraft measurements and ii) the lidar and ERA5.

### 4.6 Calibration during the WaLiNeAs field campaign

During the WaLiNeAs field campaign, the same procedure as described in Subsection 4.1.2 was used (Laly et al., 2024). The calibration curve is shown in Figure 15. As in the PARCS field campaign, the intercept is 0.27 g kg$^{-1}$, but $K_0$ is lower at 103 g kg$^{-1}$. This value remains

consistent with that of PARCS, with a non-seeded laser. Assuming no bias in the measurements, it is worth noting that a forced regression with an intercept of 0 gives $K_0 = 107$ g kg$^{-1}$. In fact, the two calibrations give very similar WVMR profiles in this case.
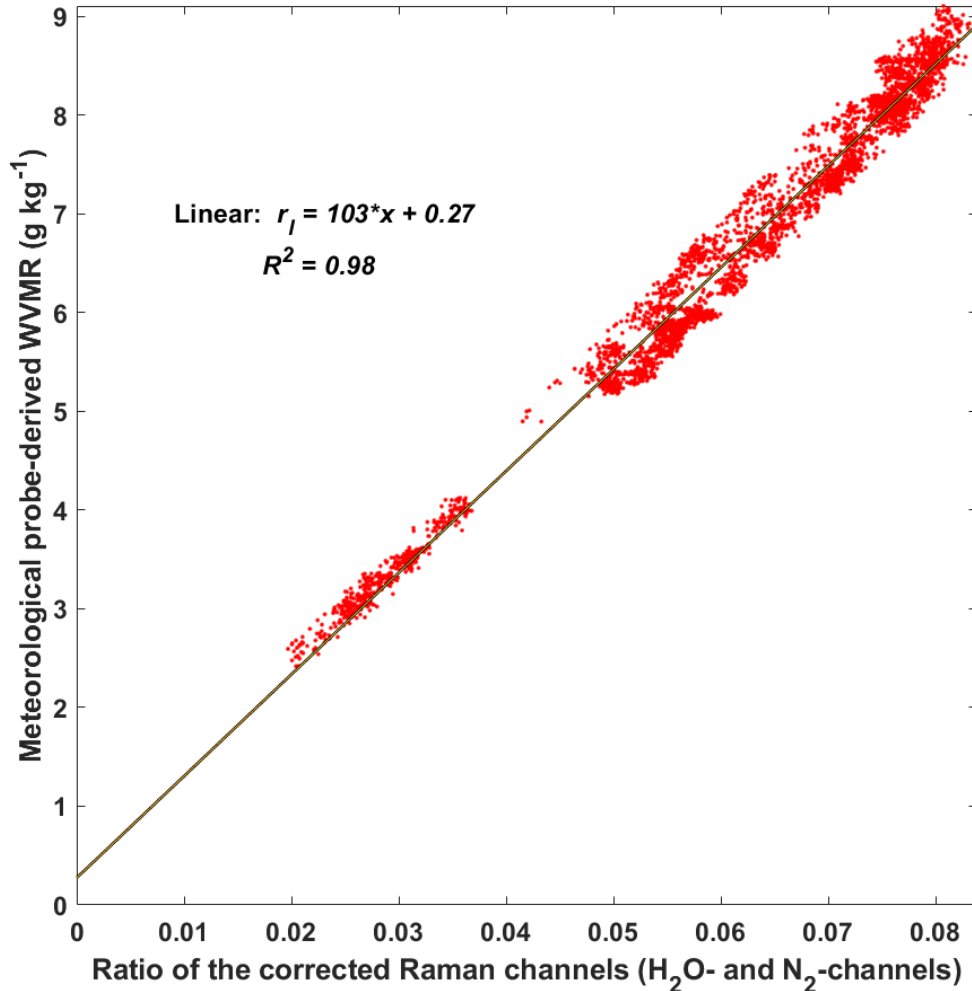
Figure 15. Calibration of the WALI lidar during the WaLiNeAs field campaign by comparison of lidar measurements at 200 m a.g.l. and in situ meteorological measurements at 5 m a.g.l. The linear regression is represented by the solid black line and the equation is given in the figure with $r_l$ the WVMR and x the corrected ration of the Raman channels. The Pearson correlation coefficient $R^2$ is also given.

## 4.7 Instrumental noise vs atmospheric variability

In this last sub-section, we show the important influence of the atmospheric variability relative to the instrumental noise on the calibration error. The detection variance and the natural atmospheric variability are assumed to be uncorrelated. Furthermore, the atmospheric variability is correlated in time and not the instrumental noise. The simplest method to separate the atmospheric variance from the noise contribution is the autocovariance method (Eq. 20) (Lenschow et al., 2000). Indeed, the atmospheric variance can be obtained from the autocovariance function of the WVMR by extrapolating the tails (non-zero lags) to zero lag

with a power-law fit. Since the autocovariance at zero lag is the total variance, the instrumental noise variance is the difference between the two (Behrendt et al., 2015).

The PARCS and L-WAIVE field campaigns are used as illustrations because they are representative of the two typical moisture variabilities encountered during this study. During the PARCS field campaign, daylight is continuous at the date and latitude of the lidar station. As shown in Figure 16a, the instrumental noise dominates, accounting for more than 56% of the variance. The detection noise then leads to a standard deviation of $\sim$0.17 g kg$^{-1}$. Arctic conditions are associated with low WVMRs, which are highly sensitive to the advection of air masses, which explains the rest of the RMSE. For the L-WAIVE field campaign, the situation is different, and the observed natural variability is the most prevalent. The proportion of signal noise decreases for the L-WAIVE field campaign, which took place on an uneven terrain. It accounts for 18% of the variance (Figure 16b), with a standard deviation of $\sim$0.25-0.30 g kg$^{-1}$. For all calibrations performed at night, the natural variability of the atmosphere dominates the variance. For the ADM-AEOLUS and post COVID field campaigns, it accounts for more than 90% of the variance is natural. It is associated with numerous advections of different origin (Chazette et al., 2017; Chazette and Royer, 2017). In the case of the LEMON field campaign, natural variability is largely dominant in the altitude range considered, accounting for more than 85% of the variance both day and night. It is associated with thunderstorm phenomena and Atlantic intrusions.

In summary, the natural variability of the atmosphere during calibration dominates the RMSE in all cases studied. To improve the accuracy in the lidar-derived WVMR, it seems preferable to perform calibrations over short time periods to limit the influence of the atmospheric variability. However, this is at the expense of the signal-to-noise ratio and therefore the range of the lidar required for calibration when compared to airborne measurements and radiosondes.
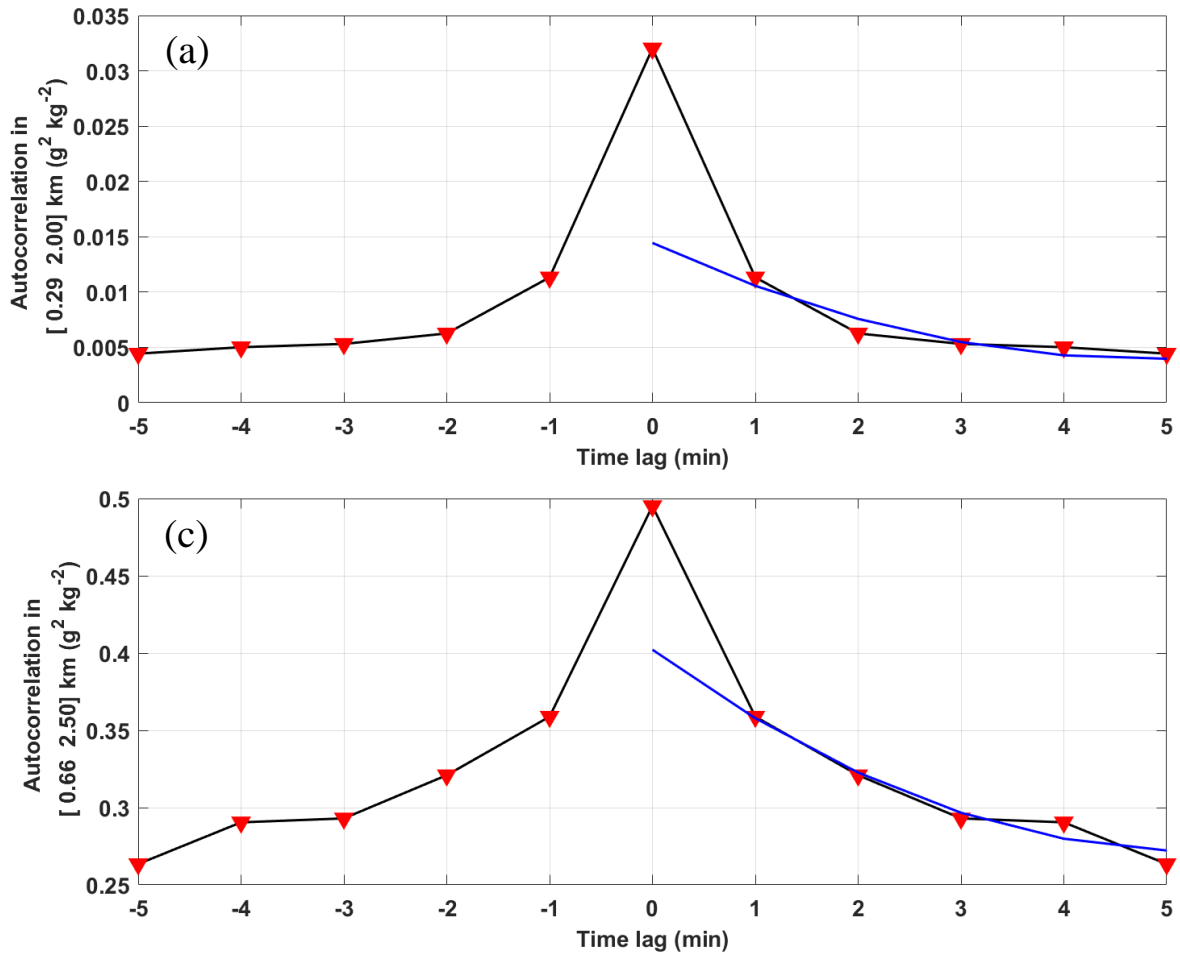
Figure 16. Autocovariance functions (dark line and red markers) around the zero-lag obtained in altitude ranges defined in the y-label for: a) the PARCS and L-WAIVE field campaigns. The blue curve represents the power function (parabola) fitted to calculate the natural contribution of the atmosphere to the origin.

## 5    Conclusive discussion

A study of the temporal evolution of the calibration of the WALI-derived WVMR was carried out based on 6 field campaigns that differed significantly in the atmospheric water vapour content, and in the temporal and vertical evolution of this essential meteorological and climatic parameter. This study showed that the calibration was quite stable over time, considering the changes in the laser emission, which did not always operate under the same conditions from one campaign to the next.

Table 2 summarises the results using the statistical variables defined in Subsection 2.4. They remain close to the recommendations of the World Meteorological Organisation (WMO), which recommends a statistical error of 0.4 g kg$^{-1}$ for atmospheric water vapour measurements (http://public.wmo.int, last access: 25 October 2024). The presence or absence of laser injection does not seem to affect the RMSE values significantly. They are significantly lower

($\sim$0.3 g kg$^{-1}$) when the calibration is based on airborne measurements taken vertically from the lidar site. The higher values (0.62 g kg$^{-1}$) for L-WAIVE are related to aircraft flights over Annecy Lake, which did not necessarily sample the same atmospheric layers as the lidar. Comparisons to radiosondes give RMSEs of between 0.35 and 0.60 g kg$^{-1}$, which are higher than expected. This is due to the respective locations of the lidar site and the reference radiosonde station for the ADM-AEOLUS and P-COVID field campaigns. The calibrations were performed in the lower troposphere (below 4 km a.m.s.l.) to obtain a good signal-to-noise ratio for the lidar data, but this zone is more subject to spatial variability than the free troposphere. In the case of the LEMON field campaign, the radiosondes were launched parallel to the lidar line of sight, but the balloons were advected horizontally very quickly and often explored different valleys. This may explain the difference between calibrating against radiosondes and airborne measurements.

The MB remains low in all cases, which demonstrates the reliability of the calibration. It can be higher compared to ERA5 when the meteorological situations encountered are more regional in origin, as in the case of thunderstorms during the LEMON field campaign. We find that there is very good agreement between the temporal changes in the WVMR profiles associated with lidar, flights or radiosondes and reanalysis, with the COR generally greater than 0.75.

The calibration coefficients are significantly different between the 6 campaigns with and without functional laser injection seeder. Values of around 105-108 g kg$^{-1}$ are found for PARCS and L-WAIVE 3 years apart. For P-COVID, the value decreases to 89 g kg$^{-1}$ ($\sim$-15%) one year after L-WAIVE. Between L-WAIVE and P-COVID, the laser was overhauled by the manufacturer. It is very possible that the emitted wavelength fluctuated, which could affect the cut-off on the rotation lines, which is linked to the interference filters (Totems et al., 2021), changed. However, the calibration is very stable when the laser is injected (117-122 g kg$^{-1}$). This is the advantage of the injector, which stabilises the wavelength of the laser emission and makes it narrower. It is therefore advisable to seed the laser to ensure the stability of the calibration coefficient of a water vapour Raman lidar but monitoring the seeder spectrum remains necessary. Although this is not ideal, calibration can be done after the campaigns using re-analyses such as ERA5. There is an agreement on the $K_0$ values to better than 5%. Such an approach is not operational and limits the value of lidar measurements for model validation purposes. The uncertainty in the value of $K_0$ can be estimated to be between 0.5 and 1.5 g kg$^{-1}$ by comparing the different approaches and taking into account the RMSD minimisation approach.

A set of vertical reference profiles must be available to ensure the calibration of the water vapour Raman lidar. Calibration on one or two radiosondes is not sufficient to guarantee the accuracy of $K_0$. Wherever possible, measurements should be taken vertically above the site, along the lidar line of sight. Ground measurements are also a valuable addition to the calibration, provided that the time series are long enough to select periods where the hypotheses of vertical stability of the lower troposphere are verified.

To our knowledge this is the first time such a long time series of lidar calibration results is published. Its analysis shows that an instrument designed to avoid vignetting and non-uniform detector sensitivity effects in the receiver, in agreement with the conclusions of David et al. (2017), is able to maintain a similar calibration, within 5% over 6 years excluding malfunctions, despite its transportation over long distances, repeated disassembly and reassembly, and the aging of components (fiber, filters and detectors).

Table 2. Statistical parameters for each field campaign using airborne (Flight) and radiosonde (RS) measurements for calibration. The calibration constant is $K_0$. $\widetilde{MB}$ stands for the altitude averaged mean bias, $\widetilde{RMSE}$ for the altitude averaged root mean square error and COR for correlation coefficient. The statistical parameters are also calculated using the colocalised ERA5 profiles. Lines in italics represent calibrations relative to a ground-based station. The bold lines represent measurements with a laser injection seeder. Values in brackets are for calibrations using ERA5.

| Field campaign | $K_0$ (g kg⁻¹) | $\widetilde{MB}$ (g kg⁻¹) | | | $\widetilde{RMSE}$ (g kg⁻¹) | | | COR | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Flight | RS | ERA5 | Flight | RS | ERA5 | Flight | RS | ERA5 |
| PARCS | 105 (109) | 0.02 | - | -0.08 | 0.23 | - | 0.53 | ≳ 0.75 | - | ≳ 0.75 |
| | *108* | *0.27* | | | *0.35* | | | *0.92* | | |
| **AEOLUS** | **117 (116)** | **-** | **0.02** | **-0.03** | **-** | **0.55** | **0.48** | **-** | **≳ 0.75** | **≳ 0.8** |
| L-WAIVE | 108 (107) | -0.08 | - | -0.04 | 0.60 | - | 0.73 | ≳ 0.85 | - | ≳ 0.70 |
| P-COVID | 89 (89) | - | -0.07 | -0.11 | - | 0.60 | 0.57 | - | ≳ 0.90 | ≳0.90 |
| **LEMON** | **121.5 (117.5)** | **-** | **0.03** | **0.24** | **-** | **0.48** | **0.91** | **-** | **≳ 0.90** | **≳ 0.75** |
| | **122** | **-0.02** | **-** | **0.29** | **0.32** | **-** | **0.65** | **≳ 0.85** | **-** | **≳ 0.80** |
| WaLiNeAs | *103* | *0.27* | | | *0.25* | | | *0.98* | | |

**Data availability.** The datasets are published open access on the AERIS database (https://en.aeris-data.fr/, last access: 25 October 2024).

**Author contributions.** PC participated to the field experiment as PI of lidar measurements, he performed the calculations, analysed the data, and wrote the paper. JT took part in the field campaigns. FL took part in WaLiNeAs field campaign. FL and JT contributed to the proofreading of the paper.

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

**Acknowledgements.** We thank Franck Toussaint, Nathalie Toussaint and Sébastien Blanchon for their help during the airborne field experiments. During PARCS, we thank the Avinor crew of Hammerfest Airport, represented by Hans-Petter Nergård, and the Air Création company are acknowledged for their hospitality. Thanks also to Xiaoxia Shang, Jean-Christophe Raut and Yoann Chazette for their help. Friendly acknowledgements to local authorities of Thales-Alenia-Space during WaLiNeAs and to Jérémy Lagarrigue for his help. Aurelien Clémençon and Nicolas Geyskens from the INSU technical department are thanked for their help with radiosonde launches during LEMON.

**Financial support.** This work was supported by the Institut National de l'Univers (INSU) of the Centre National de la Recherche Scientifique (CNRS) and the Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA). Specific financial supports were obtained for: PARCS through the INSU French Arctic Initiative; ADM-AEOLUS through the Centre National d'Etudes Spatiale (CNES); L-WAIVE through the Agence Nationale de la Recherche (ANR-16-CE01-0009); LEMON through the Horizon 2020 Framework Programme (821868); WaLiNeAs through the Agence Nationale de la Recherche (ANR-20-CE04-0001).

**References**

Agusti-Panareda, A., Vasiljevic, D., Beljaars, A., Bock, O., Guichard, F., Nuret, M., Mendez, A. G., Andersson, E., Bechtold, P., Fink, A., Hersbach, H., Lafore, J. P., Ngamini, J. B., Parker, D. J., Redelsperger, J. L., and Tompkins, A. M.: Radiosonde humidity bias correction over the West African region for the special AMMA reanalysis at ECMWF, Q. J. R. Meteorol. Soc., 135, 595–617, 2009.

Ansmann, A., Riebesell, M., Wandinger, U., Weitkamp, C., Voss, E., Lahmann, W., and Michaelis, W.: Combined raman elastic-backscatter LIDAR for vertical profiling of moisture, aerosol extinction, backscatter, and LIDAR ratio, Appl. Phys. B Photophysics Laser Chem., 55, 18–28, 1992.

Behrendt, A., Wulfmeyer, V., Hammann, E., Muppa, S. K., and Pal, S.: Profiles of second- to fourth-order moments of turbulent temperature fluctuations in the convective boundary layer: first measurements with rotational Raman lidar, Atmos. Chem. Phys., 15, 5485–5500, https://doi.org/10.5194/acp-15-5485-2015, 2015.

Bock, O. and Nuret, M.: Verification of NWP Model Analyses and Radiosonde Humidity Data with GPS Precipitable Water Vapor Estimates during AMMA, https://doi.org/10.1175/2009WAF2222239.1, 2009.

Bock, O., Bosser, 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.

Boylan, J. W. and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ., 40, 4946–4959, 2006.

Buck, A. L.: New Equations for Computing Vapor Pressure and Enhancement Factor, Cover J. Appl. Meteorol. Climatol. J. Appl. Meteorol. Climatol., 20, 1527–1532, https://doi.org/https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2, 1981.

Chazette, P. and Royer, P.: Springtime major pollution events by aerosol over Paris Area: From a case study to a multiannual analysis, J. Geophys. Res. Atmos., 122, 8101–8119, https://doi.org/10.1002/2017JD026713, 2017.

Chazette, P., Marnas, F., Totems, J., and Shang, X.: Comparison of IASI water vapor retrieval with $H_2O$-Raman lidar in the framework of the Mediterranean HyMeX and ChArMEx programs, Atmos. Chem. Phys., 14, 9583–9596, https://doi.org/10.5194/acp-14-9583-2014, 2014a.

Chazette, P., Marnas, F., and Totems, J.: The mobile Water vapor Aerosol Raman LIdar and its implication in the framework of the HyMeX and ChArMEx programs: application to a dust transport process, Atmos. Meas. Tech., 7, 1629–1647, https://doi.org/10.5194/amt-7-1629-2014, 2014b.

Chazette, P., Totems, J., and Shang, X.: Atmospheric aerosol variability above the Paris Area during the 2015 heat wave - Comparison with the 2003 and 2006 heat waves, Atmos. Environ., 170, 216–233, https://doi.org/10.1016/j.atmosenv.2017.09.055, 2017.

Chazette, P., Raut, J. C., and Totems, J.: Springtime aerosol load as observed from ground-based and airborne lidars over northern Norway, Atmos. Chem. Phys., 18, 13075–13095, https://doi.org/10.5194/acp-18-13075-2018, 2018.

Chazette, P., Totems, J., Baron, A., Flamant, C., and Bony, S.: Trade-wind clouds and aerosols characterized by airborne horizontal lidar measurements during the EUREC4A field campaign, Earth Syst. Sci. Data, 12, 2919–2936, https://doi.org/10.5194/essd-12-2919-2020, 2020.

Chazette, P., Flamant, C., Sodemann, H., Totems, J., Monod, A., Dieudonné, E., Baron, A., Seidl, A., Steen-Larsen, H. C., Doira, P., Durand, A., and Ravier, S.: Experimental investigation of the stable water isotope distribution in an Alpine lake environment (L-WAIVE), Atmos. Chem. Phys., 21, 10911–10937, https://doi.org/10.5194/acp-21-10911-2021, 2021.

Collard, A. D. and McNally, A. P.: The assimilation of infrared atmospheric sounding interferometer radiances at ECMWF, Q. J. R. Meteorol. Soc., https://doi.org/10.1002/qj.410, 2009.

Cooney, J. A.: Remote Measurements of Atmospheric Water Vapor Profiles Using the Raman Component of Laser Backscatter, J. Appl. Meteorol., 9, 182–184, https://doi.org/10.1175/1520-0450(1970)009<0182:RMOAWV>2.0.CO;2, 1970.

David, L., Bock, O., Thom, C., Bosser, P., and Pelon, J.: Study and mitigation of calibration factor instabilities in a water vapor Raman lidar, Atmos. Meas. Tech., 10, 2745–2758, https://doi.org/10.5194/amt-10-2745-2017, 2017.

Ferrare, R. A., Melfi, S. H., Whiteman, D. N., Evans, K. D., Schmidlin, F. J., and Starr, O.: A Comparison of Water Vapor Measurements Made by Raman Lidar and Radiosondes, J. Atmos. Ocean. Technol., 12, 1177–1195, https://doi.org/https://doi.org/10.1175/1520-0426(1995)012%3C1177:ACOWVM%3E2.0.CO;2, 1995.

Flamant, C., Chazette, P., Caumont, O., Di Girolamo, P., Behrendt, A., Sicard, M., Totems, J., Lange, D., Fourrié, N., Brousseau, P., Augros, C., Baron, A., Cacciani, M., Comerón, A., De Rosa, B., Ducrocq, V., Genau, P., Labatut, L., Muñoz-Porcar, C., Rodríguez-Gómez, A., Summa, D., Thundathil, R., and Wulfmeyer, V.: A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative, Bull. Atmos. Sci. Technol., 2, https://doi.org/10.1007/s42865-021-00037-6, 2021.

Fourrié, N., Nuret, M., Brousseau, P., Caumont, O., Doerenbecher, A., Wattrelot, E., Moll, P., Bénichou, H., Puech, D., Bock, O., Bosser, P., Chazette, P., Flamant, C., Di Girolamo, P., Richard, E., and Saïd, F.: The AROME-WMED reanalyses of the first special observation period of the Hydrological cycle in the Mediterranean experiment (HyMeX), Geosci. Model

Dev., 12, https://doi.org/10.5194/gmd-12-2657-2019, 2019.

Di Girolamo, P., De Rosa, B., Flamant, C., Summa, D., Bousquet, O., Chazette, P., Totems, J., and Cacciani, M.: Water vapor mixing ratio and temperature inter-comparison results in the framework of the Hydrological Cycle in the Mediterranean Experiment—Special Observation Period 1, Bull. Atmos. Sci. Technol., 1, 113–153, https://doi.org/10.1007/s42865-020-00008-3, 2020.

Hamperl, J., Capitaine, C., Dherbecourt, J. B., Raybaut, M., Chazette, P., Totems, J., Grouiez, B., Régalia, L., Santagata, R., Evesque, C., Melkonian, J. M., Godard, A., Seidl, A., Sodemann, H., and Flamant, C.: Differential absorption lidar for water vapor isotopologues in the 1.98 µm spectral region: Sensitivity analysis with respect to regional atmospheric variability, Atmos. Meas. Tech., 14, 6675–6693, https://doi.org/10.5194/amt-14-6675-2021, 2021.

Held, I. M. and Soden, B. J.: WATER VAPOR FEEDBACK AND GLOBAL WARMING1, Annu. Rev. Energy Environ., 25, 441–475, https://doi.org/10.1146/annurev.energy.25.1.441, 2000.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi A., Muñoz SabaterJ., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, Copernicus Clim. Chang. Serv. Clim. Data Store (CDS), DOI 10.24381/cds.adbb2d47, 147, 4186–4227, https://doi.org/10.24381/cds.bd0915c6, 2023.

Ingleby, B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., James, S. J., Tennant, W., and Weedon, R.: Progress toward high-resolution, real-time radiosonde reports, Bull. Am. Meteorol. Soc., 97, 2149–2161, https://doi.org/10.1175/BAMS-D-15-00169.1, 2016.

IPCC: Climate change 2022: Impacts, Adaptation and Vulnerability. Summary for policymakers. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, in: United Nations Environment Programme UNEP, vol. AR6, Cambridge University Press, 551–712, 2022.

Kim, Y., Sartelet, K., Raut, J.-C. C. J.-C., and Chazette, P.: Evaluation of the Weather Research and Forecast / Urban Model Over Greater Paris, Boundary-Layer Meteorol., 149, 105–132, https://doi.org/10.1007/s10546-013-9838-6, 2013.

Labzovskii, L. D., Papayannis, A., Binietoglou, I., Banks, R. F., Baldasano, J. M., Toanca, F., Tzanis, C. G., and Christodoulakis, J.: Relative humidity vertical profiling using lidar-based synergistic methods in the framework of the Hygra-CD campaign, Ann. Geophys., 36, 213–229, https://doi.org/10.5194/ANGEO-36-213-2018, 2018.

Laly, F., Chazette, P., Totems, J., Lagarrigue, J., Forges, L., and Flamant, C.: Water vapor Raman-lidar observations from multiple sites in the framework of WaLiNeAs, Earth Syst. Sci. Data, 16, 5579–5602, https://doi.org/10.5194/ESSD-16-5579-2024, 2024.

Lange, D., Behrendt, A., and Wulfmeyer, V.: Compact Operational Tropospheric Water Vapor and Temperature Raman Lidar with Turbulence Resolution, Geophys. Res. Lett., 46, 14844–14853, https://doi.org/10.1029/2019GL085774, 2019.

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,

https://doi.org/10.1364/AO.47.005592, 2008.

Lenschow, D. H., Wulfmeyer, V., and Senff, C.: Measuring second through fourth-order moments in noisy data, J. Atmos. Ocean. Technol., 17, 1330–1347, https://doi.org/https://doi.org/10.1175/1520-0426(2000)017<1330:MSTFOM>2.0.CO;2, 2000.

Melfi, S. H. and Whiteman, D.: Observation of Lower-Atmospheric Moisture Structure and Its Evolution Using a Raman Lidar, Bull. Am. Meteorol. Soc., 66, 1288=1292, 1985.

Melfi, S. H., Lawrence, J. D., and McCormick, M. P.: Observation of raman scattering by water vapor in the atmosphere, Appl. Phys. Lett., 15, 295–297, 1969.

Raut, J.-C. and Chazette, P.: Assessment of vertically-resolved PM<inf>10</inf> from mobile lidar observations, Atmos. Chem. Phys., 9, 2009.

Reichardt, J., Wandinger, U., Klein, V., Mattis, I., Hilber, B., and Begbie, R.: RAMSES: German meteorological service autonomous Raman lidar for water vapor, temperature, aerosol, and cloud measurements, Appl. Opt., 51, 8111–8131, https://doi.org/10.1364/AO.51.008111, 2012.

Royer, P., Chazette, P., Lardier, M., and Sauvage, L.: Aerosol content survey by mini N2-Raman lidar: Application to local and long-range transport aerosols, Atmos. Environ., 45, 7487–7495, https://doi.org/10.1016/j.atmosenv.2010.11.001, 2011.

Serreze, M. C., Barrett, A. P., and Stroeve, J.: Recent changes in tropospheric water vapor over the Arctic as assessed from radiosondes and atmospheric reanalyses, J. Geophys. Res. Atmos., 117, 10104, https://doi.org/10.1029/2011JD017421, 2012.

De Tomasi, F. and Perrone, M. R.: Lidar measurements of tropospheric water vapor and aerosol profiles over southeastern Italy, J. Geophys. Res. Atmos., 108, 4286, https://doi.org/10.1029/2002jd002781, 2003.

Tombette, M., Chazette, P., Sportisse, B., and Roustan, Y.: Simulation of aerosol optical properties over Europe with a 3-D size-resolved aerosol model: comparisons with AERONET data, Atmos. Chem. Phys., 8, 7115–7132, https://doi.org/10.5194/acp-8-7115-2008, 2008.

Totems, J. and Chazette, P.: Calibration of a water vapour Raman lidar with a kite-based humidity sensor, Atmos. Meas. Tech., 9, 1083–1094, https://doi.org/10.5194/amt-9-1083-2016, 2016.

Totems, J., Chazette, P., Shang, X., Flamant, C., Raut, J.-C., Doerenbecher, A., Ducroq, V., Bock, O., and Marnas, F.: Water Vapor Measurements by Mobile Raman Lidar over the Mediterranean Sea in the Framework of HyMex: Application to Multi-Platform Validation of Moisture Profiles, in: EPJ Web of Conferences, https://doi.org/10.1051/epjconf/201611926006, 2016.

Totems, J., Chazette, P., and Raut, J. J.-C.: Accuracy of current Arctic springtime water vapour estimates, assessed by Raman lidar, Q. J. R. Meteorol. Soc., 145, 1234–1249, https://doi.org/10.1002/qj.3492, 2019.

Totems, J., Chazette, P., and Baron, A.: Mitigation of bias sources for atmospheric temperature and humidity in the mobile Raman Weather and Aerosol Lidar (WALI), Atmos. Meas. Tech., 14, 7525–7544, https://doi.org/10.5194/amt-14-7525-2021, 2021.

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. Technol., 16, 1062–1076, 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. R. Meteorol. Soc., 114, 1471–1484, 1988.

Wang, Y., Sartelet, K. N. N., Bocquet, M., Chazette, P., Sicard, M., D'Amico, G., Léon, J. F. F., Alados-Arboledas, L., Amodeo, A., Augustin, P., Bach, J., Belegante, L., Binietoglou, I., Bush, X., Comerón, A., Delbarre, H., García-Vízcaino, D., Guerrero-Rascado, J. L. L., Hervo,

M., Iarlori, M., Kokkalis, P., Lange, D., Molero, F., Montoux, N., Muñoz, A., Muñoz, C., Nicolae, D., Papayannis, A., Pappalardo, G., Preissler, J., Rizi, V., Rocadenbosch, F., Sellegri, K., Wagner, F., Dulac, F., Rizi, V., Rocadenbosch, F., Sellegri, K., Wagner, F., and Dulac, F.: Assimilation of lidar signals: application to aerosol forecasting in the western Mediterranean basin, Atmos. Chem. Phys., 14, 12031–12053, https://doi.org/10.5194/acp-14-12031-2014, 2014.

Whiteman, D. N.: Examination of the traditional Raman lidar technique I Evaluating the temperature-dependent lidar equations, Appl. Opt., 42, 2571, https://doi.org/10.1364/AO.42.002571, 2003.

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. Opt., 50, 2170–2176, https://doi.org/10.1364/AO.50.002170, 2011.

Wulfmeyer, V., Hardesty, R. M., Turner, D. D., Behrendt, A., Cadeddu, M. P., Di Girolamo, P., Schlüssel, P., Van Baelen, J., and Zus, F.: A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles, https://doi.org/10.1002/2014RG000476, 1 September 2015.