Differential absorption lidar measurements of water vapor by the High Altitude Lidar Observatory (HALO): Retrieval framework and first results

Brian J. Carroll[1], Amin R. Nehrir[2], Susan Kooi[3], James Collins[3], Rory A. Barton-Grimley[2], Anthony Notari[2], David B. Harper[2], Joseph Lee[2]

[1] NASA Postdoctoral Program Fellow, NASA Langley Research Center, Hampton, VA, USA
[2] NASA Langley Research Center, Hampton, VA, United States
[3] Science Systems and Applications, Inc., Hampton, VA, United States

Correspondence to: Brian J. Carroll, Amin R. Nehrir (brian.j.carroll@nasa.gov, amin.r.nehrir@nasa.gov)

Abstract. Airborne differential absorption lidar (DIAL) offers a uniquely capable solution to the problem of measuring water vapor (WV) with high precision, accuracy, and resolution throughout the troposphere and lower stratosphere. The High Altitude Lidar Observatory (HALO) airborne WV DIAL was recently developed at NASA Langley Research Center and was first deployed in 2019. It uses four wavelengths near 935 nm to achieve sensitivity over a wide dynamic range, and simultaneously employs 1064 nm backscatter and 532 nm high spectral resolution lidar (HSRL) measurements for aerosol and cloud profiling. A key component of the WV retrieval framework is flexibly trading resolution for precision to achieve optimal data sets for scientific objectives across scales. An approach to retrieving WV in the lowest few hundred meters of the atmosphere using the strong surface return signal is also presented.

The five maiden flights of the HALO WV DIAL spanned the tropics through midlatitudes with a wide range of atmospheric conditions, but opportunities for validation were sparse. Comparisons to dropsonde WV profiles were qualitatively in good agreement, though statistical analysis was impossible due to systematic error in the dropsonde measurements. Comparison of HALO to in situ WV measurements onboard the aircraft showed no substantial bias across three orders of magnitude, despite variance ($R^2 = 0.66$) that may be largely attributed to spatiotemporal variability. Precipitable water vapor measurements from the spaceborne sounders AIRS and IASI compared very well to HALO with $R^2 > 0.96$ over ocean.

1 Introduction

Water vapor (WV) is a key component of the Earth's atmosphere and water cycle, playing major roles in cloud, weather, and climate processes, including radiative balance as the most dominant greenhouse gas (Trenberth et al. 2007). The need for accurate WV measurement across scales is widely recognized, as is the value of remote sensing for providing such measurements with desirable spatiotemporal coverage (Bony et al. 2006, Bony et al. 2015, Sherwood et al. 2010, Teixeira et al. 2021, Wulfmeyer et al. 2015).

Radiosonde networks have long provided the most consistent and extensive in situ WV profile measurements globally, though the operational network is limited by resources and personnel to a finite number of sites, only over land, and two launches per day (Ferreira et al. 2019). Spaceborne passive sounders have the great advantage of daily global coverage, and their column and cloud products are important contributors to operational forecasting and climate research (Wulfmeyer et al. 2015). However, their WV retrieval vertical resolution is roughly $1 - 2$ km at best in the troposphere and accuracy is closely tied to the observed scene as well as the prior inputs to the retrievals. This makes spaceborne passive sounders incapable of capturing many lower tropospheric, cloud, and planetary boundary layer (PBL) processes relevant on the weather and climate timescales. Measurement techniques with higher spatiotemporal resolution, such as lidars, have become vital to studying WV processes as our understanding improves and more demanding measurement criteria have become prevalent (Wulfmeyer et al. 2015). Due to the challenging nature of developing and deploying airborne lidars, there are not many within the research community. These sparse but capable airborne lidars are complimented with emerging technologies that show promise for enabling dense, low-cost surface-based networks (Nehrir et al. 2012, Spuler et al. 2015, Spuler et al. 2021). The mobility of airborne lidar cements it as an important asset for Earth science, able to observe large regions with high precision and resolution, including over oceans where permanent observing networks are impractical.

Two types of lidar are commonly used for atmospheric WV measurements: Raman lidar (e.g., Cooney 1970, Eichinger et al 1999, Goldsmith et al. 1998, Leblanc et al. 2012, Philbrick 1994, Whiteman et al. 1992) and differential absorption lidar (DIAL; e.g., Browell et al. 1998, Ehret et al. 1993, Ferrare et al. 2004, Nehrir et al. 2011, Nehrir et al. 2017, Späth et al. 2016, Spuler et al. 2021, Wirth et al. 2009). Raman lidar systems provide the advantageous capability of monitoring multiple gas species simultaneously, but this comes at the cost of requiring large, high peak power ultraviolet lasers to overcome small Raman scattering cross-sections, and a need for frequent calibration. WV DIAL avoids the need for high peak power ultraviolet lasers, but only measures one species and requires the use of single frequency and frequency-agile pulsed lasers which have been the focus of decades of laser research within the DIAL community. A significant benefit of the DIAL technique that results from the stringent laser transmitter requirements is the lack of a need for external calibration, as this technique relies on the ratio of elastic backscatter signals from a closely spaced wavelength pair, one online and one offline of a WV absorption line. The challenges of signal attenuation and limited dynamic range with a single DIAL wavelength pair can be overcome by utilizing multiple wavelength pairs spread along the side of the WV absorption line. The direct, calibration-free measurement of WV profiles with reduction in overall complexity resulting from emerging laser technologies make WV DIAL a unique and important tool for atmospheric studies, and well-suited for airborne and space-based implementations.

Airborne WV DIAL systems were first employed in the 1980s (Browell 1983), with subsequent decades of literature documenting advances in the theoretical framework and instrument designs (e.g., Behrendt et al. 2007, Browell et al. 1998,

65 Ehret et al. 1993, Higdon et al. 1994, Ismail & Browell 1989, Wulfmeyer 1998) as well as measurements and applications (e.g., Carroll et al. 2021, Ismail et al. 2009, Kiemle et al. 2017, Schafler et al. 2021, Wakimoto et al. 2006, Wulfmeyer et al. 2006). Two airborne WV DIALs of particular note for this manuscript are the NASA Langley Research Center (LaRC) Laser Atmospheric Sensing Experiment (LASE, Moore et al. 1997) and the Deutsches Zentrum für Luft- und Raumfahrt (DLR) airborne WAter vapour Lidar Experiment in Space (WALES) instrument (Wirth et al. 2009). LASE is the

70 predecessor to the focus of this work, the High Altitude Lidar Observatory (HALO) WV DIAL. HALO has similar capabilities to WALES with simultaneous WV DIAL and high spectral resolution lidar (HSRL, Hair et al. 2008) capability for aerosols and clouds, however, HALO's modular design and exploitation of emerging laser and receiver technologies allows for a more compact form factor that permits operation on a wide range of aircraft.

75 This manuscript reports on the HALO WV DIAL retrieval framework and measurement capabilities based on its maiden flights in 2019, including comparisons against the other instrumentation that was available. Section 2 describes the field campaign and instruments utilized in this work. Section 3 provides a brief review of WV DIAL principles followed by relevant specifics of HALO's design. Section 4 details the HALO WV retrieval methodology. Section 5 presents comparisons of HALO WV measurements with available in situ instrumentation and spaceborne sounders. Conclusions and

80 future directions are discussed in Section 6.

2 Instruments and comparison methodology

2.1 The NASA Aeolus Cal/Val Test Flight Campaign

The HALO measurements described in this manuscript were collected during the NASA Aeolus Cal/Val Test Flight campaign from 17-30 April 2019. The campaign was designed to provide calibration/validation comparisons with the

85 European Space Agency (ESA) space-borne Doppler wind lidar mission ADM-Aeolus (Stoffelen et al. 2005). The NASA Doppler Aerosol WiNd lidar (DAWN, Kavaya et al. 2014) provided wind profiles for this purpose, and HALO was deployed to validate aerosol measurements via HSRL as well as opportunistically test the WV DIAL capability. Further details of this campaign, including overviews of each flight, are given in Bedka et al. (2021). Data is archived and publicly available online (NASA/LARC/SD/ASDC, n.d.).

90

This campaign consisted of five flights of the NASA DC-8 aircraft: four from the NASA Armstrong Flight Research Center in Palmdale, California and one from Kona, Hawaii. These flights were mainly over the northeastern Pacific spanning midlatitudes to the tropics and observing a wide range of atmospheric conditions. There were only a few hours of observations over land. Installed along with HALO and DAWN on the DC-8 was the Diode Laser Hygrometer (DLH, Diskin

95 et al. 2002) to provide in situ WV measurements for comparison with HALO. Dropsondes were also deployed to validate

wind and WV measurements. This campaign provided the maiden flights for the HALO WV DIAL configuration which also employed simultaneous HSRL measurements of aerosols and clouds.

2.2 HALO

The HALO airborne lidar was developed at NASA LaRC to address observational needs in the Earth Sciences, specifically focusing on atmospheric dynamics, composition, radiation, and the carbon cycle, in addition to serving as a technology testbed for future space-borne DIAL. HALO has a modular design with the capability to measure WV or methane mixing ratios via the DIAL technique along with aerosol, cloud, and ocean optical properties via the HSRL technique. This multi-functional design allows any two measurement capabilities simultaneously (i.e., WV DIAL/HSRL, WV DIAL/methane DIAL, or methane DIAL/HSRL) with rapid reconfiguration by using three modular laser transmitters and a single multi-channel, multi-wavelength receiver. Because of its compact design, the instrument can be flown on most research aircraft including autonomous operation aboard the high-altitude NASA ER-2. The methane DIAL/HSRL configuration has successfully flown in multiple field campaigns (e.g., Davis et al. 2021, Wu et al. 2021). This manuscript focuses on the WV DIAL retrieval framework and results from its maiden flights during the Aeolus cal/val campaign. Details of the HALO instrument design and performance will follow in an instrument paper.

2.3 Dropsondes

The dropsondes deployed from the DC-8 were expendable digital dropsondes (XDD) with the Yankee Environmental Systems Inc. High-Definition Sounding System (HDSS). The HDSS and XDD are presented in full in Black et al. (2017) and will be referred to as sondes in this manuscript. The pressure, temperature, and humidity measurements were taken at 2 Hz, which is roughly 8 meters in vertical resolution. These sondes have been used in previous field campaigns (e.g. Doyle et al. 2017), but the Aeolus cal/val campaign was the first deployment with a new relative humidity sensor. This new sensor was found during the campaign to have a time-lagged response that varied with altitude, leading to large errors in the WV profiles (Bedka et al. 2021). This prevents a typical approach of using sonde measurements to quantitatively validate HALO WV profiles. Only a cursory qualitative comparison is made in Section 5. More rigorous comparison to dropsondes will be an objective of future HALO flights.

2.4 DLH

The NASA Langley/Ames DLH (hereafter just DLH) is designed to measure in situ WV in the troposphere and lower stratosphere while flying on board research aircraft, in this case the NASA DC-8 (Diskin et al. 2002). It reports WV mixing ratio at 1 Hz and is considered to be a community standard for accurate measurements of WV from the upper troposphere and lower stratosphere (UT/LS) down to the surface. The measurement relies on WV absorption of diode laser light in the 1.4 µm spectral region, tuning to either a weak or strong absorption line to accommodate the wide range of WV concentrations in the atmosphere. It uses an open-path double-pass configuration between the laser transceiver inside the

DC-8 pointing out a modified window panel and a retroreflecting panel mounted on an outboard engine nacelle. The calibration, algorithm, and validation study by Podolske et al. (2003) found a 1σ error estimate of 3.7%. This qualifies the DLH as a viable validation instrument for the HALO WV measurements, although conclusions are limited by the instruments' different sampling volumes. Comparisons between HALO and DLH were made in two ways for this study: Examining a single spiral descent profile, which was preceded by HALO overpasses, and a more statistically robust near-field comparison throughout the campaign.

2.5 Satellites

The Atmospheric Infrared Sounder (AIRS) and Infrared Atmospheric Sounding Interferometer (IASI) are community standards for spaceborne WV measurement and are used here for comparison to HALO WV products. The AIRS and IASI precipitable water vapor (PWV) products were validated against ground-based measurements by Roman et al. (2016), finding that both satellites generally fell within a 5% error range, except for very dry (< 5 mm) or very moist (> 50 mm) regions which tended to have larger wet or dry biases, respectively.

2.5.1 Atmospheric Infrared Sounder (AIRS)

The AIRS instrument onboard the polar-orbiting NASA Aqua satellite retrieves temperature, humidity, and trace gas information with daily global coverage to support weather prediction, study the water and energy cycles, and provide a record of several critical greenhouse gases (Chahine et al. 2006, Le Marshall et al. 2006). The hyperspectral sounder measures infrared radiation from 3.7 to 15 µm in a cross-track scanning pattern with a 13.5 km diameter nadir field of view (FOV), stretching to 22.4 km along-track by 41.0 km cross-track at scan edge.

The AIRS Level 2 Version 7 data products used in this work (PWV and WV mixing ratio profiles) are derived from 3-by-3 arrays of AIRS FOVs, and uncertainty estimates are provided with each retrieval (Thrastarson et al 2020). The AIRS profiles used here are from the support data files which have 100 vertical levels, but it is important to note that the 100 levels represent a much finer resolution than the actual independent information within the profile. AIRS WV vertical resolution is actually limited to 1-2 km at best within the troposphere (Gettelman et al. 2004, Thrastarson et al 2020, Wong et al. 2015, Wulfmeyer et al. 2015).

2.5.2 Infrared Atmospheric Sounding Interferometer (IASI)

The IASI instrument onboard the polar-orbiting ESA MetOp satellite series is designed to support numerical weather prediction by providing information on temperature, humidity, and some trace gases with global coverage twice per day (e.g. Clerbaux et al. 2009, Hilton et al. 2009, Hilton et al. 2012, Klaes et al. 2007). IASI is composed of an imaging system and a Fourier transform spectrometer to analyze infrared spectra between 3.6 µm to 15.5 µm. IASI scans across-track with 30 elementary fields of view (EFOV), each of which contains a 2-by-2 grid of 4 instantaneous fields of view (IFOV). The IFOV

footprint is a circle of 12 km diameter at nadir and an ellipse of 20 km along-track by 39 km across-track at swath edge. While IASI humidity is reported at up to 101 vertical levels, there are only a maximum of 10 independent pieces of information within the profile and sensitivity is lowest in the lowest few kilometers of the troposphere.

The IASI data used in this work is the operational Level 2 data Version 6 for the instrument onboard MetOp-B. The data files report PWV and WV mixing ratio profiles for each IFOV, but do not provide associated uncertainties beyond the instrument accuracy requirement of 10%. To provide some estimate of uncertainty related to spatiotemporal variability in comparison to HALO, the mean and standard deviation of the 4 IFOV products within a given EFOV were calculated and used here.

3 DIAL principle and HALO measurements

3.1 DIAL theory

Derived from the lidar equation, the single-scattering DIAL equation can be written as

$$N\left(r + \frac{\Delta r}{2}\right) = \frac{1}{2\,\Delta r\,\Delta\sigma(\lambda_{on}, \lambda_{off}, r)} \ln\left(\frac{P(\lambda_{on}, r)}{P(\lambda_{on}, r+\Delta r)} \frac{P(\lambda_{off}, r+\Delta r)}{P(\lambda_{off}, r)}\right), \tag{1}$$

where N is the number density of WV (molecules/cm^3), r is range from the lidar (cm), Δr is the range cell length over which N is calculated, $\Delta\sigma$ is the difference between online and offline absorption cross sections (cm^2), λ is the online or offline wavelength as indicated by subscript, and P is the measured backscatter signal (Schotland 1974). This dependence on the ratio of online to offline signals yields a direct, calibration-free measurement of WV assuming the following conditions are true: the spectral separation between the on and off wavelengths is small enough to neglect differences in atmospheric transmission properties such that differential attenuation is only due to WV, the spatiotemporal variations in the atmosphere are negligible between the two laser pulses, the pulsed laser exhibits high spectral purity (Ismail & Browell 1989), Doppler broadening is constant across Δr, and instrument differential transmission is constant within a retrieval (Schotland 1974). It is also assumed that the WV concentration in the retrieval bin is constant.

In practice, accurate measurements require careful consideration of $\Delta\sigma$ calculation, close monitoring of instrument and laser characteristics, and corrections to the measured signals. These sources of error have been presented at length in the literature (e.g., Schotland 1974, Remsberg & Gordley 1978, Ismail & Browell 1989, Wulfmeyer & Bosenberg 1998) and reproduction is unwarranted here. Some details of implementation with HALO will be discussed further in Section 4. We will highlight here the sensitivity of statistical uncertainty to the temporal (or along-track) and range resolutions of the DIAL measurement:

$$\delta N \propto (\Delta t)^{-0.5}(\Delta r)^{-1.5}, \tag{2}$$

where δN is statistical uncertainty in the DIAL measurement of N and Δt is temporal resolution (Ismail & Browell 1989). This proportionality indicates that the statistical uncertainty of a DIAL measurement can be reduced by decreasing

resolution, with greater sensitivity to changes in Δr than Δt. This is caused by Poisson statistics in both dimensions giving
190 the 0.5 exponent, and the differential absorption along the path length contributing 1.0 Δr (equation 2 in Nehrir et al. 2009; Ismail & Browell 1989). Eq. 2 assumes validity of Poisson sampling statistics for the DIAL system in question, which is typically applicable to photon counting systems but can also be applied to analog detection systems such as HALO via a noise scale factor (NSF, Liu et al. 2006).

195 Optical depth (OD) is fundamental to the DIAL principle, affecting signal-to-noise ratio and factoring into the δN calculation (Wulfmeyer & Bosenberg 1998). The OD between range r and $r+\Delta r$ can be calculated in a very straightforward manner from the DIAL signals:

$$OD = \frac{1}{2} \ln \left(\frac{P(\lambda_{on}, r)}{P(\lambda_{on}, r+\Delta r)} \frac{P(\lambda_{off}, r+\Delta r)}{P(\lambda_{off}, r)} \right), \tag{3}$$

assuming as in Eq. 1 that online and offline wavelengths are spectrally close enough to have identical transmission except for
200 the absorption due to WV. The one-way OD at the online wavelength increases along the beam path until the signal becomes too attenuated for accurate measurement. The maximum WV OD up to which an accurate retrieval can be made is dependent on many system parameters such as the detection noise floor and pulse energy, and in practice is typically in the range $1 - 2$, and for a single range cell $0.01 - 0.05$ (Bosenberg 1998, Wulfmeyer & Bosenberg 1998, Remsberg and Gordley 1978). To meet these criteria, flexibility in selection of online wavelengths and Δr is incorporated into the HALO WV retrieval.

205 **3.2 HALO water vapor DIAL**

HALO transmits four wavelengths around 935 nm to cover the large WV dynamic range from the moist surface to the dry UT/LS. These four wavelengths are spread between varying strength absorption features and wings of lines in the 935.5 nm line complex as shown in Fig. 1a, thus providing sensitivity across the wide dynamic range of WV within the troposphere and UT/LS. We refer to the most strongly absorbed HALO wavelength as λ_1, progressing sequentially to λ_4 as the least-
210 absorbed wavelength. The strong absorption of λ_1 makes it an ideal online wavelength in very dry UT/LS, suitable for ~8-20 km altitude range and therefore compatible with high-altitude aircraft such as the ER-2. λ_2 follows as the typical online wavelength for the mid-troposphere, and λ_3 is chosen for the lower troposphere and PBL. Due to the high variability of lower tropospheric moisture (e.g., tropics versus arctic), λ_3 is tunable within a 22 GHz range (0.064 nm) when referenced to λ_2. λ_4 is offset locked to λ_3 and thus tunes in tandem. The details of this tunability and transmitter design will be explored in a
215 subsequent instrument paper.

In selecting online and offline pairs for a given DIAL calculation, consecutive wavelengths are used, e.g., λ_2 online with λ_3 offline. This optimizes accuracy concerning the DIAL equation assumptions by minimizing spectral differences in transmission and minimizes the angular dependence of near-field returns through the narrowband optical filter (Nehrir et al.
220 2009). Figure 1b shows an example of nadir lidar signals at the four wavelengths, attenuating with increasing range from the

aircraft. The increase in signal below 1 km in Fig. 1b is due to enhanced aerosol backscatter within the PBL. Figure 1b also shows the low end of the dynamic range of these channels, with λ_1 and λ_2 reaching the noise floor around 10^0 counts. Each wavelength pair is sensitive to the altitude range that has sufficiently high differential absorption optical depth (DAOD) and sufficient online signal strength, which is demonstrated in Fig. 1c. Fig. 1c is a manually drawn estimate of the DAOD and

225 Poisson statistics (shot noise) contributions to the WV uncertainty, meant to illustrate the typical sensitivity of each wavelength pair. It should be noted that the plots of Fig. 1b and 1c will differ from profile to profile based on the vertical distribution and magnitude of the WV profile as well as the aerosol loading and scattering properties within the sampled volume. The precision of the WV DIAL retrieval is also highly dependent on instrument parameters and spatial averaging.

230 The HALO WV DIAL uses an injection seeded and frequency doubled Nd:YAG laser to pump an injection seeded optical parametric oscillator (OPO) to generate output at 935 nm with 1 KHz pulse repetition frequency (PRF). The 935 nm seed laser and OPO cavity are tuned to switch between the four wavelengths on a shot-by-shot basis. The residual 532 nm and 1064 nm pulsed energy left over from the OPO conversion process is transmitted collinearly with the 935 nm output to enable simultaneous aerosol/cloud profiling utilizing HSRL (532 nm) and backscatter (1064 nm) techniques, including

235 depolarization measurement at 532 and 1064 nm. The simultaneous WV and HSRL sampling approach employed with HALO is similar to previous work (e.g., Wirth et al. 2009) but the design to utilize a single transmitter for all wavelengths makes HALO a uniquely capable and compact instrument with the ability to support airborne campaigns from a wide range of aircraft optimized for different sampling strategies.

240 Real-time onboard processing of the 1 KHz PRF signal sums shots at each wavelength to improve the signal-to-noise ratio (SNR) and reduce the data rate to 2 Hz. The data system sampling rate is 120 MHz or 1.25 m in range, but the 935 and 1064 nm channels are limited by the detection chain electrical bandwidth to 15 m vertical resolution. The high vertical resolution allows for oversampling of surface or cloud signals to allow for high spatial resolution total and partial WV columns and future cloud retrievals, respectively. To keep the file sizes manageable, the atmospheric signals for all channels are filtered

245 and downsampled to 15 m vertical resolution. Subsequent temporal and vertical averaging of the WV DIAL data are employed to improve the precision of the DIAL retrieval, and this is discussed in further detail in Section 4. The 532 nm HSRL detector chain has sufficient bandwidth to maintain 1.25 m vertical resolution, however, the vertical resolution is limited by the 532 nm laser pulse width to ~3 m. The high vertical resolution will be used for future cloud and ocean profiling measurements, but the atmospheric data shown in this manuscript are digitally filtered and binned to 15 m vertical

250 resolution in post-processing to increase SNR and match the WV DIAL and 1064 nm data resolution. Working with 15 m resolution for calculating the NSF and all subsequent data products also ensures that data from each vertical bin is uncorrelated (any range correlations are small enough to be negligible.)

The receiver dynamic range is extended by implementing a high and low optical split. Dynamic range is further increased as both optical channels have an adjustable dual-gain (high and low) electronic output. The high-optical high-electrical gain channel is used for all atmospheric measurements in this manuscript and will be referred to simply as the "high gain" channel. The low-optical high-electrical gain channel is used in surface return calculations and will be referred to as the "low gain" channel. The other channels are used for diagnostics and will be optimized for measurements in future campaigns.

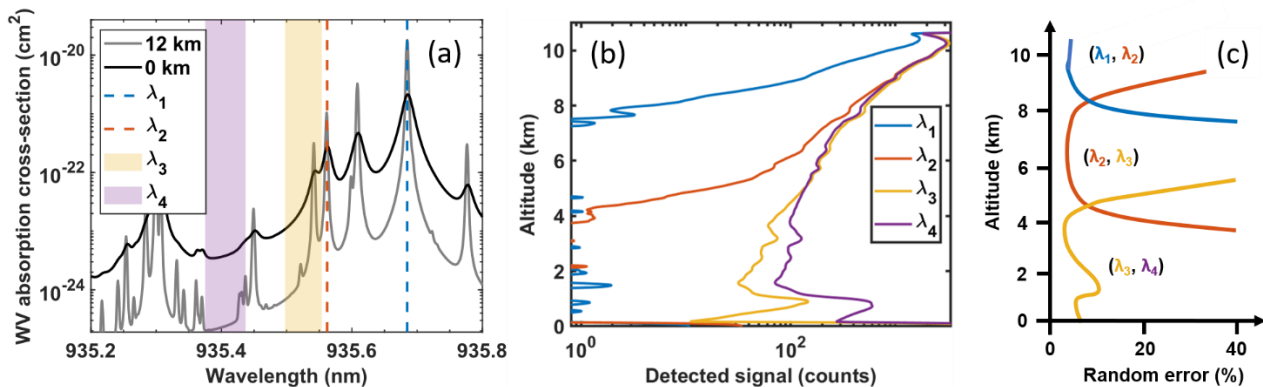



Figure 1. (a) Voigt spectrum of WV absorption cross-section at 12 km (210 hPa, 22K) and at sea level (1013 hPa, 300K), with vertical lines and shading indicating the fixed and tunable HALO wavelengths. (b) Example profile of HALO measured signals for the four transmitted wavelengths, averaged temporally and vertically to match a typical final WV product resolution (60 seconds temporal, 315 m vertical). (c) Sensitivity of line pairs to different parts of the atmosphere, expressed as an estimation of random error versus altitude. Line pairs are in parentheses as (online, offline).

Some key parameters of the HALO instrument and processing are listed in Table 1 and will be discussed in greater detail below. Another relatively brief summary of HALO WV DIAL was presented in the Aeolus cal/val campaign overview by Bedka et al. (2021), and further technical details will be presented in an instrument paper.

Table 1. HALO parameters.

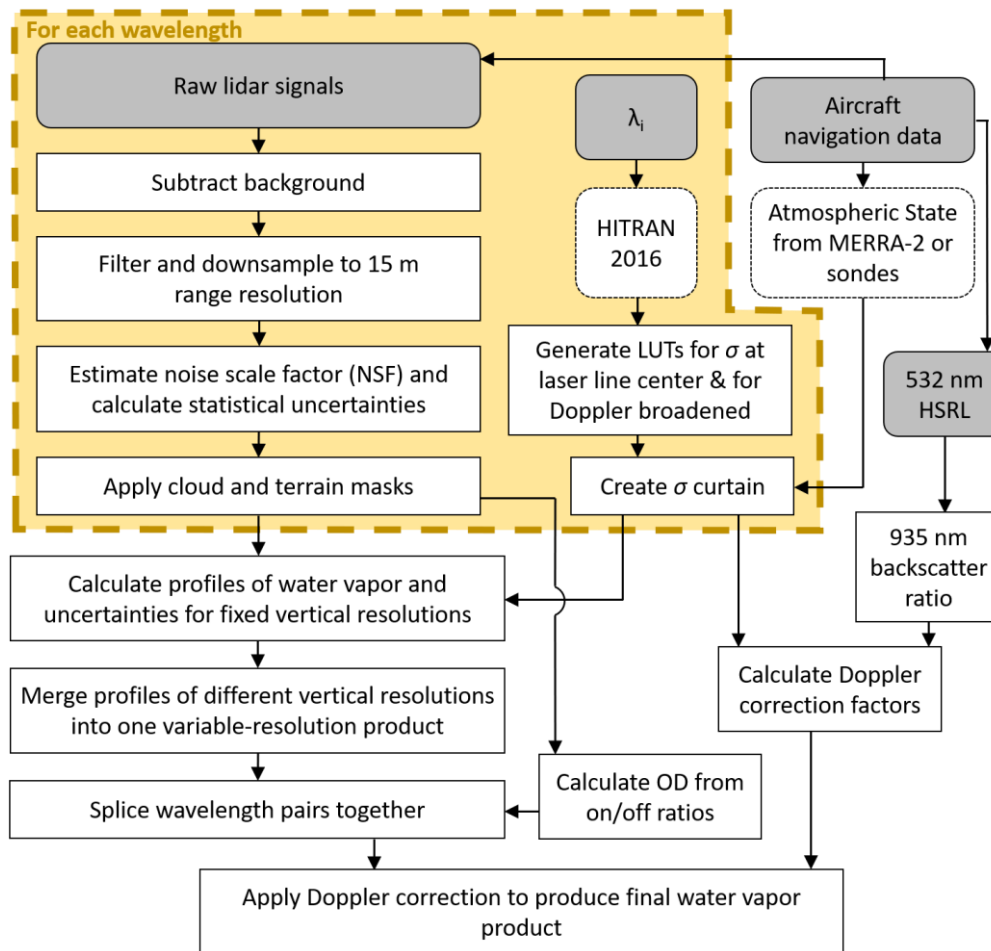
Transmitter type	Custom Fibertek Nd:YAG pumped OPO
Wavelengths	532, 935, 1064 nm
Pulse energy: 532, 935, 1064 nm	6, 1.5, 6 mJ
Pulse width: 532, 935, 1064 nm	18, 11, 22 ns
PRF	1000 Hz (effectively 250 Hz for each DIAL λ)
Average power: 532, 935, 1064 nm	6, 1.5, 6 W
Measurement principal: 532, 935, 1064 nm	HSRL, DIAL, backscatter
Detector type: 532, 935, 1064 nm	PMT, APD, APD
Telescope Diameter	40 cm
Receiver FOV: 532, 935, 1064 nm	1000, 300, 1000 µrad
DIAL temporal resolution *	5 – 60 s **
DIAL vertical resolution *	250 – 585 m
DIAL measurement dynamic range	0.001 – 25 g/kg

*Resolution is variable, with increasing statistical uncertainty accompanying finer resolutions, as discussed in Section 4A.
**For typical NASA DC-8 flight speed, including the data this paper, this is roughly 1 – 12 km.

4 HALO water vapor retrieval methodology

285 The HALO DIAL data processing and WV calculation is based on the preceding decades of DIAL research within the community and implements heritage techniques developed for HALO's predecessor LASE (e.g., Ismail & Browell 1989, Moore et al. 1997). This section gives an overview of the HALO WV retrieval, with additional details in subsections where warranted.

290 Two components of the DIAL equation (Eq. 1) require extensive consideration: the lidar signals and the absorption cross-sections. Their treatment is outlined in Fig. 2. First, electronic and atmospheric background signals are removed from the raw lidar signals before digitally filtering and downsampling to 15 m range resolution. Data below cloud top or terrain are then removed based on cloud-top heights (CTHs) identified from HSRL and terrain elevation from the Global Land One-Kilometer Base Elevation (GLOBE, Hastings & Dunbar 1999) digital elevation model, respectively. An additional 45 m

295 back-off is added to both cloud and terrain to account for any spatial heterogeneity within the observation time. High-altitude cirrus clouds are not automatically masked because they are often thin enough to be penetrated. Manual inspection of final datasets is employed for any additional masking that may be necessary.

 Figure 2. Flowchart of the HALO WV retrieval steps. Gray boxes are data or user inputs.

Calculation of absorption cross-section profiles with proper consideration of collisional (pressure) and Doppler broadening is critical for accurate DIAL measurements. This begins with the creation of lookup tables (LUTs) of WV absorption cross-section for each transmitted HALO wavelength over the reasonable expected range of pressures and temperatures. These LUTs utilize the 2016 High Resolution Transmission database (HITRAN, Gordon et al. 2017). The atmospheric state for each profile is taken from the MERRA-2 reanalysis model (Gelaro et al. 2017), or radiosondes/dropsondes when MERRA-2 is not yet available during field campaign operation. The atmospheric state at each point gives the pressure and temperature for the LUT, and thus the absorption cross-section curtains are generated for the DIAL calculation. One consideration not shown in Fig. 2 is an additional step to address the contribution of WV to collisional broadening instead of assuming dry air (Ismail & Browell 1989). The contribution of this collisional self-broadening is first calculated with WV mixing ratios from MERRA-2, then the resultant HALO WV mixing ratio product is fed back into the absorption cross-section curtain creation to fine-tune the contribution of self-broadening. The subsequent code is then rerun to produce the final HALO WV product.

Once the lidar signals and $\Delta\sigma$ have been implemented in Eq. 1 to calculate WV number density, the collocated MERRA-2
315 molecular number density is used to convert the HALO backscattered power measurements to WV mass mixing ratio. This
is done for each of the three wavelength pairs and at multiple range resolutions, and those resolutions are then combined into
a single curtain of WV measurements for each wavelength pair (Section 4A). Wavelength pairs are then spliced together for
each profile based on optical depths (Section 4B), and surface return measurements are appended below the minimum
altitude of atmospheric calculations (Section 4C). Lastly, Doppler correction factors are applied to produce the final WV
320 curtain (Section 4D).

Parallel to the WV processing is the calculation of statistical uncertainty in the WV product. The uncertainties calculated for
each wavelength pair and range resolution are merged into one final curtain in the same manner as the WV mixing ratios.
This uncertainty is the statistical error based on Poisson statistics (e.g., Eq. 9 of Ismail & Browell 1989) adapted to the
325 analog detection used by HALO via NSFs (Liu et al. 2006). A mean NSF is used for each channel in each campaign,
calculated from the NSFs determined from each flight. Because relevant instrument parameters are kept constant to the
greatest extent possible, the NSFs are fairly constant over the campaign, with standard deviations from the campaign mean
<15%. Additional sources of systematic uncertainty such as errors in HITRAN (Birk et al. 2017, Hodges et al. 2008),
knowledge of atmospheric state (namely temperature), or propagating errors from HSRL products used in Doppler correction
330 are estimated to have <1% impact on the WV product and thus are not accounted for here. Additionally, the magnitude of
systematic errors resulting from uncertainties in the performance of the instrument such as knowledge of the transmitted
wavelength, spectral purity, or linearity of the receiver detector chain are estimated to be <2%, far below the magnitude of
the random error resulting from detector electronic noise and shot noise. These sources of systematic error account for a
larger fraction of the error budget in drier environments such as in the UT/LS, however, the comparison with DLH
335 demonstrates that systematic sources of error are to a large extent well understood and don't drive the overall error budget of
the retrieval.

4.1 Variable resolution

As discussed in Section 3.1, statistical uncertainty in a DIAL measurement can be reduced by using a coarser temporal
(horizontal) or range (vertical) resolution, though this may introduce error in representativeness of gradients smaller than the
340 resolution. This allows for flexibility in processing to optimize for a given scientific objective. HALO WV processing is
typically done at a fixed temporal resolution with vertical resolution that varies between two values, because the statistical
uncertainty of the DIAL measurement is more sensitive to the vertical resolution. For most of the HALO data shown in this
manuscript, the initial DIAL calculation vertical resolution is 315 m. If the statistical uncertainty for a given mixing ratio
value at 315 m resolution exceeds 6%, the value is replaced with the corresponding 585 m vertical resolution value. These
345 vertical resolutions are user-defined and can be made finer or coarser when reprocessing the data, as can the conservative

choice of 6% uncertainty threshold. 6% was chosen for this dataset by empirical investigation to provide precise measurements even in the most challenging environments. A linear weighting function is applied to a 165 m vertical window centered on each bin where the range resolution changes to ensure a smooth transition in the WV profile.

350   To illustrate how the chosen resolutions for the DIAL calculation enable applicability to science targets across scales, Fig. 3 explores the WV product and its statistical uncertainty for temporal resolutions ranging from 1 second to 70 seconds with fixed vertical resolution, either 315 m or 585 m. This temporal range corresponds to 200 m to 14 km along-track for the typical DC-8 flight speed in the data shown here but would differ when HALO is deployed on other aircraft. Figure 3c and 3d exemplify Eq. 2, that statistical uncertainty for a given measurement decreases proportional to $\Delta t^{0.5}$ and that the statistical

355   uncertainty is more sensitive to Δr than Δt because a larger Δr results in a larger DAOD and the precision of the WV retrieval is directly proportional to the DOAD. It should be noted that Eq. 2 holds for any given profile, but the appearance of plots such as those in Fig. 3 will vary depending on the scene. For example, this profile shows a jump in uncertainty below ~4.5 km, where the λ_3 / λ_4 pair is used, because λ_3 for this flight was optimized for wetter environments than were observed at that time. Since this paper is not investigating specific targets, a conservative choice of 60 s temporal resolution

360   has been implemented in the retrieval, except where stated otherwise.

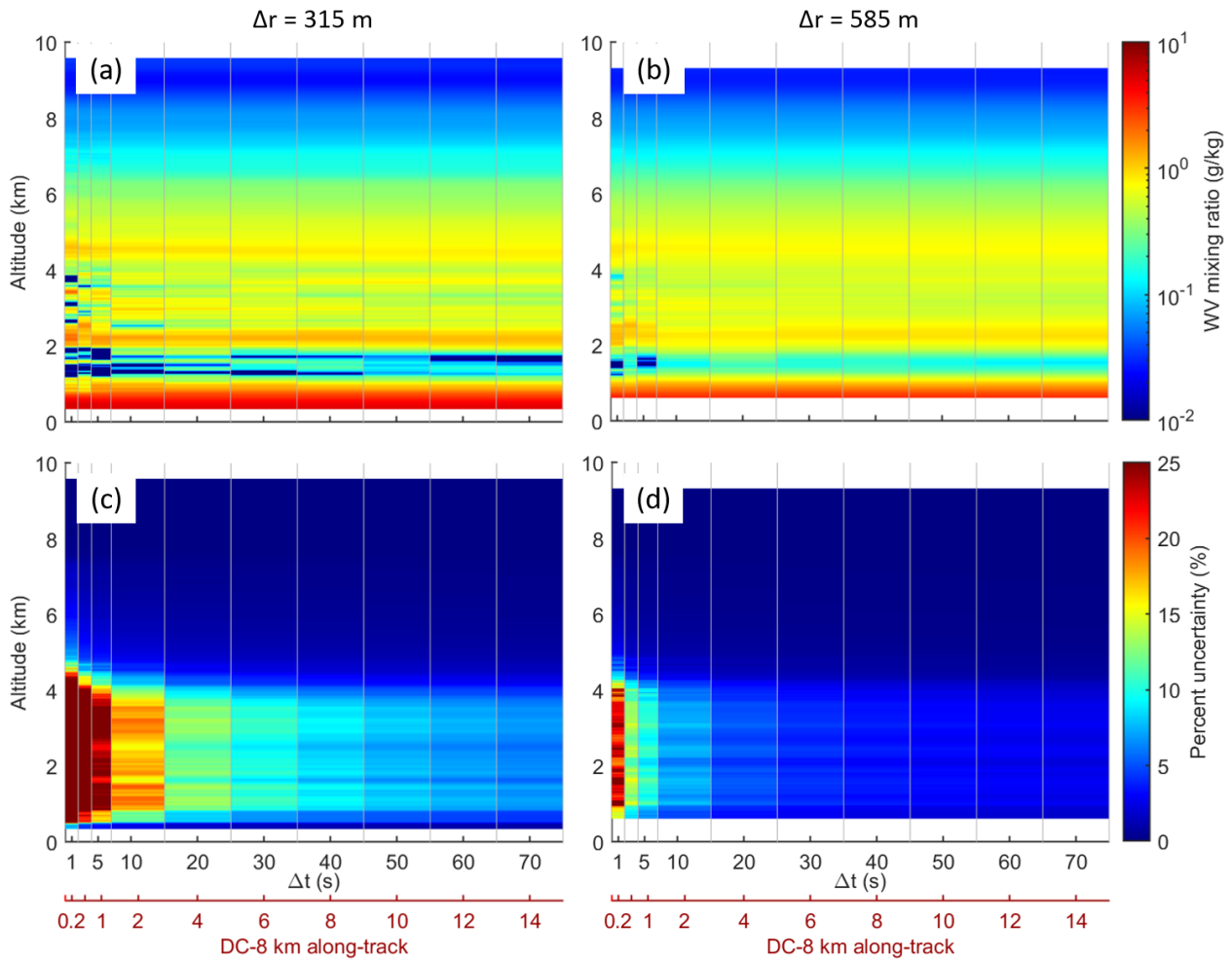



Figure 3. Plots showing (a, b) WV mixing ratio and (c, d) the associated statistical uncertainty for different temporal and range resolutions for a single profile centered at 25.04 UTC from the 2019-04-29 flight (this profile is also shown in Fig. 4). (a, c) have Δr = 315 m throughout the profile while (b, d) have Δr = 585 m. The x-axis indicates different temporal resolutions (Δt), which translate to along-track horizontal resolution calculated for the typical DC-8 flight speed.

Figure 4 shows HALO data for a section of a midlatitude flight. Figure 4a is the final WV mixing ratio product, with corresponding statistical uncertainty and range resolution shown in Fig. 4b and 4c. Various environmental conditions and features were sampled, including broken and unbroken marine stratocumulus, a very dry layer above the PBL, and moist layers throughout the troposphere that in one location extended towards the tropopause (visible and infrared satellite imagery for this flight was shown in Bedka et al. 2021). In Fig. 4b, many of the sharper gradients in uncertainty are a result of a change in vertical resolution (compare to Fig. 4c). The largest uncertainties for this scene were driven by insufficient DAOD in the dry layer above cloud-top. Because the layer was very dry, the small DAOD was difficult to measure with the weakly absorbed λ_3. Aerosol scattering in the PBL provided good SNR and thus low uncertainties. Aircraft altitude also indirectly

affected uncertainty by effectively shifting the SNR profile; this can be seen around 0.4 – 0.9 UTC. Figure 4d exhibits the utility of the simultaneous HSRL measurements to identify clouds and the PBL while providing valuable information for WV-aerosol-cloud interaction studies.

4.2 Splicing profiles from multiple wavelength pairs

380 After the variable range resolution has been determined for each wavelength pair, the WV profiles of the three wavelength pairs are spliced together based on WV DAOD thresholds. The extent of each splicing region is plotted in Fig. 4c as an example. Starting from the highest altitude, WV calculated from the first wavelength pair (λ_1 and λ_2) is used until a DAOD of 1.0 is reached, at which point a linearly weighted average incorporates an increasing contribution from the second wavelength pair until an OD of 1.6. The second wavelength pair is then used alone until the OD range 1.0 to 1.5, wherein

385 again a linearly weighted average controls the transition to the third wavelength pair. The third wavelength pair is used alone for the rest of the profile. The OD thresholds can change from these default values based on manual inspection of instrument performance and atmospheric scene. Results are typically not affected by threshold changes within ±0.1 of chosen values because the overlap region where both wavelength pairs perform well is sufficiently broad.

390 Fig. 4 illustrates a fundamental challenge of the DIAL technique with the very dry layer below 2 km. Because the more absorbing wavelengths are attenuated close to the instrument, only the less sensitive λ_3 has useable signal at this altitude. The very low DAOD gives accurate identification of the dry layer but low precision. During operation the HALO λ_3 signal is monitored and the wavelength is tuned as needed to ensure maximum DOAD in conjunction with the signal staying above the instrument noise floor at all altitudes. This effectively maintains maximum possible sensitivity to any dry layer that may

395 be present while still capturing the full profile to the surface.

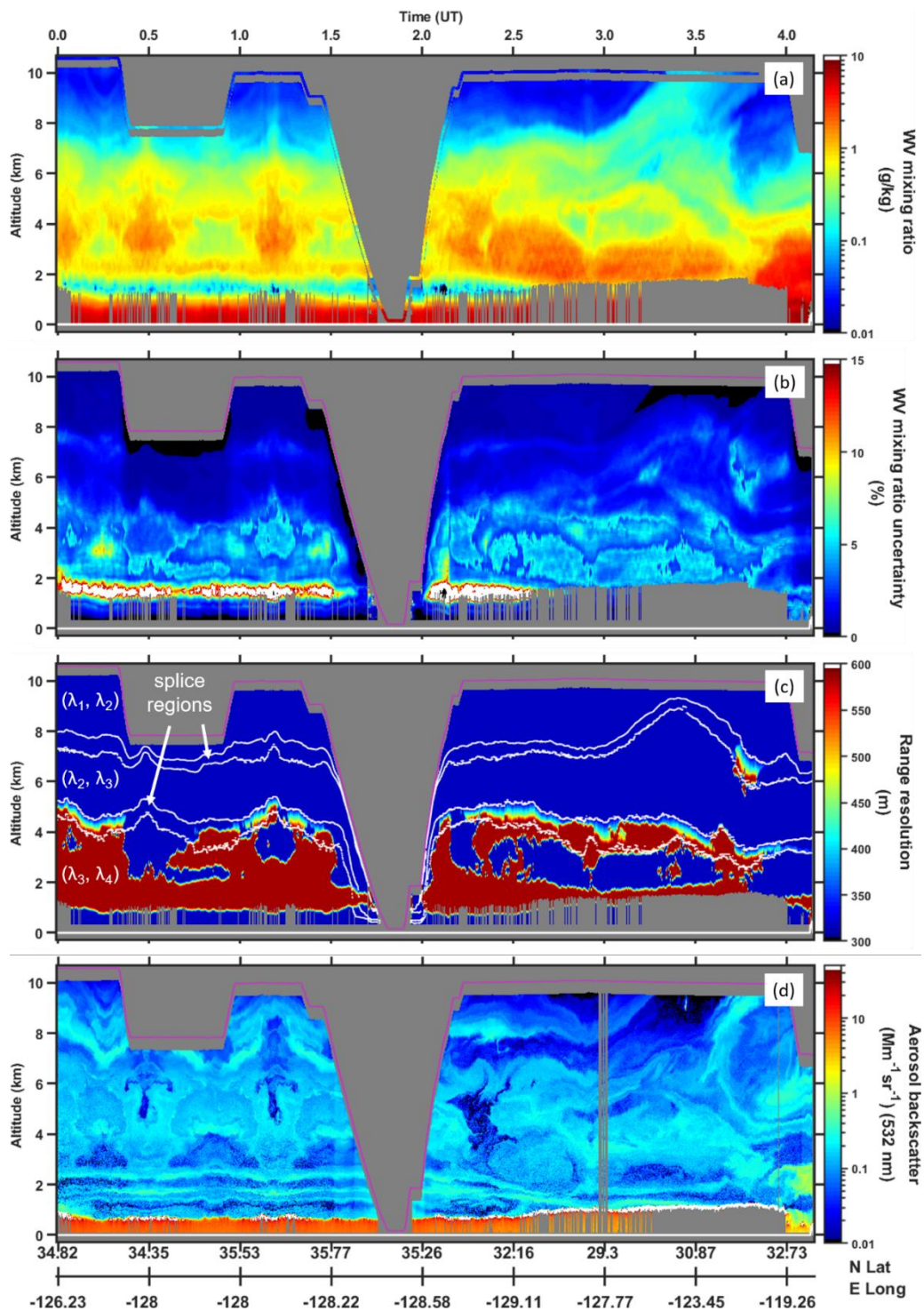



Figure 4. Time-altitude curtains of HALO (a) WV mixing ratio, (b) percent uncertainty, (c) vertical resolution, (d) HSRL 532 nm aerosol backscatter on 30 April. DLH measurements are also shown in (a) at the aircraft altitude, whereas aircraft altitude is a magenta line in the other panels. The white lines in (c) mark wavelength pair and splicing regions, as indicated.

4.3 Near-surface water vapor measurement via surface return signals

The center of a nadir-pointing atmospheric DIAL measurement window (i.e., the altitude at which the value is reported) is by its nature limited to a distance of Δr, the range resolution, above a hard target scattering surface (e.g., land, ocean, or cloud top). As discussed above, Δr is driven by the measurement SNR and desired precision. However, the unresolved lowest bin (from the surface up to an altitude of Δr) can be retrieved by carrying out a DIAL retrieval using the strong surface return and the last atmospheric bin above the surface. This is accomplished by employing Eq. 1 with λ_3 and λ_4 where $P(\lambda, r + \Delta r)$ is replaced with $P(\lambda, r_{surface})$.

These signals are depicted in Fig. 5, utilizing high gain and low gain signals for $P(\lambda, r)$ and $P(\lambda, r_{surface})$, respectively. The high gain atmospheric signals are temporally and vertically averaged matching the atmospheric DIAL resolutions from Fig. 4 of 60 s and 315 m. Due to the increase in SNR from aerosol scattering in the PBL, future work may reduce this averaging to a finer resolution. Only atmospheric signals more than about 100 m above the surface are considered to ensure no contamination from surface heterogeneity or other effects such as sea spray. The low gain channels have an optimal receiver dynamic range for the surface return signals, ensuring that the strong signal is captured within the linear regime of the channel's digitizer (i.e., neither saturating nor reaching the noise floor.) The low gain channel consistently best captured the surface return signal over ocean during this campaign, but this channel choice may differ for other datasets, e.g., over ice or flying at lower altitudes. To ensure accurate representation of the energy capture within the impulse response from the surface return, the data from 5 bins centered about the peak of the surface return are integrated (highlighted with color in Fig. 5). The online and offline surface return signals were ratioed then temporally smoothed to 60s to match the atmospheric data. This order of operations was empirically determined as the best approach for reducing non-physical outliers. WV retrieved in this manner is included in the HALO profiles throughout this paper (except Fig. 4) as well as the entire campaign curtains in Bedka et al. (2021).

While this method was previously employed with LASE and is similar to other techniques (e.g., integrated path differential absorption lidar (Abshire et al. 2013, Amediek et al. 2017, Dobler et al. 2013)), to the authors' knowledge it has not been previously published with application to range-resolved WV DIAL profiling. This method has only been applied to HALO for clear-sky data over oceans thus far because it is a relatively uniform surface compared to cloud or land. Moving forward, efforts will be made to adapt the surface return methodology to extend HALO measurements down to those more complex surfaces. Topographic and albedo variability of cloud and land will require a more detailed treatment to ensure surface-related changes in signal are separated from atmospheric OD variation, and may ultimately prohibit usefulness of the retrieval over sufficiently complex scenes. The HALO low gain channels will be optimized to keep the signals over clouds and land on scale, something that was not the focus of the maiden HALO flights. Furthermore, the surface return will be examined with the native HALO 1.25 m resolution, which despite being oversampled still allows for much more accurate

peak finding and representation of the energy captured within the surface impulse response. This latter topic is an area of ongoing research and will be the subject of a follow-on study.

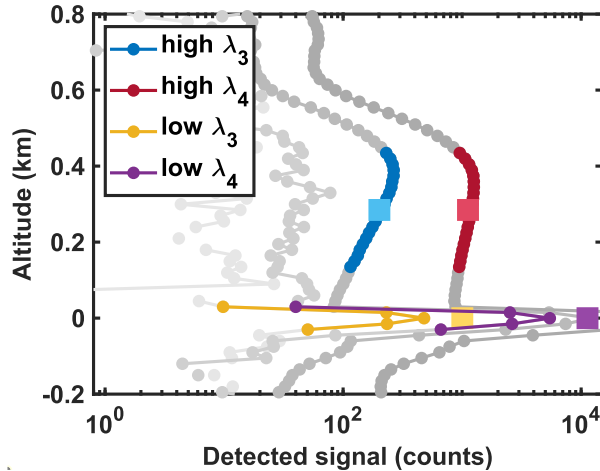



Figure 5. Illustration of surface return DIAL concepts. The legend "high" and "low" refer to high gain and low gain channels, plotted at 15 m vertical resolution. The high gain data is temporally averaged to 60 sec, while low gain is a single half second averaged record. The colored circles are averaged (atmospheric) or summed (surface return) to use in the DIAL calculation. These are the $P(\lambda, r)$ inputs to Eq. 1 and are shown as colored squares on the plot. Grayscale points are not used in the surface return but are still shown here for context.

4.4 Doppler broadening correction

Another consideration in optimizing the HALO WV retrieval is accounting for the Doppler broadening of the backscattered signal. The backscattered light from aerosols experiences negligibly small Doppler shifts (MHz), whereas the backscattered light from molecules is significantly Doppler-broadened (GHz) because of the temperature-dependent molecular velocity distribution. The Doppler-broadened return signal must be considered carefully due to the wavelength-dependent OD of the return path. In particular, the signal from the start of a range cell and the end of that range cell can experience different aerosol ODs based on the aerosol gradient within that Δr. Thus, the aerosol backscatter ratio profile is necessary for correct consideration of Doppler broadening effects, and in this regard the HALO HSRL measurements are a unique advantage of the HALO architecture. A full description and mathematical framework of the Doppler broadening correction for WV DIAL measurements is presented in Ansmann (1985) and Ansmann and Bosenberg (1987) (with more recent studies including Fan et al. 2015, Spath et al. 2020). The correction's application to HALO was streamlined by experience with HALO's predecessor LASE (Ismail & Browell 1989).

Doppler broadening correction is implemented in the HALO retrievals beginning with creation of 2 LUTs for absorption cross sections with the transmitted wavelength, pressure, and temperature as user inputs to the Voigt function which utilizes

the 2016 HITRAN database for fundamental line parameters. The first LUT is the same as the single-wavelength lookup tables described above and is used for the aerosol portion of the backscatter (Mie scattering). A second LUT is comprised of the Doppler (Gaussian) weighted sum of the Voigt absorption cross sections across a range of +/- 0.05 wavenumbers from each transmitted laser wavelength (Rayleigh scattering). The backscatter Doppler broadening for the profile is then determined from the atmospheric state (MERRA-2 or sondes), and the aerosol backscatter profile at 935 nm is estimated from the backscatter angstrom exponent at 532 nm / 1064 nm (Burton et al. 2012). This angstrom exponent comes from the HSRL 532 nm aerosol backscatter and the 1064 nm aerosol backscatter retrieved using Equation 10 of Hair et al. (2008). If there are brief periods when DIAL data is available but HSRL is not (e.g., 3 UTC in Fig. 4), the last available HSRL profile is used. The propagation of the broadened signal is then calculated (e.g., Ansmann & Bosenberg 1987) and thus the Doppler corrected WV profile is determined. This processing is done parallel to the non-corrected WV, and ultimately applied to the WV product as a correction factor. This allows optional application of the correction in the field where processing is abbreviated, as well as examination of the magnitudes of the Doppler correction. Throughout the Aeolus cal/val campaign, the Doppler correction to the WV product only had an impact of 3.5% at most.

5 Data product comparisons and discussion

Sondes were deployed during the Aeolus cal/val campaign to validate the wind measurements by Aeolus and DAWN. The sondes also opportunistically provided WV measurements for comparison to HALO WV DIAL. However, due to the poor performance of the sonde WV measurements, only a very limited qualitative comparison is possible (Bedka et al. 2021). This section therefore places emphasis on HALO WV comparisons to the other observations that were available, namely the DLH and satellites, which proved useful despite lacking the profiling spatiotemporal resolution to constitute a thorough validation dataset.

5.1 Dropsondes

A representative set of WV profiles from sondes and HALO is shown in Fig. 6, taken throughout the campaign including 4 out of the 5 flights, spanning the tropics to relatively dry midlatitudes. HALO profiles with 30 s and 60 s temporal resolution are shown to further exemplify the resolution flexibility discussed above. Vertical resolution is the same for both, using the default algorithm for variable vertical resolution applied to the 60 s dataset. Both resolutions have very good agreement with the sondes and with each other, as expected from Fig. 3. However, as mentioned in Section 2C only a qualitative comparison to dropsondes is possible from this campaign because the dropsonde moisture sensors had a nonlinear damped response and erroneously low values in the first few km of descent. This damped response is seen as a vertical shift, i.e., for a given feature in the HALO data, the same feature is in the sonde data but its location is shifted downwards by a variable amount (this is especially clear in Fig. 6b, 6c). This shift starts large then diminishes in an inconsistent way, making a correction infeasible and limiting to qualitative comparison.

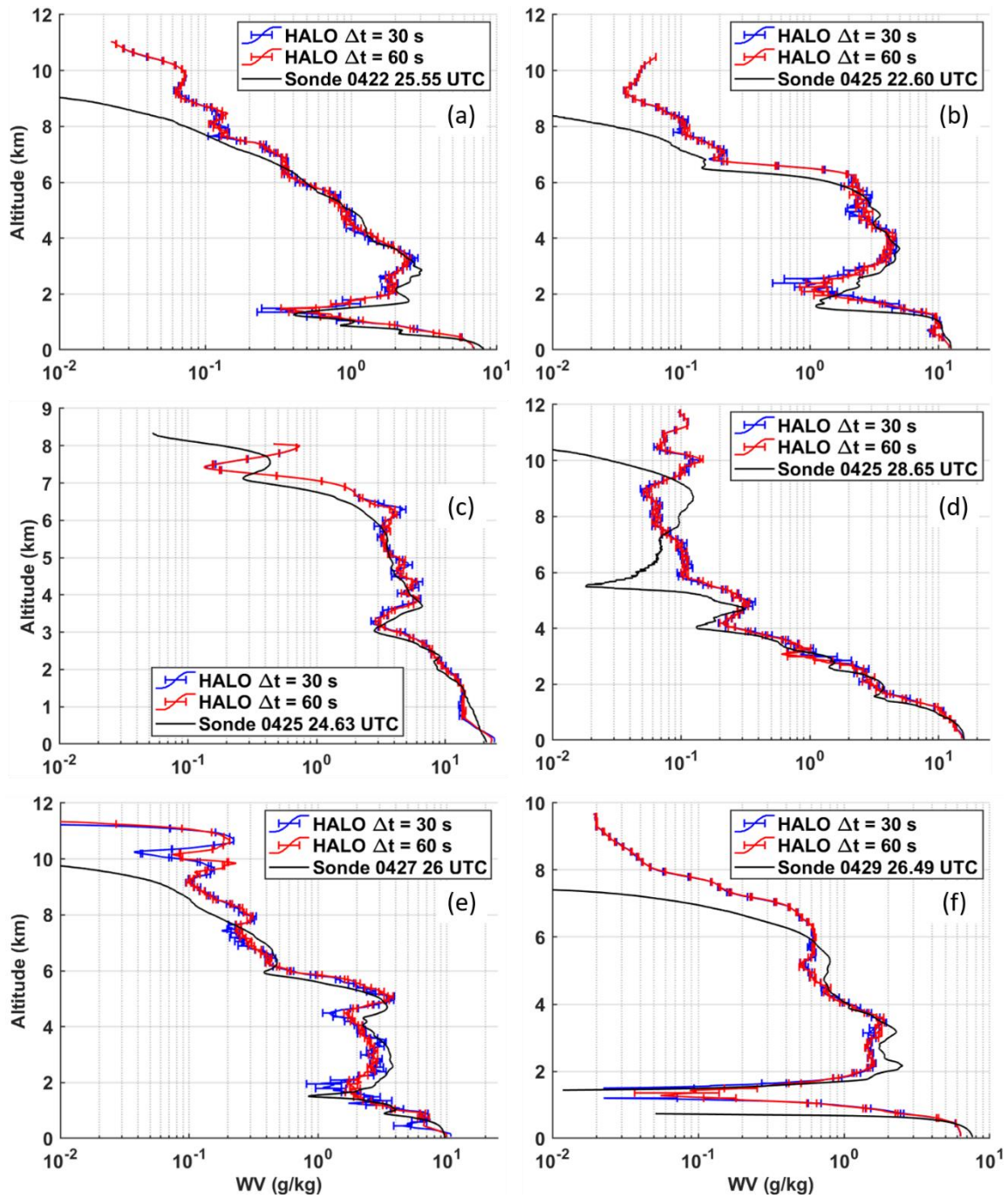



Figure 6. HALO WV mixing ratio profiles with statistical uncertainty bars, and collocated dropsonde profiles for comparison. The bars have been decimated for legibility. The sonde WV sensors had erroneous nonlinear damped responses and thus cannot be used for a rigorous quantitative comparison.

In the middle and lower troposphere, the dropsondes and HALO resolved the same moisture structures with similar results across two orders of magnitude, approximately 0.1 to 20 g/kg. Dry layers within the lower troposphere were captured by both instruments. The lowest few hundred meters where the HALO measurements utilized the surface return signal over ocean also showed good agreement, with differences typically less than 15%. Across the wide range of aircraft altitudes and environmental conditions that were sampled, there were no apparent systematic biases or other deficiencies in HALO WV for the atmospheric signal profiles or the surface return retrieval.

5.2 DLH

The DLH onboard the DC-8 was the best available option for validation of a HALO WV profile. Only one spiral descent was carried out during the campaign, from 1.34 UTC to 1.82 UTC on 30 April (shown in Fig. 4), and was preceded by multiple overpasses. The DLH profile from 10 km down to 160 meters above the surface, with an average descent rate of ~1300 ft/min, is shown alongside HALO profiles in Fig. 7. DLH data is shown at native resolution and with 315 m vertical smoothing to match the HALO measurements. The DC-8 descent occupied 50 km^2 and the preceding overpasses were tangential to this area due to vectoring by air traffic control, so spatiotemporal differences in the measured profiles are inevitable and there is no single best HALO profile to compare to the DLH. Furthermore, the DC-8 in-situ spiral diameter was approximately a quarter of the length of the remote sensing overpass leg and substantial variability was observed by HALO within this volume, which could explain the high frequency variability in the DLH data around 4 km. The gray shaded area in Fig. 7 shows the range of values measured by HALO in the overpass preceding the spiral, from 0.93 to 1.06 UTC, which included 8 independent profiles. An example HALO profile taken near the start of the descent is also shown (1.51 UTC), with data above 7 km appended from a slightly earlier profile with a higher aircraft altitude (1.38 UTC). The vast majority of the DLH profile was within the range of values observed by HALO, with two very small discrepancies in the upper troposphere and another at PBL top, which were potentially caused by spatiotemporal variability between the measurements. The single HALO profile also shows good agreement, exemplifying the measurement precision across the wide dynamic range of the profile and even capturing the very dry layer above the PBL, which is particularly challenging from a range of several km.

A sonde profile is also shown in Fig. 7. The sonde was launched at 1.43 UTC, near the start of the spiral descent. The sonde is not in agreement with the DLH nor the range of HALO values until the lower half of the profile, and even then appears to have temporal lag as it exceeds the range of HALO values around the sharp gradients below 3 km. This reinforces the previous statements that the sondes were not reliable enough for quantitative validation of HALO WV.

We do not present error statistics for this one DLH profile because the dataset is too small, and differences are likely dominated by specific atmospheric features and sampling differences between the instruments. Many profiles would be needed (e.g., from DLH or functional sondes) to objectively assess precision or any systematic bias in the HALO profiles.

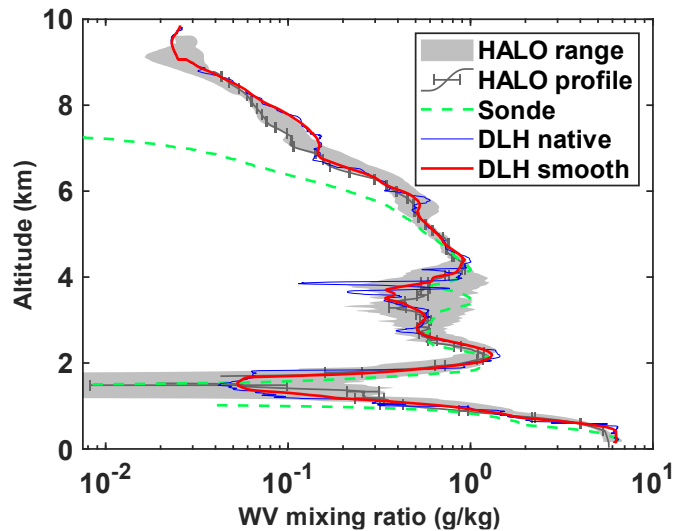



Figure 7. HALO comparisons to DLH and sonde from the descent on 04/30. The blue line is the DLH data at native resolution. The red line is DLH smoothed with a 315 m rolling average to match the constant 315 m vertical resolution of the HALO profile. The gray area shows the range of HALO measurements in the area of the spiral descent from the preceding overpass, indicating the expected spatiotemporal variability of the region. An example HALO profile from right before the descent is also shown, with bars denoting measurement statistical uncertainty. The dashed green line shows the malfunctioning sonde profile.

A much more statistically robust approach to comparing the DLH and HALO measurements is to utilize the near-field HALO measurements over the course of the campaign to DLH taken at flight altitude. The closest comparable HALO measurement was typically ~400 m below the aircraft due to geometric overlap of the transmitter and receiver. Figure 8 and Fig. 4a show DLH WV data from the four flights where it was operational, plotted at aircraft altitude with the HALO WV profiles below. The continuity of features between the DLH and HALO data in these plots confirms HALO's ability to capture UT/LS WV, plus the relatively moist environments of the intertropical convergence zone (ITCZ) troposphere in Fig. 8c and the spiral descent in Fig. 4a. This agreement is quantified with the scatterplot in Fig. 9 and associated statistics in Table 2. Comparing the 86300 available DLH measurements with the nearest measurements in the coincident HALO profiles, ranging 0.004 g/kg to 2 g/kg, yielded an $R^2 = 0.66$ and a very small wet bias of 0.003 g/kg. A small wet bias could be expected due to the typical increase of WV with decreasing altitude, plus the range and distribution of the values in Fig. 9 suggest that this calculation may be dominated by the spatial variability over the ~400 m vertical separation between the measurements rather than indicative of a systematic bias in HALO. The low end of the comparison, $\leq 10^{-2}$ g/kg, most clearly shows a moist bias for HALO in Fig. 9, but this data is sourced entirely from 4-5 UTC on 23 April (Fig. 8b) where the aircraft was clipping the tropopause and large moisture gradients appear to be influencing the comparison. Removing this flight from the comparison drops the HALO wet bias by about one order of magnitude, to $6.8 \cdot 10^{-4}$ g/kg.

550 It should also be noted that due to the high-altitude nature of this comparison, 86% of the observations are with HALO wavelength pair λ_1/λ_2, and another 7% are in the splicing region of λ_1/λ_2 and λ_2/λ_3. When considering only the remaining 7% which are all pair λ_2/λ_3 and constitute most points >0.04 g/kg, bias is still very low at -0.005 g/kg. Overall, this good agreement of DLH and HALO WV with a large dataset spanning 3 orders of magnitude and a range of atmospheric conditions is a validation of HALO WV in the near-field, as well as the general measurement principle and implementation.

555 This is also a notable comparison as the UT/LS is a region rarely profiled with the accuracy, precision, and resolution of a lidar.

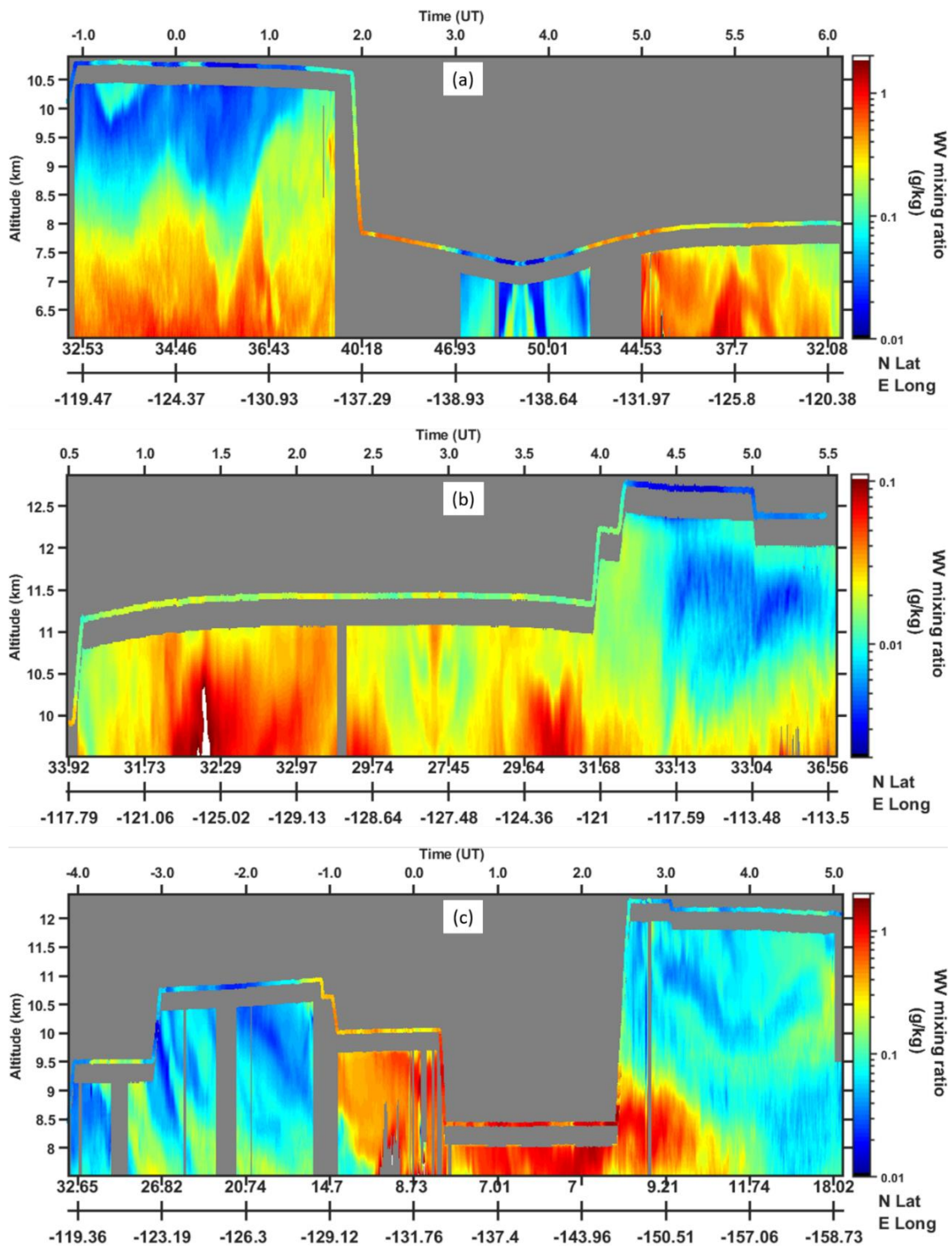



Figure 8. HALO WV curtains with DLH plotted at the aircraft altitude, as in Fig. 4a. Segments of three different flights are shown
560 **in (a, 18 April; b, 23 April) midlatitudes and (c, 26 April) the tropics. Only upper altitudes are shown to focus on comparison with DLH.**

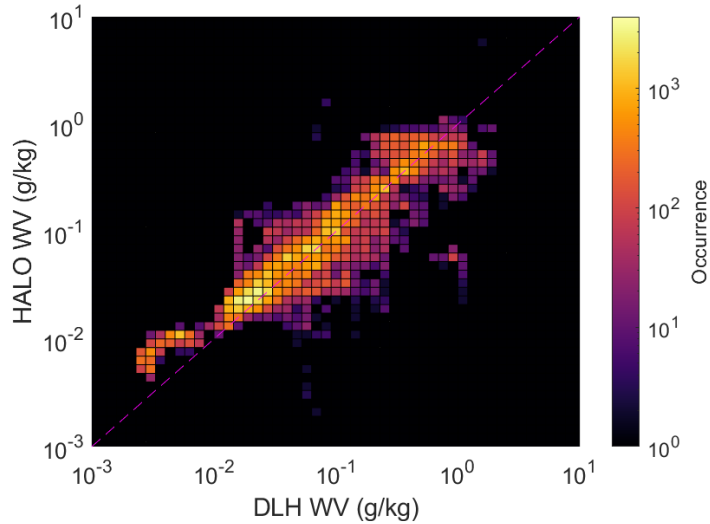



Figure 9. Plot of DLH versus HALO WV on log scales and aggregated over all available data, with color indicating the number of points in each bin. The dashed line is a one-to-one line. HALO data is the average of the 5 highest-altitude measurements, ~400 m below the DLH.

Table 2. Statistics of HALO comparisons with other instrumentation. DLH comparison is against the average of the closest 5 HALO measurements. AIRS and IASI comparisons are of PWV, as shown in Fig. 11. *n* is the number of data points in the comparison.

	n	Bias (HALO – x)	R^2
DLH	86300	0.003 g/kg	0.66
AIRS (ocean)	44	-2.55 mm	0.96
IASI (ocean)	81	-1.93 mm	0.98
IASI (land)	24	-4.07 mm	0.85

5.3 Satellites

5.3.1 PWV

Geophysical observables such as WV profiles and PWV from satellites provide another opportunity to assess HALO products against community standards (e.g., Chazette et al. 2014, Martins et al. 2010, Roman et al. 2016). During this campaign there were several under-flights of opportunity with AIRS and IASI that allowed comparisons to HALO. HALO PWV is calculated by vertical integration of a given WV profile. It should be noted for the HALO and spaceborne sounder comparisons that HALO only gives a partial column PWV, limited to altitudes between the aircraft and surface. HALO records were omitted from PWV comparison if a cloud was detected or the aircraft was below 8 km. AIRS overpasses were

580   suitable for comparison to HALO on 25 and 27 April, both passing through the northern edge of the ITCZ. IASI overpasses were suitable for comparison to HALO on 17, 22, 25, and 27 April, which included tropical and midlatitude flights. The temporal average and standard deviation of HALO measurements within 25 km (20 km) of the AIRS (IASI) footprint centers were used to mitigate the instrumental and spatiotemporal sampling differences between the instruments, similar to previous studies (e.g., Bedka et al. 2010, Chazette et al. 2014, Diao et al. 2013, Roman et al. 2016).

585

PWV from HALO, AIRS, and IASI are shown on maps in Fig. 10 for comparison. Overall agreement is good, capturing the large moisture gradient near the ITCZ as well as features in relatively dry midlatitudes. An important note for these comparisons is that some differences may arise from the brevity of the satellite overpass (minutes or less) versus the hours of DC-8 flight that fall within the satellite swath. The DC-8 location at overpass time is marked with a green circle in Fig. 10.

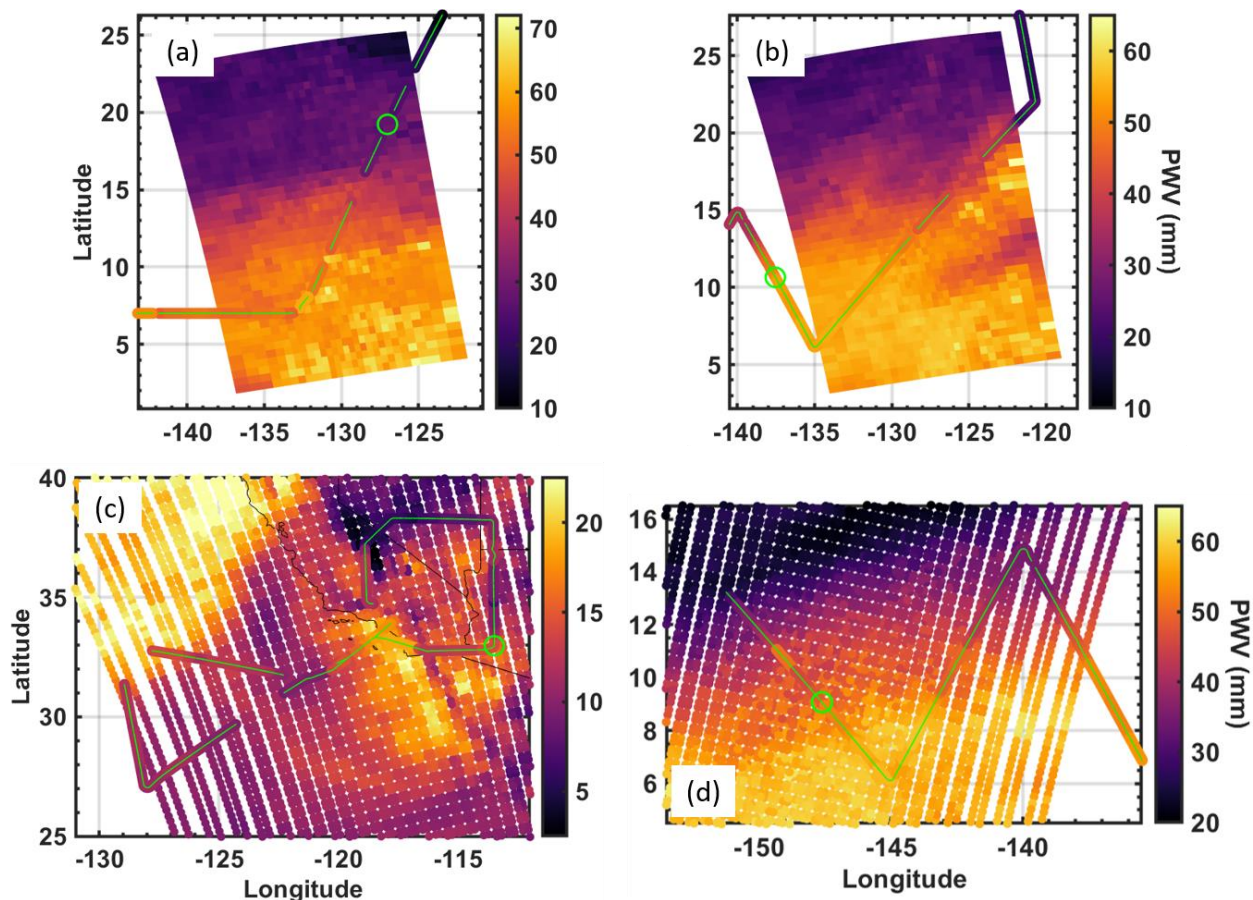
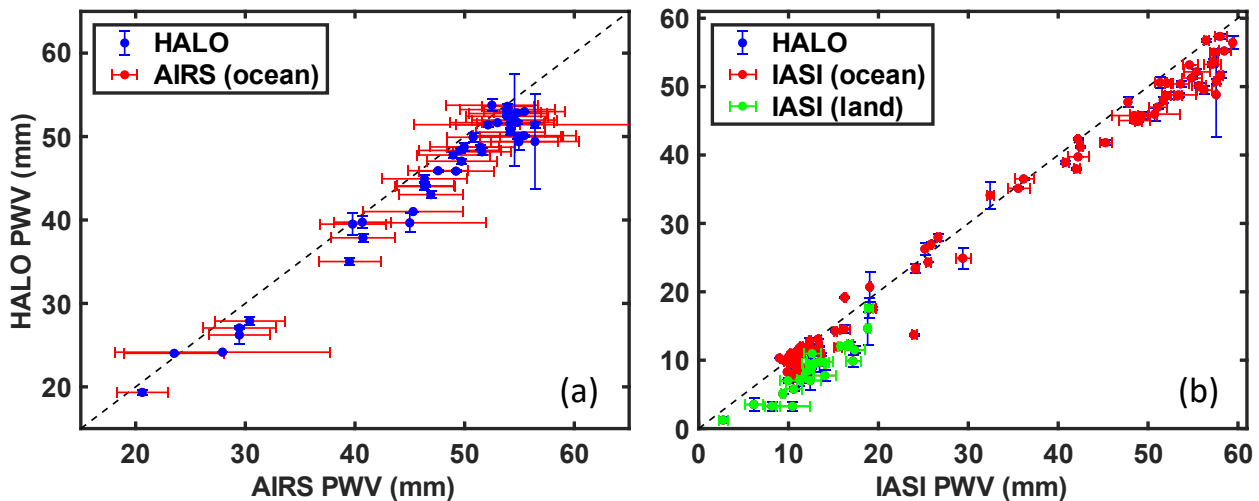
590   This spatiotemporal offset may be responsible for some differences, including the northwest corner of the oceanic portion of Fig. 10c, but cannot be corrected for in a straightforward manner.

595

Figure 11 and Table 2 show that there was excellent agreement of HALO PWV with AIRS and IASI, with HALO measurements over ocean having a dry bias of 2.55 mm and 1.93 mm and R^2 of 0.96 and 0.98 against AIRS and IASI,

600 respectively. A dry bias is expected for these comparisons because HALO is only capturing the partial atmospheric column, i.e., below the DC-8 flight altitude. The bias is most prevalent at high PWV values, which correspond to tropical flight legs where the DC-8 was often flying within the upper reaches of deep ITCZ moisture plumes, e.g. Fig. 8c. On midlatitude flights with low or moderate PWV and a relatively dry upper troposphere, such a bias was not clearly evident. The overall good agreement may also be interpreted as an indirect indicator of accuracy in the near-surface retrieval presented in Section 4.3,

605 since PWV is generally dominated by the high moisture content of the PBL (e.g., Richardson et al. 2021, Thompson et al. 2021), but this is not a direct relationship appropriate for robust validation.

Some of the IASI overpass time included data over land (e.g., Fig. 10c), wherein the surface return DIAL technique was not employed to measure the lowest few hundred meters of the atmosphere, and thus measured PWV is expected to have a larger

610 dry bias. The bias from this limitation is shown in Fig. 11 and Table 2 to be 2.14 mm larger than the PWV comparison over ocean which included the surface return DIAL technique. This discrepancy may be eliminated in the future as the surface return retrieval is applied to HALO data collected over land.

615 **Figure 11.** Comparison of (a) AIRS or (b) IASI PWV with the average HALO PWV calculated within each satellite retrieval footprint. Bars for AIRS are the reported retrieval uncertainties. Bars for HALO and IASI are standard deviations. Dashed line is 1:1, for reference.

5.3.2 Water vapor profiles

Figure 12 shows four HALO WV profiles alongside AIRS and IASI WV profile products. As noted in Section 2E, the spaceborne sounders report at finely spaced vertical levels but are actually limited to 1 – 2 km vertical resolution of independent information in the troposphere under ideal conditions. While the sounders generally agree with the trends of HALO, and some differences may be attributed to spatiotemporal changes, the spaceborne sounders are ultimately limited relative to HALO in their ability to resolve sharp gradients and small-scale variability in moisture that are important for transport and convection. This is especially apparent in these profiles around PBL top and lower tropospheric dry layers, where discrepancies of 1-2 g/kg or greater are present. The logarithmic plots also highlight large differences in the UT/LS.

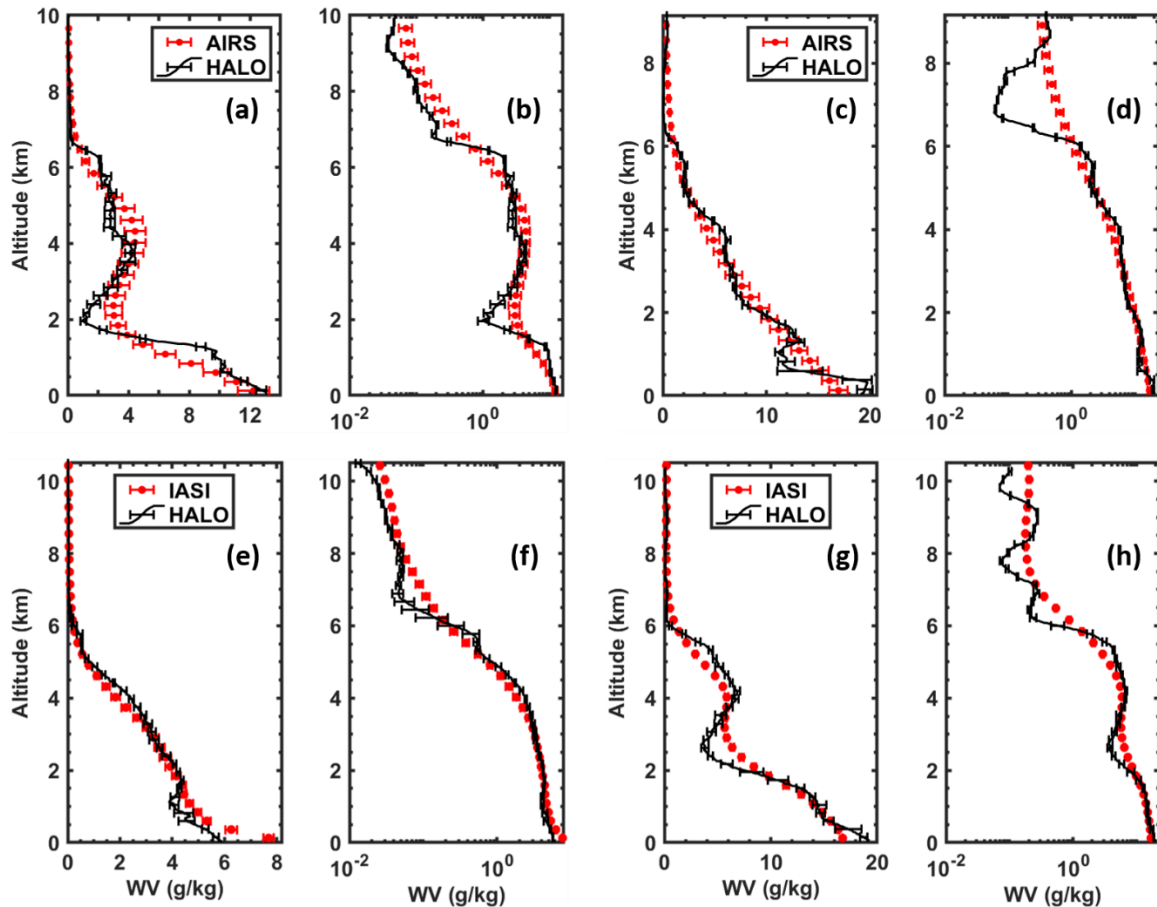



Figure 12. WV mixing ratio profiles from HALO versus (a-d) AIRS and (e-h) IASI with two locations each and both (a, c, e, g) linear and (b, d, f, h) log scales to emphasize capabilities across scales. The AIRS profiles are from 25 April 2215 UTC overpass, with the HALO profiles from (a, b) 2236 and (c, d) 2316 UTC. IASI profiles are from (e, f) April 23 0400 UTC overpass and 0430 HALO profile, and (g, h) 27 April 1926 UTC overpass and 1912 UTC and HALO profile.

AIRS and IASI provide important research and operational data to the community and have the strength of frequent global coverage, but HALO or similar active remote sensing is clearly advantageous for supporting process studies and other applications requiring higher spatial resolution and accuracy. The synergistic strengths of combining active and passive sounders will be an important resource moving forward, such as the work by Turner and Lohnert (2021) combining passive and active remote sensing observations to increase information content and improve vertical resolution and accuracy of passive retrievals.

6 Conclusions

This manuscript provides an overview of implementation and retrieval methodology for the new HALO airborne WV DIAL system and comparison to other instrumentation from its five maiden flights, serving to inform the community of a new capability within the suborbital portfolio. The HALO instrument and WV retrieval were designed based on decades of legacy of related DIAL efforts at NASA LaRC and in the global community. The HALO WV DIAL measurements are carried out in the 935 nm spectral range, transmitting four wavelengths to achieve sensitivity to moisture from the UT/LS down to the PBL within a single profile. HALO is the first WV DIAL system to employ shot-by-shot switching between the four wavelengths using a single laser transmitter, and in doing so reduces instrument size, complexity, and potential for certain systematic errors. The retrieval methodology incorporates the flexibility of the DIAL technique to trade resolution for precision, with streamlined reprocessing to optimize for scientific applications across scales. Another unique advantage of HALO is the combination of the WV DIAL with HSRL (or methane DIAL). The HSRL measurements provide cloud and aerosol optical properties, which are assets to both the WV retrieval calculations and scientific analysis.

The maiden flights of the HALO WV DIAL were flights of opportunity that enabled a first assessment of the WV measurements in various forms over a wide range of atmospheric conditions spanning the tropical and midlatitude eastern Pacific, though opportunities for rigorous validation were limited at best. The DLH was operational onboard the DC-8 for most of the campaign, providing a large dataset of in situ WV observations for comparison to the HALO measurements nearest the aircraft (~400 m below). Values spanned 0.004 g/kg to 2 g/kg, and a very small bias of only 0.003 g/kg was found which was likely dominated by spatial variability rather than systematic bias. There was also considerable spread with R^2=0.66, again due in part to spatial variability. A single DLH in situ profile was available for comparison to HALO profiles, and the DLH data was found to be within the range of values measured by HALO during the preceding overpass (with the exception of three very small deviations), indicating good agreement within the expected spatiotemporal variation for the area. Unfortunately, a typical assessment of profiles via statistical comparison to sondes was not possible due to deficiencies in the sonde moisture sensors throughout the campaign. Despite this limitation, the HALO and sonde profiles were found to capture WV features and magnitudes with good agreement.

665 HALO PWV comparisons to community standard spaceborne sounders AIRS and IASI showed excellent agreement. PWV observations over ocean ranged 10 to 60 mm with R^2=0.96 and 0.98 for AIRS and IASI, respectively. HALO measurements were often slightly drier than the sounders in accordance with expectations, considering HALO only measures a partial column based on DC-8 flight altitude. HALO profile comparisons to the sounders' WV profiles highlighted the different capabilities of the instruments, with HALO resolving much more detail and sharper gradients throughout the troposphere.

670 Synergies between active and passive sounders such as these will be critical for improving measurements and information content on global scales in the future (Teixeira et al. 2021).

Although substantial conclusions were drawn, the opportunities for validation of HALO WV during the Aeolus cal/val campaign were not ideal due to the focus and brevity of the campaign. Further validation of HALO will be performed during

675 future campaigns to quantify any potential systematic or other sources of error beyond the statistical uncertainty that is currently reported, though no such errors are evident to date. Improvements to the HALO architecture and processing are also ongoing, including efforts to expand the tunability of the transmitted wavelength targeting the lower troposphere, which will improve performance and expand measurement capability in very moist environments such as the ITCZ. Another major advancement will be the refining of the surface return DIAL algorithm over oceans and attempting to extend this capability

680 to land as well as cloud top, which will further increase utility for PBL and cloud process studies in addition to column products such as PWV.

Data availability

Data from the Aeolus cal/val campaign is available through the NASA Atmospheric Science Data Center (ASDC) at https://doi.org/10.5067/SUBORBITAL/AEOLUSCALVAL2019/DATA001 (NASA/LARC/SD/ASDC).

685 ### Author contribution

Co-authors worked collaboratively on each aspect as part of the HALO team. ARN led conceptualization and development of the HALO instrument with contributions from AN and DBH. SK and JC led data curation and the HALO WV retrieval code with contributions from ARN and BJC. JC, JL, RAB contributed to the integration and operation of HALO during the campaign. BJC and ARN led the analyses presented here. BJC prepared the manuscript with contributions from co-authors.

690 ### Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We acknowledge funding support from the NASA Headquarters Earth Science Division, the NASA Earth Science Technology Office, and NASA Langley Research Center. We thank the DC-8 team at the NASA Armstrong Flight Research Center and the National Suborbital Education and Research Center for their support of the Aeolus cal/val campaign. We thank Joshua Digangi and Glenn Diskin for providing the DLH data used in the preparation of this manuscript. We thank Yankee Environmental Systems for providing the dropsondes and providing technical guidance on the use of the data presented in this manuscript. We also thank Kris Bedka for his leadership during the Aeolus cal/val campaign and fruitful discussions regarding synergistic DIAL retrievals. BJC was supported by the NASA Postdoctoral Program.

References

Abshire, J. B., Riris, H., Weaver, C. J., Mao, J., Allan, G. R., Hasselbrack, W. E., & Browell, E. V. (2013). Airborne measurements of CO 2 column absorption and range using a pulsed direct-detection integrated path differential absorption lidar. *Applied Optics*, 52(19), 4446-4461.

Amediek, A., Ehret, G., Fix, A., Wirth, M., Büdenbender, C., Quatrevalet, M., ... & Gerbig, C. (2017). CHARM-F—a new airborne integrated-path differential-absorption lidar for carbon dioxide and methane observations: measurement performance and quantification of strong point source emissions. *Applied Optics*, 56(18), 5182-5197.

Ansmann, A., 1985. Errors in ground-based water-vapor DIAL measurements due to Doppler-broadened Rayleigh backscattering. Applied optics, 24(21), pp.3476-3480.

Ansmann, A. and Bosenberg, J., 1987. Correction scheme for spectral broadening by Rayleigh scattering in differential absorption lidar measurements of water vapor in the troposphere. Applied optics, 26(15), pp.3026-3032.

Bedka, S., Knuteson, R., Revercomb, H., Tobin, D. and Turner, D., 2010. An assessment of the absolute accuracy of the Atmospheric Infrared Sounder v5 precipitable water vapor product at tropical, midlatitude, and arctic ground-truth sites: September 2002 through August 2008. Journal of Geophysical Research: Atmospheres, 115(D17).

Bedka, K.M., Nehrir, A.R., Kavaya, M., Barton-Grimley, R., Beaubien, M., Carroll, B., Collins, J., Cooney, J., Emmitt, G.D., Greco, S. and Kooi, S., 2021. Airborne lidar observations of wind, water vapor, and aerosol profiles during the NASA Aeolus calibration and validation (Cal/Val) test flight campaign. Atmospheric Measurement Techniques, 14(6), pp.4305-4334.

Behrendt, A., Wulfmeyer, V., Schaberl, T., Bauer, H.S., Kiemle, C., Ehret, G., Flamant, C., Kooi, S., Ismail, S., Ferrare, R. and Browell, E.V., 2007. Intercomparison of water vapor data measured with lidar during IHOP_2002. Part II: Airborne-to-airborne systems. Journal of Atmospheric and Oceanic Technology, 24(1), pp.22-39.

Birk, M., Wagner, G., Loos, J., Lodi, L., Polyansky, O. L., Kyuberis, A. A., ... & Tennyson, J. (2017). Accurate line intensities for water transitions in the infrared: comparison of theory and experiment. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *203*, 88-102.

Black, P., Harrison, L., Beaubien, M., Bluth, R., Woods, R., Penny, A., Smith, R.W. and Doyle, J.D., 2017. High-definition Sounding System (HDSS) for atmospheric profiling. Journal of Atmospheric and Oceanic Technology, 34(4), pp.777-796.

Bony, S., Colman, R., Kattsov, V.M., Allan, R.P., Bretherton, C.S., Dufresne, J.L., Hall, A., Hallegatte, S., Holland, M.M., Ingram, W. and Randall, D.A., 2006. How well do we understand and evaluate climate change feedback processes?. Journal of Climate, 19(15), pp.3445-3482.

Bony, S., Stevens, B., Frierson, D.M., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T.G., Sherwood, S.C., Siebesma, A.P., Sobel, A.H. and Watanabe, M., 2015. Clouds, circulation and climate sensitivity. Nature Geoscience, 8(4), pp.261-268.

Bösenberg, J., 1998. Ground-based differential absorption lidar for water-vapor and temperature profiling: methodology. Applied optics, 37(18), pp.3845-3860.

Browell, E.V., 1983. Remote sensing of tropospheric gases and aerosols with an airborne DIAL system. In Optical and Laser Remote Sensing (pp. 138-147). Springer, Berlin, Heidelberg.

Browell, E.V., Ismail, S. and Grant, W.B., 1998. Differential absorption lidar (DIAL) measurements from air and space. Applied Physics B, 67(4), pp.399-410.

Burton, S.P., Ferrare, R.A., Hostetler, C.A., Hair, J.W., Rogers, R.R., Obland, M.D., Butler, C.F., Cook, A.L., Harper, D.B. and Froyd, K.D., 2012. Aerosol classification using airborne High Spectral Resolution Lidar measurements–methodology and examples. Atmospheric Measurement Techniques, 5(1), pp.73-98.

Carroll, B.J., Demoz, B.B., Turner, D.D. and Delgado, R., 2021. Lidar observations of a mesoscale moisture transport event impacting convection and comparison to Rapid Refresh model analysis. Monthly Weather Review, 149(2), pp.463-477.

Chahine, M.T., Pagano, T.S., Aumann, H.H., Atlas, R., Barnet, C., Blaisdell, J., Chen, L., Fetzer, E.J., Goldberg, M., Gautier, C. and Granger, S., 2006. AIRS: Improving weather forecasting and providing new data on greenhouse gases. Bulletin of the American Meteorological Society, 87(7), pp.911-926.

Chazette, P., Marnas, F., Totems, J. and Shang, X., 2014. Comparison of IASI water vapor retrieval with H 2 O-Raman lidar in the framework of the Mediterranean HyMeX and ChArMEx programs. Atmospheric Chemistry and Physics, 14(18), pp.9583-9596.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S. and Wespes, C., 2009. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. Atmospheric Chemistry and Physics, 9(16), pp.6041-6054.

Cooney, J., 1970. Remote measurements of atmospheric water vapor profiles using the Raman component of laser backscatter. Journal of Applied Meteorology (1962-1982), 9(1), pp.182-184.

Davis, K.J., Browell, E.V., Feng, S., Lauvaux, T., Obland, M.D., Pal, S., Baier, B.C., Baker, D.F., Baker, I.T., Barkley, Z.R. and Bowman, K.W., 2021. The Atmospheric Carbon and Transport (ACT)–America Mission. Bulletin of the American Meteorological Society, pp.1-54.

Diao, M., Jumbam, L., Sheffield, J., Wood, E.F. and Zondlo, M.A., 2013. Validation of AIRS/AMSU-A water vapor and temperature data with in situ aircraft observations from the surface to UT/LS from 87° N–67° S. Journal of Geophysical Research: Atmospheres, 118(12), pp.6816-6836.

Diskin, G. S., Podolske, J. R., Sachse, G.W., and Slate, T. A.: Open-path airborne tunable diode laser hygrometer, Diode Lasers and Applications in Atmospheric Sensing, Proc. SPIE, 4817, https://doi.org/10.1117/12.453736, 2002.

Dobler, J. T., Harrison, F. W., Browell, E. V., Lin, B., McGregor, D., Kooi, S., ... & Ismail, S. (2013). Atmospheric CO 2 column measurements with an airborne intensity-modulated continuous wave 1.57 μm fiber laser lidar. *Applied optics*, *52*(12), 2874-2892.

Doyle, J. D., Moskaitis, J. R., Feldmeier, J. W., Ferek, R. J., Beaubien, M., Bell, M. M., Cecil, D. L., Creasey, R. L., Duran, P., Elsberry, R. L., Komaromi, W. A., Molinari, J., Ryglicki, D. R., Stern, D. P., Velden, C. S., Wang, X., Allen, T., Barrett, B. S., Black, P. G., Dunion, J. P., Emanuel, K. A., Harr, P. A., Harrison, L., Hendricks, E. A., Herndon, D., Jeffries, W. Q., Majumdar, S. J., Moore, J. A., Pu, Z., Rogers, R. F., Sanabia, E. R., Tripoli, G. J., and Zhang, D.: A View of Tropical

Cyclones from Above: The Tropical Cyclone Intensity Experiment, B. Am. Meteorol. Soc., 98, 2113–2134, https://doi.org/10.1175/BAMSD-16-0055.1, 2017.

795 Ehret, G., Kiemle, C., Renger, W. and Simmet, G., 1993. Airborne remote sensing of tropospheric water vapor with a near–infrared differential absorption lidar system. Applied optics, 32(24), pp.4534-4551.

Eichinger, W.E., Cooper, D.I., Forman, P.R., Griegos, J., Osborn, M.A., Richter, D., Tellier, L.L. and Thornton, R., 1999. The development of a scanning Raman water vapor lidar for boundary layer and tropospheric observations. Journal of
800 Atmospheric and Oceanic Technology, 16(11), pp.1753-1766.

Fan, L., Zhang, Y., Chen, S., Guo, P., & Chen, H. (2015, November). Rayleigh-backscattering doppler broadening correction for differential absorption lidar. In *Selected Papers of the Photoelectronic Technology Committee Conferences held June–July 2015* (Vol. 9795, p. 979517). International Society for Optics and Photonics.
805

Ferrare, R.A., Browell, E.V., Ismail, S., Kooi, S.A., Brasseur, L.H., Brackett, V.G., Clayton, M.B., Barrick, J.D.W., Diskin, G.S., Goldsmith, J.E.M. and Lesht, B.M., 2004. Characterization of upper-troposphere water vapor measurements during AFWEX using LASE. Journal of Atmospheric and Oceanic Technology, 21(12), pp.1790-1808.

810 Ferreira, A.P., Nieto, R. and Gimeno, L., 2019. Completeness of radiosonde humidity observations based on the Integrated Global Radiosonde Archive. Earth System Science Data, 11(2), pp.603-627.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R. and Wargan, K., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-
815 2). Journal of climate, 30(14), pp.5419-5454.

Gettelman, A., Weinstock, E.M., Fetzer, E.J., Irion, F.W., Eldering, A., Richard, E.C., Rosenlof, K.H., Thompson, T.L., Pittman, J.V., Webster, C.R. and Herman, R.L., 2004. Validation of Aqua satellite data in the upper troposphere and lower stratosphere with in situ aircraft instruments. Geophysical research letters, 31(22).
820

Goldsmith, J.E.M., Blair, F.H., Bisson, S.E. and Turner, D.D., 1998. Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols. Applied Optics, 37(21), pp.4979-4990.

Gordon, I.E., Rothman, L.S., Hill, C., Kochanov, R.V., Tan, Y., Bernath, P.F., Birk, M., Boudon, V., Campargue, A., Chance, K.V. and Drouin, B.J., 2017. The HITRAN2016 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer, 203, pp.3-69.

Hair, J.W., Hostetler, C.A., Cook, A.L., Harper, D.B., Ferrare, R.A., Mack, T.L., Welch, W., Izquierdo, L.R. and Hovis, F.E., 2008. Airborne high spectral resolution lidar for profiling aerosol optical properties. Applied optics, 47(36), pp.6734-6752.

Hastings, D.A., Dunbar, P.K., Elphingstone, G.M., Bootz, M., Murakami, H., Maruyama, H., Masaharu, H., Holland, P., Payne, J., Bryant, N.A. and Logan, T.L., 1999. The global land one-kilometer base elevation (GLOBE) digital elevation model, version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325, pp.80305-3328.

Higdon, N.S., Browell, E.V., Ponsardin, P., Grossmann, B.E., Butler, C.F., Chyba, T.H., Mayo, M.N., Allen, R.J., Heuser, A.W., Grant, W.B. and Ismail, S., 1994. Airborne differential absorption lidar system for measurements of atmospheric water vapor and aerosols. Applied optics, 33(27), pp.6422-6438.

Hilton, F., Atkinson, N.C., English, S.J. and Eyre, J.R., 2009. Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 135(639), pp.495-505.

Hilton, F., Armante, R., August, T., Barnet, C., Bouchard, A., Camy-Peyret, C., Capelle, V., Clarisse, L., Clerbaux, C., Coheur, P.F. and Collard, A., 2012. Hyperspectral Earth observation from IASI: Five years of accomplishments. bulletin of the american meteorological Society, 93(3), pp.347-370.

Hodges, J. T., Lisak, D., Lavrentieva, N., Bykov, A., Sinitsa, L., Tennyson, J., ... & Tolchenov, R. N. (2008). Comparison between theoretical calculations and high-resolution measurements of pressure broadening for near-infrared water spectra. *Journal of Molecular Spectroscopy*, *249*(2), 86-94.

Ismail, S. and Browell, E.V., 1989. Airborne and spaceborne lidar measurements of water vapor profiles: a sensitivity analysis. Applied Optics, 28(17), pp.3603-3615.

Ismail, S., Ferrare, R.A., Browell, E.V., Chen, G., Anderson, B., Kooi, S.A., Notari, A., Butler, C.F., Burton, S., Fenn, M. and Dunion, J.P., 2010. LASE measurements of water vapor, aerosol, and cloud distributions in Saharan air layers and tropical disturbances. Journal of the atmospheric sciences, 67(4), pp.1026-1047.

860 Kavaya, M. J., Beyon, J. Y., Koch, G. J., Petros, M., Petzar, P. J., Singh, U. N., Trieu, B. C., and Yu, J.: The Doppler Aerosol Wind (DAWN) Airborne, Wind-Profiling Coherent-Detection Lidar System: Overview and Preliminary Flight Results, J. Atmos. Ocean. Tech., 31, 826–842, https://doi.org/10.1175/JTECH-D-12-00274.1, 2014.

Kiemle, C., Groß, S., Wirth, M. and Bugliaro, L., 2017. Airborne lidar observations of water vapor variability in tropical
865 shallow convective environment. In Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity (pp. 253-271). Springer, Cham.

Klaes, K.D., Cohen, M., Buhler, Y., Schlüssel, P., Munro, R., Luntama, J.P., von Engeln, A., Clérigh, E.Ó., Bonekamp, H., Ackermann, J. and Schmetz, J., 2007. An introduction to the EUMETSAT polar system. Bulletin of the American
870 Meteorological Society, 88(7), pp.1085-1096.

Leblanc, T., McDermid, I.S. and Walsh, T.D., 2012. Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring. Atmospheric Measurement Techniques, 5(1), pp.17-36.

875 Liu, Z., Hunt, W., Vaughan, M., Hostetler, C., McGill, M., Powell, K., Winker, D. and Hu, Y., 2006. Estimating random errors due to shot noise in backscatter lidar observations. Applied optics, 45(18), pp.4437-4447.

Le Marshall, J., Jung, J., Derber, J., Chahine, M., Treadon, R., Lord, S.J., Goldberg, M., Wolf, W., Liu, H.C., Joiner, J. and Woollen, J., 2006. Improving global analysis and forecasting with AIRS. Bulletin of the American Meteorological Society,
880 87(7), pp.891-894.

Martins, J.P., Teixeira, J., Soares, P.M., Miranda, P.M., Kahn, B.H., Dang, V.T., Irion, F.W., Fetzer, E.J. and Fishbein, E., 2010. Infrared sounding of the trade-wind boundary layer: AIRS and the RICO experiment. Geophysical Research Letters, 37(24).
885

Moore, A.S., Brown, K.E., Hall, W.M., Barnes, J.C., Edwards, W.C., Petway, L.B., Little, A.D., Luck, W.S., Jones, I.W., Antill, C.W. and Browell, E.V., 1997. Development of the Lidar Atmospheric Sensing Experiment (LASE)—an advanced airborne DIAL instrument. In Advances in Atmospheric Remote Sensing with Lidar (pp. 281-288). Springer, Berlin, Heidelberg.

890

NASA/LARC/SD/ASDC: Aeolus CalVal HALO Aerosol and Water Vapor Profiles and Images, NASA Langley Atmospheric Science Data Center DAAC [data set], https://doi.org/10.5067/Airborne/Aeolus-CalVal-HALO_DC8_1.

Nehrir, A. R., Repasky, K. S., Carlsten, J. L., Obland, M. D., & Shaw, J. A. (2009). Water Vapor Profiling Using a Widely
895 Tunable, Amplified Diode-Laser-Based Differential Absorption Lidar (DIAL), Journal of Atmospheric and Oceanic Technology, 26(4), 733-745.

Nehrir, A. R., Repasky, K. S., & Carlsten, J. L., 2011. Eye-safe diode-laser-based micropulse differential absorption lidar (DIAL) for water vapor profiling in the lower troposphere. Journal of Atmospheric and Oceanic Technology, 28(2), 131-147.
900

Nehrir, A.R., Repasky, K.S. and Carlsten, J.L., 2012. Micropulse water vapor differential absorption lidar: transmitter design and performance. Optics express, 20(22), pp.25137-25151.

Nehrir, A.R., Kiemle, C., Lebsock, M.D., Kirchengast, G., Buehler, S.A., Löhnert, U., Liu, C.L., Hargrave, P.C., Barrera-
905 Verdejo, M. and Winker, D.M., 2017. Emerging technologies and synergies for airborne and space-based measurements of water vapor profiles. Surveys in geophysics, 38(6), pp.1445-1482.

Philbrick, C. R., 1994. Raman lidar measurements of atmospheric properties. In Atmospheric Propagation and Remote Sensing III (Vol. 2222, pp. 922-931). International Society for Optics and Photonics.
910

Podolske, J.R., Sachse, G.W. and Diskin, G.S., 2003. Calibration and data retrieval algorithms for the NASA Langley/Ames Diode Laser Hygrometer for the NASA transport and chemical evolution over the pacific (TRACE-P) mission. Journal of Geophysical Research: Atmospheres, 108(D20).

915 Remsberg, E.E. and Gordley, L.L., 1978. Analysis of differential absorption lidar from the Space Shuttle. Applied optics, 17(4), pp.624-630.

Richardson, M. T., Thompson, D. R., Kurowski, M. J., & Lebsock, M. D. (2021b). Boundary layer water vapour statistics from high-spatial-resolution spaceborne imaging spectroscopy. *Atmospheric Measurement Techniques Discussions*, 1-33.
920

Roman, J., Knuteson, R., August, T., Hultberg, T., Ackerman, S. and Revercomb, H., 2016. A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations. Journal of Geophysical Research: Atmospheres, 121(15), pp.8925-8948.

925 Schäfler, A., Fix, A. and Wirth, M., 2021. Mixing at the extratropical tropopause as characterized by collocated airborne H 2 O and O 3 lidar observations. Atmospheric Chemistry and Physics, 21(6), pp.5217-5234.

Schotland, R.M., 1974. Errors in the lidar measurement of atmospheric gases by differential absorption. Journal of Applied Meteorology, 13(1), pp.71-77.

930

Sherwood, S.C., Roca, R., Weckwerth, T.M. and Andronova, N.G., 2010. Tropospheric water vapor, convection, and climate. Reviews of Geophysics, 48(2).

Späth, F., Behrendt, A., Muppa, S.K., Metzendorf, S., Riede, A. and Wulfmeyer, V., 2016. 3-D water vapor field in the 935 atmospheric boundary layer observed with scanning differential absorption lidar. Atmospheric Measurement Techniques, 9(4), pp.1701-1720.

Späth, F., Behrendt, A., & Wulfmeyer, V. (2020). Minimization of the Rayleigh-Doppler error of differential absorption lidar by frequency tuning: a simulation study. Optics Express, 28(20), 30324-30339.

940

Spuler, S.M., Repasky, K.S., Morley, B., Moen, D., Hayman, M. and Nehrir, A.R., 2015. Field-deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor. Atmospheric Measurement Techniques, 8(3), pp.1073-1087.

Spuler, S.M., Hayman, M., Stillwell, R.A., Carnes, J., Bernatsky, T. and Repasky, K.S., 2021. MicroPulse DIAL (MPD)–a 945 diode-laser-based lidar architecture for quantitative atmospheric profiling. Atmospheric Measurement Techniques, 14(6), pp.4593-4616.

Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J.M., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A. and Meynart, R., 2005. The atmospheric dynamics mission for global wind field measurement. Bulletin of the 950 American Meteorological Society, 86(1), pp.73-88.

Teixeira, J., J. R. Piepmeier, A. R. Nehrir, C. O. Ao, S. S. Chen, C. A. Clayson, A. M. Fridlind, M. Lebsock, W. McCarty, H. Salmun, J. A. Santanello, D. D. Turner, Z. Wang, and X. Zeng, 2021: Toward a Global Planetary Boundary Layer Observing System: The NASA PBL Incubation Study Team Report. NASA PBL Incubation Study Team. 134 pp.

955

Thrastarson, H. T., Manning, E., Kahn, B., Fetzer, E., Yue, Q., Wong, S., ... & Wilson, R. C., 2020. AIRS/AMSU/HSB Version 7 Level 2 Product User Guide. Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA, 83-92.

960 Thompson, D. R., Kahn, B. H., Brodrick, P. G., Lebsock, M. D., Richardson, M., & Green, R. O. (2021a). Spectroscopic imaging of sub-kilometer spatial structure in lower-tropospheric water vapor. *Atmospheric Measurement Techniques*, *14*(4), 2827-2840.

Trenberth, K.E., Jones, P.D., Ambenje, P., Bojariu, R., Easterling, D., Klein Tank, A., Parker, D., Rahimzadeh, F., Renwick, 965 J.A., Rusticucci, M. and Soden, B., 2007. Observations. Surface and atmospheric climate change. Chapter 3.

Turner, D.D., Ferrare, R.A., Brasseur, L.H., Feltz, W.F. and Tooman, T.P., 2002. Automated retrievals of water vapor and aerosol profiles from an operational Raman lidar. Journal of Atmospheric and Oceanic Technology, 19(1), pp.37-50.

970 Turner, D.D. and Löhnert, U., 2021. Ground-based temperature and humidity profiling: combining active and passive remote sensors. Atmospheric Measurement Techniques, 14(4), pp.3033-3048.

Wakimoto, R.M., Murphey, H.V., Browell, E.V. and Ismail, S., 2006. The "triple point" on 24 May 2002 during IHOP. Part I: Airborne Doppler and LASE analyses of the frontal boundaries and convection initiation. Monthly weather review, 134(1), 975 pp.231-250.

Whiteman, D.N., Melfi, S.H. and Ferrare, R.A., 1992. Raman lidar system for the measurement of water vapor and aerosols in the Earth's atmosphere. Applied optics, 31(16), pp.3068-3082.

980 Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F. and Ehret, G., 2009. The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Applied Physics B, 96(1), pp.201-213.

Wong, S., Fetzer, E.J., Schreier, M., Manipon, G., Fishbein, E.F., Kahn, B.H., Yue, Q. and Irion, F.W., 2015. Cloud-induced uncertainties in AIRS and ECMWF temperature and specific humidity. Journal of Geophysical Research: Atmospheres, 985 120(5), pp.1880-1901.

Wu, Y., Nehrir, A.R., Ren, X., Dickerson, R.R., Huang, J., Stratton, P.R., Gronoff, G., Kooi, S.A., Collins, J.E., Berkoff, T.A. and Lei, L., 2021. Synergistic aircraft and ground observations of transported wildfire smoke and its impact on air quality in New York City during the summer 2018 LISTOS campaign. Science of The Total Environment, 773, p.145030.

990

Wulfmeyer, V., 1998. Ground-based differential absorption lidar for water-vapor and temperature profiling: development and specifications of a high-performance laser transmitter. Applied optics, 37(18), pp.3804-3824.

Wulfmeyer, V. and Bösenberg, J., 1998. Ground-based differential absorption lidar for water-vapor profiling: assessment of

995 accuracy, resolution, and meteorological applications. Applied Optics, 37(18), pp.3825-3844.

Wulfmeyer, V., Bauer, H.S., Grzeschik, M., Behrendt, A., Vandenberghe, F., Browell, E.V., Ismail, S. and Ferrare, R.A., 2006. Four-dimensional variational assimilation of water vapor differential absorption lidar data: The first case study within IHOP_2002. Monthly weather review, 134(1), pp.209-230.

1000

Wulfmeyer, V., Hardesty, R.M., Turner, D.D., Behrendt, A., Cadeddu, M.P., Di Girolamo, P., Schlüssel, P., Van Baelen, J. and Zus, F., 2015. 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. Reviews of Geophysics, 53(3), pp.819-895.