Assessment of Mixed-Layer Height Estimation from Single-wavelength Ceilometer Profiles

T. N. Knepp[1,2], J. S. Szykman[3,4], R. Long[3], R. Duvall[3], J. Krug[3], M. Beaver[3], K. Cavender[3], K. Kronmiller[5,3], M. Wheeler[5,3], R. Delgado[6], R. Hoff[6], T. Berkoff[2], E. Olson[8], R. Clark[9], D. Wolfe[10], D. Van Gilst[11], and D. Neil[2]

[1]Science Systems and Applications Inc., Hampton, VA, United States

[2]NASA Langley Research Center, Hampton, VA, United States

[3]US EPA, Research Triangle Park, Durham, NC, United States

[4]Currently assigned to NASA Langley Research Center, Hampton, VA 23681, USA

[5]Jacobs Technology Inc., Tullahoma, TN 37388, USA

[6]Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, MD 21250, USA

[7]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[8]Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI 53706, USA

[9]Department of Earth Sciences, Millersville University, Millersville, PA 17551, USA

[10]NOAA/ESRL Physical Sciences Division, Boulder, CO 80305, USA

[11]Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, CO 80309, USA

[12]National Suborbital Education and Research Center, University of North Dakota, Grand Forks, ND 58202, USA

Correspondence to: T. N. Knepp
(travis.n.knepp@nasa.gov)

Abstract. An assessment of differing boundary/mixed-layer height measurement methods was performed with a focus on the Vaisala CL51 instrument and BLView and STRAT softwares. Of primary interest was determining how these differing methodologies will intercompare when deployed as part of a larger instrument network. The intercomparisons were performed as part of ongoing measurements at the Chemistry And Physics of the Atmospheric Boundary Layer Experiment (CAPABLE) site in Hampton, VA and during the 2014 Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) field campaign that took place in the Denver, CO area. It was observed that data collection methodology is not as important as the processing algorithm, and that, generally speaking, sonde-derived boundary layer heights are higher than LIDAR-derived mixed-layer heights.

1   Introduction

The atmospheric boundary layer (ABL) is the lowermost portion of the troposphere that is directly influenced by the Earth's surface and responds to surface forcing of heat, moisture, pollutant emissions, and momentum on timescale of an hour or less Stull (1988). The ABL can be defined via a number of criteria depending on the particular interest (e.g. the thermodynamic boundary layer, chemical boundary layer (CBL), aerosol mixed layer, etc.). Traditionally, the ABL has been defined by thermodynamic data (i.e. potential temperature and/or skew-T plot) obtained from meteorological sondes. While meteorological sondes have excellent vertical resolution, the temporal resolution is generally poor, and ongoing regular sonde launches are labor intensive. Conversely, mixed-layer heights (MLH) as calculated from backscatter LIDAR instruments provide both excellent vertical and temporal resolution. With respect to air quality, the top of the ABL often acts like a lid within the lowest layer of the atmosphere and temporarily traps the majority of near-surface anthropogenic and biogenic emissions. As a result, the vertical distribution of ambient air pollutants, and associated precursors, within the ABL and lower-troposphere are strongly influenced by the height of, and vertical mixing within, the ABL.

In 2009 the National Research Council highlighted planetary (atmospheric) boundary layer height as a high priority observation needed to address current inadequacies at the meso-scale for improved predictions of air quality, short-range severe-weather forecasting, and regional climate modeling. More recently, The National Plan for Civil Earth Observation (NSTC, 2014) called out the need for improved observation density and sampling of the boundary layer. In 2015, as part of the revisions to the ozone (O_3) National Ambient Air Quality Standard (NAAQS), EPA finalized a new requirement under the Photochemical Assessment Monitoring Stations (PAMS) program for the collection of continuous mixing layer height observations. By 2019, the PAMS program will involve the implementation of approximately fifty air-quality sites around the United States, providing measurements of MLH on a continuous basis.

From 2011 through 2014 NASA conducted the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) Earth Venture Suborbital Mission with four field deployments: Baltimore/Washington region of Maryland during 2011; the San Joaquin Valley (SJV) of California during January-February 2013; Houston, Texas during September 2013; and the Front Range region of Colorado in July-August 2014. A primary objective of DISCOVER-AQ was to investigate the use of satellite remote sensing ability to inform air quality at the surface. Since the ABL limits vertical exchange of primary pollutants, and controls near-surface pollutant concentrations, the ABL height can directly influence air quality and chemistry. Therefore, measurements during these missions focused on the vertical distribution of trace gases and aerosols within the ABL and lower troposphere, and the diurnal variability of these distributions in conjunction with the ABL.

ABL variability poses a complication in quantitative determination of surface trace-gas levels from a remote-sensing platform Herman et al. (2009); Knepp et al. (2015); Lamsal et al. (2008, 2014); Petritoli et al. (2004); Piters et al. (2012). Therefore, properly accounting for ABL variability from a continuous measurement system such as LIght Detection And Ranging (LIDAR) will provide invaluable information to policy, health, modeling, and remote-sensing communities for applications sensitive to the vertical profiles of tracers Compton et al. (2013); Martin (2008); Scarino et al. (2014). Herein is presented results from an intercomparison of three backscatter LIDAR's from the 2014 DISCOVER-AQ field campaign in Colorado and coincident sonde launches from the Chemistry and Physics of the Atmospheric Boundary Layer Experiment (CAPABLE) site at NASA's Langley Research Center (LaRC) in Hampton, VA.

2   Instrumentation

2.1   CL51

The Vaisala CL51 ceilometer is a single-wavelength (eye safe Class 1M InGaAs diode LASER emitting at 910 ± 10 nm, pulsed at 6.5 kHz with a 110 ns pulse width, and average pulse power of 19.5 mW, with avalanche photodiode detector centered at 915 nm), single-lens, LIDAR system originally designed to report cloud-base heights and visibility. More recently, ceilometers have been used to estimate MLH (Emeis and Schäfer (2006); Emeis et al. (2008a, b); Haeffelin et al. (2012); Morille et al. (2007); Schäfer et al. (2012, 2013); Schween et al. (2014); Sokol et al. (2014); Wiegner et al. (2014)). These ceilometers have 10 m vertical resolution (with 10 m overlap) to a maximum altitude of 15.4 km ($\pm$ greater of 1% or 5 m precision), and up to 2 s temporal resolution (depending on the control software), though profiles are generally averaged over 16–36 s to improve signal-to-noise (see Sec. 3.1 for more details). An example backscatter plot that includes increased signal at 3 km due to transport of smoke from a Canadian forest fire is presented in Fig. 1.

The CL51 was designed to operate continuously, regardless of meteorological conditions, in an autonomous manner with minimal user support. Due to the emission wavelength's proximity to the near-infrared water vapor bands these ceilometers experience water vapor interference, thereby mitigating their utility in retrieval of aerosol optical properties. However, the interference on aerosol profile and MLH estimation is negligible (Wiegner et al. (2014)).

Two CL51's were deployed as part of the 2014 DISCOVER-AQ mission in Colorado (Golden, and Erie, CO). Before and after deployment these ceilometers were stationed at CAPABLE and the EPA Ambient air Innovative Research Site (AIRS) in Durham, NC, continually collecting data. The ceilometers were collocated with met-sonde launch sites during the DISCOVER-AQ campaign and at CAPABLE, allowing a direct intercomparison of the sonde and LIDAR ABL/MLH methodologies. Furthermore, during the DISCOVER-AQ campaign the ceilometers were collocated with other LIDAR instruments. Intercomparisons are presented below.

2.1.1   Ceilometer Full-profile Collection

The BLView software not only provides data analysis (e.g. MLH and cloud-height estimates), but also provides data-logging/archiving capability. While the CL51 reports backscatter up to 15.4 km, BLView truncates the profile data collection at 4.5 km. Generally speaking, there is little need to collect higher-altitude backscatter data for reprocessing purposes due to the relative simplicity of detecting cloud bases. However, failure to log the full-profile reduces the ability to monitor upper-troposphere/lower-stratosphere (UTLS) transport of aerosol, smoke, or ash from major

events. Therefore, a full-profile collection method that can run side-by-side with the standard data-collection software was developed and implemented.

Data transmission from the ceilometer to the logging computer was achieved over a simple RS-232 connection that can be split into two ports on the logging computer; one port logging to BLView, the other logging to a custom script (e.g. as written in Python, or terminal emulation). The primary drawback of using a secondary script to log the full profile (as opposed to logging in BLView) is the inability to apply calibration coefficients to the logged data. However, as shown in subsequent sections, this impacts neither the MLH estimates nor the general profile shape substantially.

2.2 MPL

Elastic lidar observations were performed with using a Sigma Space Micropulse Lidar (MPL), and the instrument has been described previously (Spinhirne, 1993; Welton et al., 2000). Briefly, the MPL transmitter consists of a Nd:YLF laser emitting at 527 nm, pulsed at 2.5 kHz and average pulse power of $6 - 10$ μJ. The receiver consists of a 178 mm telescope that collects the backscattered light. The output from the telescope is conveyed to a photon counting silicon avalanche photo-diode (APD). The APD output is recorded by a field programmable gate array (FPGA) data system that enables display and storage of range dependent averaged count rates on a laptop computer. The raw data are converted to aerosol attenuated backscatter by taking into account instrumental factors that include corrections for detector dead time, geometrical overlap, background subtraction, and range-squared normalization. Recorded lidar profiles have temporal and vertical resolution of one minute and 30 meters, respectively.

The Micro Pulse LIDAR (MPL) system is a single-wavelength (eye safe Class 2 diode LASER emitting at 532 nm, pulsed at 2.5 kHz with an average pulse power of $3 - 4$ μJ) LIDAR for continuous recording of aerosol profiles, optical properties, and calculating MLH values. The MPL has a software programmable vertical resolution, with possible values of 15, 30, and 75 m (up to 25 km), and temporal resolutions ranging from 1 s – 15 minutes.

2.3 Meteorological/Ozone Sondes

The traditional method of identifying the ABL is using meteorological sondes (herein referred to as sondes) to identify steep gradients within the potential temperature (theta) profile (Fig. 2 A) as identified by the Heffter criteria (Heffter (1980); Marsik et al. (1995)), which is a product of atmospheric turbulent kinetic energy. Similar gradients can be seen in chemical and aerosol profiles as well (Fig. 2 B-C). For the current study, meteorological sondes from International Met Systems (iMet) and ozone sondes from Droplet Measurement Technologies (DMT, now En-Sci) were used. iMet sondes require no preparation and were used as received

from the manufacturer, while ozone sondes were conditioned according to the procedure defined by the World Meteorological Organization recommendations (Smit (2013)).

Numerous analyses have been presented to illustrate differences between the various chemical and meteorological sensors, and how differing meteorological sensors influence secondary chemical measurements such as ozone (Deshler et al. (2008); Dirksen et al. (2014); Johnson et al. (2002); Miloshevich et al. (2004); Nash et al. (2006, 2011); Smit (2013); Stauffer et al. (2014)). Therefore, discussion of these differences and their influence on ABL will not be made within the current manuscript.

3 Algorithms

3.1 BLView

The Vaisala standard MLH retrieval is based on a proprietary wavelet/gradient technique built within the logging/analysis software BLView. BLView makes use of variable time and altitude averaging when calculating the MLH. Typical averaging time ranges from 14 minutes at night to 52 minutes during the clear-sky, daytime conditions, and is automatically adjusted within the software according to signal-to-noise. Altitude averaging is variable with altitude, and ranges from 80 m near the surface to 360 m above 1.5 km. Further, BLView selectively removes false-positive MLH identifications by requiring a minimum number of similar MLH values (±140 m) be within the last several minutes.

BLView contains the ability to discriminate between MLH inversions and changes in backscatter intensity induced by cloud, precipitation, and fog.

An advantage of the BLView software is the standardization of the retrieval parameters, and a user interface that provides flexibility in setting user-specified sensitivities. This comes at the cost of a database system that makes access to raw data difficult, and the inability to batch process archived data, posing a severe limitation on reprocessing datasets with a long record history. However, BLView is expected to move from these limitations in future releases.

3.2 STRAT

The STRucture of the ATmosphere (STRAT v1.04) algorithm was developed under a GNU General Public License to analyze aerosol vertical profiles, as measured via LIDAR, estimate cloud heights and aerosol MLH from a variety of LIDAR instruments, and is currently in use by the European Aerosol Research LIdar NETwork (EARLINET) (Haeffelin et al. (2012); Hirsikko et al. (2014); Morille et al. (2007); Pappalardo et al. (2014)). STRAT utilizes a covariance wavelet technique (CWT), of which the full details can be found in Morille et al. (2007) and Haeffelin et al. (2012). STRAT can be run exclusively in MatLab, or a combination of MatLab and Python. Due to its wide use throughout the
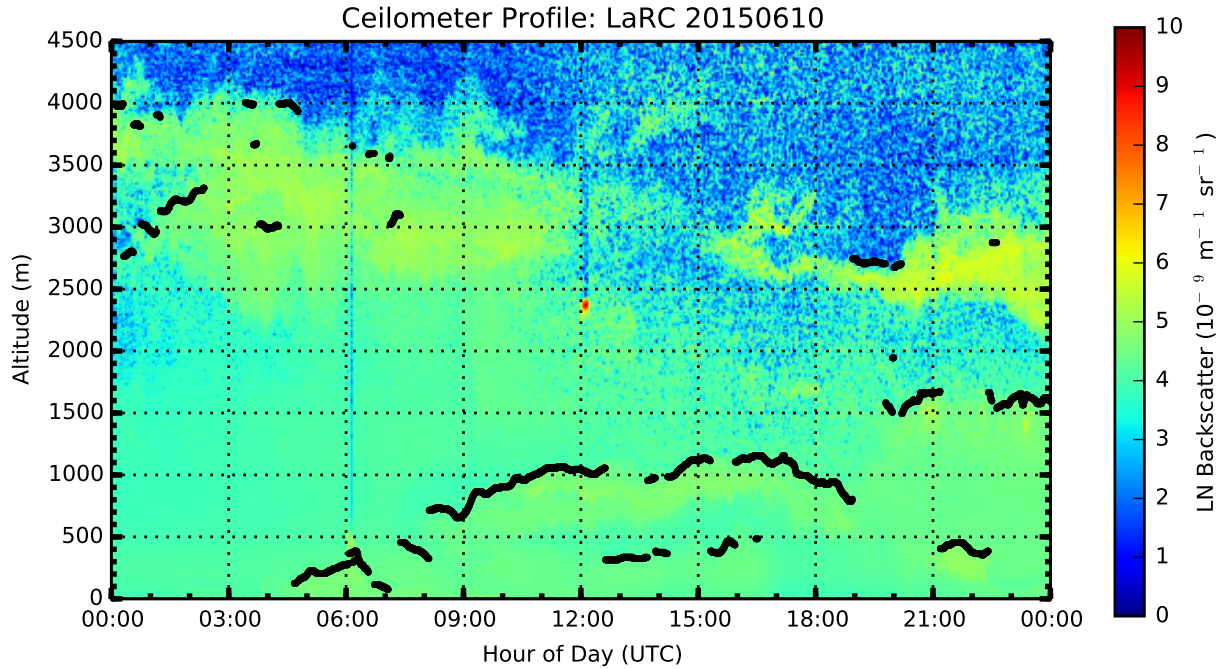
Figure 1. Backscatter curtain plot collected on 10-June, 2015 when smoke from a Canadian forest fire was transported over CAPABLE. The smoke is observed by increased backscatter in the 2500 – 4000 m range.
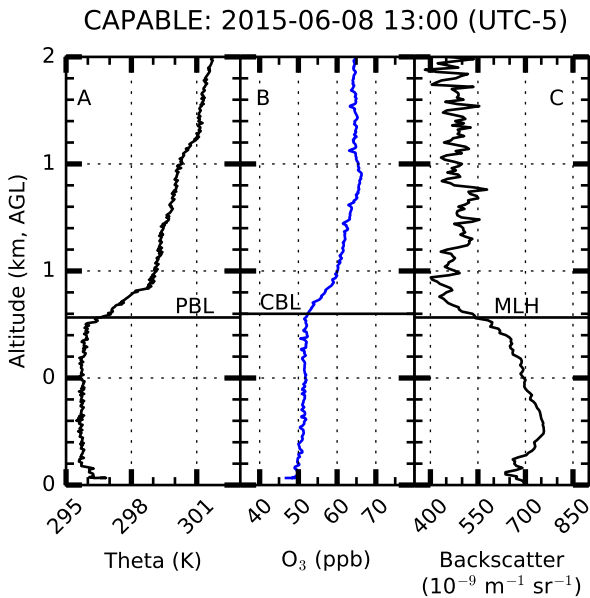

Figure 2. Potential temperature, ozone, and backscatter profiles recorded on 8-June 2015. The ABL, CBL, and MLH can be seen by the horizontal lines

European network it is considered here as a viable open-source alternative to BLView.

While BLView provides limited user control of the retrieval process, which is beneficial in regards to standardizing the retrieval process across a network, STRAT provides a significantly greater amount of user control. Such control is beneficial since retrieval parameters in a heavily polluted region will likely be different than retrievals done in a clean environment. Further, STRAT is provided as raw scripts as opposed to BLView's compiled executable, making the STRAT platform independent and highly user-configurable. STRAT also has the ability to run batch jobs, which is beneficial when reprocessing data from instruments that have a long record history.

The STRAT algorithm implements a user-defined normally-distributed weighting function in both the temporal and vertical domains to smooth the data, similar to BLView. In the current study, the STRAT parameters were set to match the BLView settings as much as possible for intercomparison. An analysis of how well the two MLH algorithms agree is presented below.

3.3 UMBC Algorithm

The UMBC algorithm was developed for estimating MLH from lidar backscatter profiles using a CWT similar to STRAT. By using the aerosols as tracers of the atmospheric dynamics, the lidar is a powerful tool for visualizing, in real

time, with high temporal and spatial evolution of the MLH. The MLH contains greater aerosol concentration because the aerosols are trapped in the PBL by a potential temperature inversion. Therefore, the backscatter signal strength is dramatically reduced when it transits from the PBL into the free troposphere. A covariance wavelet technique (CWT) was applied to the lidar signal to estimate these sharp gradient changes in the lidar backscatter profiles to determine the MLH (Davis et al. 2000; Brooks 2003).

Detailed description of the UMBC algorithm has been published elsewhere Compton et al. (2013). The first step in the MLH algorithm defines the dilation and center of the Haar function values considered in the CWT. The second step consists of applying the CWT to the lidar profile for the appropriate dilation and center of the Haar function values. The sharp gradients in the profile that are of interest are identified by local minima in the resulting wavelet covariance profile. The local minimum is selected as the MLH, and the process is repeated for each profile in the data set.

4 Locations

4.1 CAPABLE

The CAPABLE site (37.103º N, 76.387º W, 5 m ASL) was established at LaRC, in the greater Hampton Roads region (collection of cities on coastal Virginia, also known as Tidewater Virginia: Virginia Beach, Norfolk, Chesapeake, Newport News, Hampton, Portsmouth, Suffolk, Poquoson, Williamsburg), for continuous monitoring of air-quality and meteorological parameters to bridge the gap between satellite observations and ground conditions (i.e. where pollutants directly impact living organisms), improve applicability of satellite data to the air-quality user community, and act as a long-term satellite validation site. CAPABLE has a suite of in-situ and remote-sensing instruments, including a CL51 ceilometer and sounding station, that allows thorough sampling of the atmosphere to provide valuable in-situ and profile information within the lower troposphere in a highly complex (i.e. due to bay-breeze events, see Martins et al. (2012)) and moderately polluted environment that will provide valuable satellite ground-truthing and model a-priori estimates.

CAPABLE is located on a peninsula between the James River to the southwest, Chesapeake Bay to the north, and the Atlantic Ocean to the east. Immediate emission sources and their locations relative to CAPABLE are: commuter traffic (Wythe Creek Rd to the west ≈15,000 vehicles per day, Commander Shepard Blvd to the south at ≈20,000 vehicles per day; Commander Shepard and Wythe Creek share much of the same traffic, so it is not reasonable to estimate a total traffic flow of 35,000 vehicles per day), Yorktown Power Station (approximately 350 MW, 1150 MW peak) and Yorktown oil refinery to the north-northwest, Langley Air Force Base to the southeast, Richmond, VA to the west, and Baltimore/Washington D.C. further to the north.

4.2 Erie, CO/BAO-Tower

Data were collected at the Erie, CO site (40.045º N, 105.005º W, 1500 m ASL) from 14-July–12-August 2014 as part of the DISCOVER-AQ field mission. The Erie, or BAO-Tower, site was located at NOAA's Earth System Research Laboratory's (ESRL) Boulder Atmospheric Observatory (BAO) in Erie, CO, a rural community surrounded by agricultural activity. The Erie site's primary feature was a 300 m tower (known as BAO-Tower), which provided a unique profiling ability for in-situ samplers by mounting them on the tower for static sampling, or on the carriage to collect "active" profiles. Further, a CL31 is permanently located at the site.

During DISCOVER-AQ 2014, the University of Wisconsin's (UW) Space Science and Engineering Center trailer, which housed a High Spectral Resolution LIDAR (HSRL) and from which regular sonde launches were performed, was stationed at the site. The UW trailer temporarily housed a CL51 during the mission. Due to the proximity of the UW trailer and the CL31, both ceilometers experienced the same chemical, aerosol, and meteorological conditions.

4.3 Golden, CO

Data were collected at the Golden, CO site (39.750º N, 105.183º W, 1850 m ASL) from 14-July–12-August 2014 as part of the DISCOVER-AQ field mission. The Golden site was located next to the National Renewable Energy Laboratory (NREL) on Table Mountain mesa. Due to the site's elevation on the mesa, and limited emissions sources, conditions at the Golden site were generally clean from an aerosol perspective and did not typically experience a well-developed ABL/ML.

The Golden site housed the EPA trailer, the LaRC ozone LIDAR, micro-pulse LIDAR (MPL) and LEOSPHERE ALS-450 LIDAR operated by UMBC, a SOund Detection and Ranging (SODAR) instrument operated by Millersville University (MU), and regular met-sonde launches from the MU group.

5 Analysis

LIDAR data collected during the DISCOVER-AQ campaign had sampling times that ranged from 36–60 s, while sonde-profile data had instantaneous measurement times of 1 s. To harmonize all datasets onto a common time frame the data were averaged to 5-minute resolution unless otherwise specified. Further, it is well known that the atmosphere changes throughout the day due to surface heating, etc. (hence, driving ABL variability). Therefore, to remove biases caused by "time-of-day" influences some of the analyses were broken into four-hour segments. Since the primary objective of the

assessment is to understand how the CL51 MLH compares with other instruments/methods, all work is presented in relation to the CL51.

The analysis was performed using several ceilometer MLH products to do a thorough comparison of instruments (i.e. CL51, MPL, and met-sondes), collection method (i.e. allowing BLView to collect profile data with application of calibration factors vs. logging raw data with a custom Python script), and data processing algorithm (i.e. BLView vs. STRAT and custom MLH scripts from UMBC). Assessment of data acquisition methodology will be presented first, followed by a comparison of MLH retrieval algorithm on data collected by a single instrument, and finally a comparison of the various instrumentation.

5.1 Data Acquisition

The objective of the current subsection is to determine whether the CL51 data-logging methodology influenced the MLH estimate. As described above, CL51 profile data were logged using two methodologies: BLView and a custom Python routine. The BLView software has the advantage of applying the ceilometer's calibration factors and preconditioning the profiles (here referred to as BLView; note: this refers to the backscatter-profile that is logged by BLView, not the BLView-calculated MLH), while the Python script logged the raw incoming data stream up to the full profile height (i.e. 15.4 km, this dataset is referred to as full-profile, or FP). The question being: does application of the LIDAR calibration factor influence the MLH estimate? This question will be addressed in section 5.1.2, but first, a viable filtering criteria that removes spurious MLH fluctuations from the dataset must be defined prior to analysis. Defining this criteria will be the topic of 5.1.1.

5.1.1 Filtering Procedure

Regardless of the method of data acquisition (i.e. via BLView or Python) a pragmatic data-selection criteria must be established that provides reasonable assurance that the MLH estimates, which will be fed into chemical models in subsequent studies, are representative of MLH/ABL conditions. Since the ABL/MLH vary in a generally-smooth manner it is expected that the variance within a short time interval will be likewise minimal, and that any larger variance is indicative of other events (e.g. precipitation, contamination). Therefore, it remains to identify these cutoff criteria for implementing data filtering. Since the effect of the implementation of these cutoff criteria will influence the data-acquisition comparison (i.e. BLView-corrected data vs. raw data collected via the Python script), this portion of the analysis is presented first.

Despite the atmosphere's smooth variation in ABL and MLH, these parameters do change substantially over long periods of time (e.g. an hour or day), which significantly increases the standard deviation over significantly long time periods. Therefore, the current analysis must be performed on short-time-series data (e.g. 5 – 10 minutes) to eliminate a bias caused by natural low-frequency changes. Figure 3 shows a series of histogram plots for data collected at LaRC (the largest dataset within the current analysis), where the standard deviation was calculated over five-minute intervals. This figure is elucidative as it shows the distribution of the MLH standard deviation for both collection methods, with the vertical dashed lines representing percentiles of the total data collected. It is observed that, excluding the afternoon period (12:00 – 19:00 local time where the variability is slightly increased), 85% of the data fall within one standard deviation (≈ 0.20 km) regardless of time of day. Therefore, data that have a five-minute standard deviation greater than 0.20 km were removed from subsequent analysis (labeled "filtered") and data that have a relative standard deviation greater than, or equal to, 20% were also removed.

This filtering method is further supported by observing the variability in the BLView and Python-collected datasets (both processed in STRAT) in relation to backscatter curtains (Fig. 4) where it is observed that much of the difference between the BLView and Python-collected data occurs during times of high variability or precipitation (e.g. 19:00 – 24:00 in Fig. 4). During such events, neither collection method is expected to provide valid MLH estimates; rather, to overcome such discrepancies, if possible, the MLH algorithms must be adjusted accordingly.

5.1.2 Collection Method Dependence

To determine whether the data-collection method influenced MLH estimates, both BLView and Python-collected backscatter profiles were processed on a common algorithm (STRAT), using identical input configuration files. Both the BLView and FP profiles were processed using the STRAT algorithm as described above, followed by a 5-minute block average. Figure 5 presents the data as correlation plots with the z-axis representing the relative standard deviation (i.e. standard deviation divided by mean; non-filtered data) within the 5-minute interval. The data were replotted with the z-axis being representative of the immediate data density (a dimensionless value that has been scaled to 1). The data density was calculated by implementing a Gaussian-based kernel-density estimation (Scott (1992); Silverman (1986)) as supplied in Python's scipy.stats.kde module (represented mathematically in Eqs. 1–3 where $\mathbf{X}$ is the 2 x n vector of the $\mathbf{x}$ and $\mathbf{y}$ vectors (i.e. flattened and stacked atop one another), n represents the number of points within each dataset (assuming datasets are of equal length), f is the Scott's factor ($n^{\frac{-1}{d+4}}$), d is the number of independent datasets analyzed, and Eq. 3 is evaluated over the range 1 to n. As these density values are later used as weights in subsequent calculations, the output vector is labeled $\mathbf{w}$ here. It is observed that the majority of the MLH estimates fall along the 1:1 line (center column in Fig. 5), though there is significant scatter along both axes. The source of the
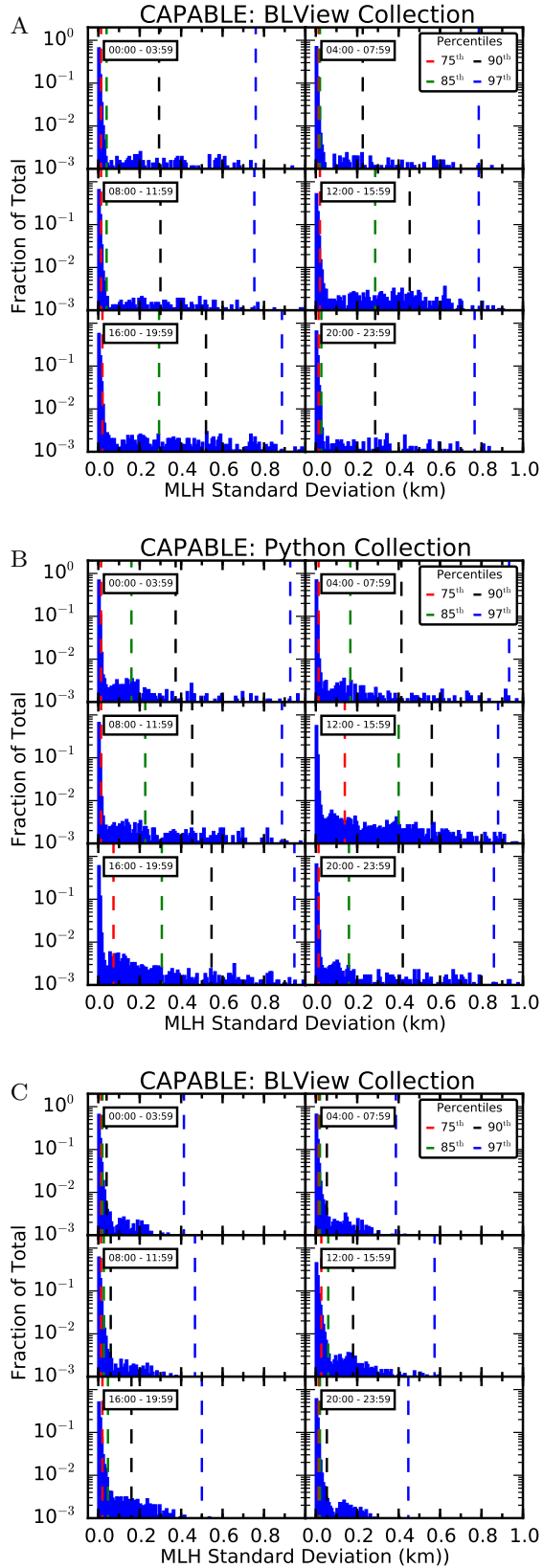
Figure 3. Histogram plots showing the distribution of MLH standard deviations. Vertical lines represent percentiles. Panels A and B were processed in STRAT; panel C was processed in BLView.

scatter, as can be seen in the relative standard deviation intensities, is the variability within each five-minute averaging block, supporting the filter selection criteria.

$$\Delta \mathbf{X} = \mathbf{X} - \mathbf{X}[:, i] \tag{1}$$

$$\mathbf{E} = \sum_{j=1} \Delta \mathbf{X}_j \cdot \frac{cov(\mathbf{X})^{-1}}{f^{-2}} \bullet \Delta \mathbf{X}_j \tag{2}$$

$$\mathbf{w}[i] = \frac{\sum\limits_{k=1} e^{-\mathbf{E}_k}}{\sqrt{det\left[2\pi \cdot cov(\mathbf{X}) \cdot f^2\right]}}\bigg|_i^n \qquad \{i \in \mathbb{N} : i \le n\} \tag{3}$$

Figure 5 was divided into four-hour blocks to identify any time-of-day dependence. It is observed that regardless of the time of day most of the data continued to fall along the 1:1 line, as indicated in the density plots, for CAPABLE and BAO-Tower, while the Golden site displays some disruption in the 16:00 – 19:59 panel. The source of this discrepancy is currently unknown. However, it has become clear that the meteorology at the Golden site is different from that observed at CAPABLE and BAO. It is suggested that this difference is primarily driven by orographic perturbations and the Golden site being located atop a mesa, both of which may inhibit formation of stable ABL and ML (Bossert et al. (1989); Bossert and Cotton (1994); Tripoli and Cotton (1989)).

Ceilometer-derived MLH values have application as model a-priori inputs that have been averaged down to one-hour resolution. The impact of the filtering criteria and re-sampling to one-hour resolution throughout the day is seen in Fig. 5 (panels C, F, I), and Table 1 presents statistics on the aggregate analysis. While the aggregate coefficients of correlation and line-of-best-fit (LOBF) equations do not change substantially after re-sampling to one-hour blocks, the scatter is dramatically reduced (Fig. 5 panels C, F, I). This is likely due to the scatter being evenly distributed about the 1:1 line and the majority of data points falling along the 1:1 line as observed in the data-density panels of Fig. 5.

From the current analysis we conclude that the majority of the variability was driven by local atmospheric fluctuations and events that cannot be readily accounted for within the algorithms, and that on the timescales relevant to model inputs and atmospheric variations, there is no significant difference between the BLView/Python-collected datasets when processed on a common algorithm. Findings presented in Section 5.1.3 further support this conclusion.

5.1.3 MLH Algorithm Dependence

In the previous section it was demonstrated that the data collection method (i.e. raw serial logger vs. BLView) has little impact on the derived MLH values when the two datasets are
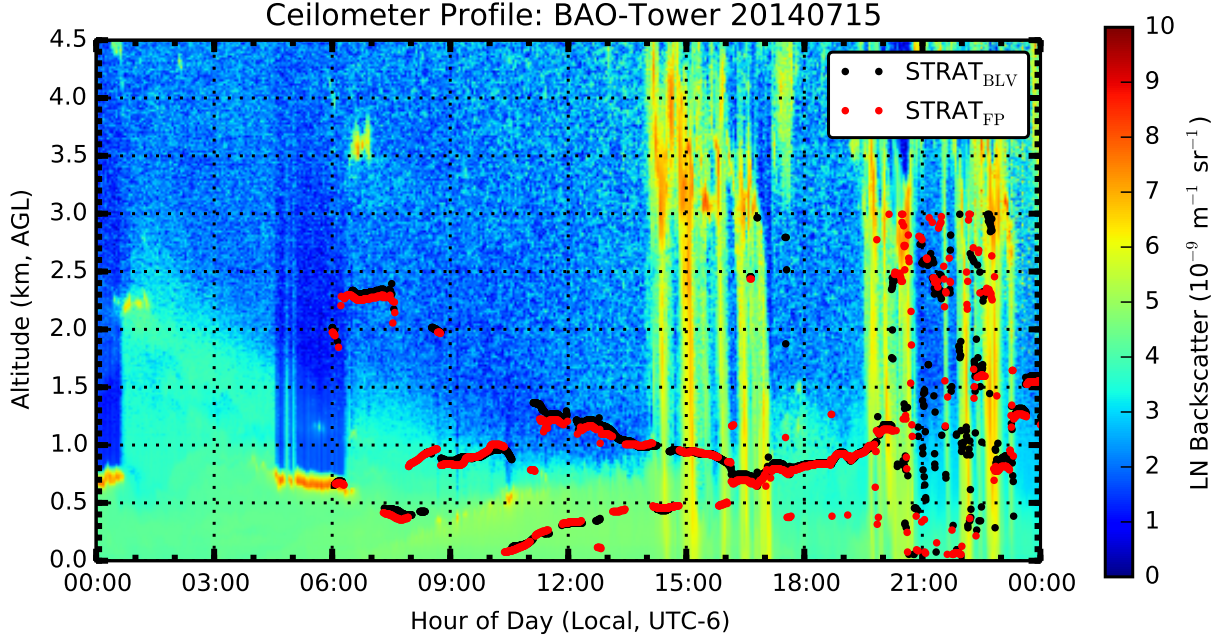
Figure 4. Backscatter curtain plot with STRAT-derived MLH values from the BLView (BLV) and Python (FP) collection methods.

	R	LOBF	$\langle FP - BL \rangle$ (km)
CAPABLE 5-min	0.87	$y = 0.913 \cdot x + 0.11$	-0.02 (1.4)
CAPABLE 1-Hr	0.87	$y = 0.925 \cdot x + 0.11$	-0.03 (2.7)
BAO 5-min	0.76	$y = 0.817 \cdot x + 0.25$	-0.08 (9.1)
BAO 1-Hr	0.77	$y = 0.814 \cdot x + 0.32$	-0.14 (15.1)
Golden 5-min	0.72	$y = 0.777 \cdot x + 0.30$	-0.08 (8.1)
Golden 1-Hr	0.77	$y = 0.792 \cdot x + 0.35$	-0.14 (13.0)

Table 1. Summary of aggregate statistics for the Python-collected/STRAT-processed and the BLView-collected/STRAT-processed MLH estimates (y and x, respectively). Filtering criteria were applied to both datasets. Values in parentheses indicate percent of the difference value with respect to the BLView-derived MLH.

processed on a common algorithm (here, STRAT). It remains to be seen how the two datasets compare when processed in different algorithms. Whereas collection methods were compared in the previous section, the algorithms will be compared here. Data collected by the Python script were processed using the STRAT algorithm, and are compared with data collected, and processed, by BLView.

Figure 6 presents scatter plots similar to Fig. 5, but with data collected and processed by differing means. It is observed that the majority of data continued to fall along the 1:1 line, as attested by the density plots, and that much of the scatter is caused by short-term variability. However, in contrast to Fig. 5, the scatter is neither as evenly distributed nor as tightly grouped about the 1:1 line. The STRAT-derived MLHs were generally lower than those calculated in BLView (according to slopes) at all sites, while the aggregate mean
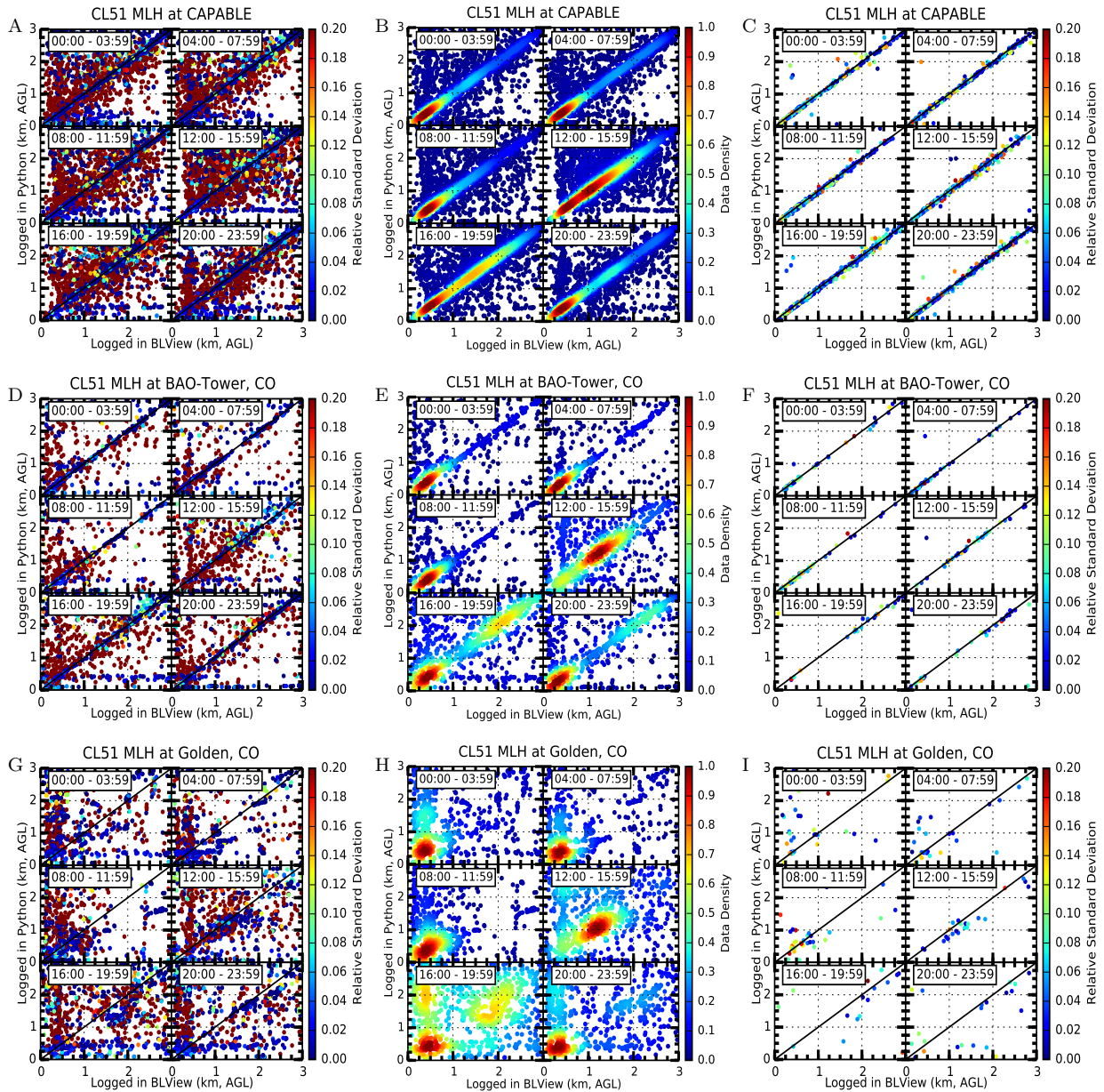
difference shows the opposite for the Colorado sites (Table 2), which is likely being driven by outliers.

It is readily observed that the agreement between the two datasets is less than when a common algorithm was employed (Table 2). Despite the increased scatter, there remains a significant subset of data that lies along the 1:1 line. As a test for how well the data sit along the 1:1 line the R and LOBF values were re-calculated with weights applied according to Eq. (3) (i.e. weighting according to data density). Therefore, points that have more data points surrounding them received more weight, while more isolated points received less weight. Weighted coefficients of correlation were calculated by Eq. (4) where variables with a w subscript indicate weighted means. Weighted regressions were performed by simultaneously solving the modified normal equations of regression as shown in Eqs. 5 and 6 with weighting factors applied.

$$R = \frac{\sum\limits_{i=1}^{N} w_i \cdot (x_i - \bar{x}_w) \cdot (y_i - \bar{y}_w)}{\sqrt{\sum\limits_{i=1}^{N} w_i \cdot (x_i - \bar{x}_w)^2 \cdot \sum\limits_{i=1}^{N} w_i \cdot (y_i - \bar{y}_w)^2}} \quad (4)$$

$$m = \frac{N \sum\limits_{i=1}^{N} w_i x_i y_i - \left(\sum\limits_{i=1}^{N} w_i x_i\right)\left(\sum\limits_{i=1}^{N} w_i y_i\right)}{N \sum\limits_{i=1}^{N} w_i x_i^2 - \left(\sum\limits_{i=1}^{N} w_i x_i\right)^2} \quad (5)$$

Figure 5. Correlation plots for data collected at the three sites under study. At all sites the data have been collected by both the Python script and BLView, and subsequently processed in STRAT. The center-column plots show the data density to better understand the distribution within the scatter plots. Data were averaged to five-minute resolution, without application of filtering criteria (left and center columns), and averaged to one-hour resolution with application of filtering criteria (right column).
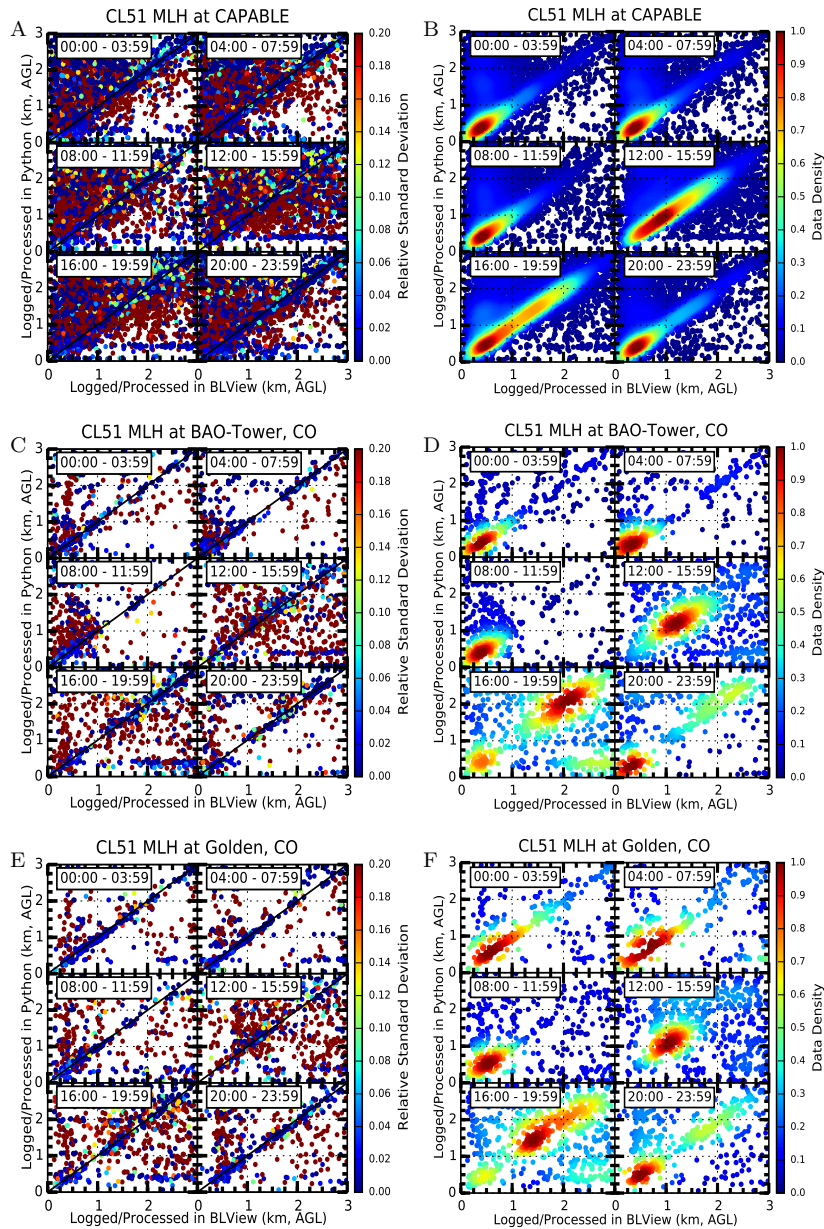
Figure 6. Correlation plots for data collected at the three sites under study. At all sites the data were collected by, and processed in, Python/STRAT and BLView/BLView. The right-hand plots show the data density to better understand the distribution within the scatter plots. Data were averaged to five-minute resolution, without application of filtering criteria.

	R	Line of best fit	$\langle FP - BL \rangle$ (km)	R_w	$LOBF_w$	$\langle FP - BL \rangle_w$ (km)
CAPABLE 5-min	0.47	$y = 0.499 \cdot x + 0.70$	-0.24 (26.8)	0.836	$y = 0.986 \cdot x + 0.70$	-0.02 (3.8)
CAPABLE 1-Hr	0.48	$y = 0.467 \cdot x + 0.74$	-0.23 (24.4)	0.799	$y = 0.997 \cdot x + 0.74$	-0.08 (12.5)
BAO 5-min	0.43	$y = 0.374 \cdot x + 0.65$	0.04 (3.4)	0.789	$y = 0.905 \cdot x + 0.65$	-0.01 (0.9)
BAO 1-Hr	0.39	$y = 0.305 \cdot x + 0.72$	0.17 (13.1)	0.740	$y = 0.881 \cdot x + 0.72$	-0.01 (1.2)
Golden 5-min	0.24	$y = 0.193 \cdot x + 0.90$	0.25 (17.7)	0.541	$y = 0.629 \cdot x + 0.90$	0.09 (10.3)
Golden 1-Hr	0.12	$y = 0.086 \cdot x + 1.12$	0.39 (23.6)	0.316	$y = 0.361 \cdot x + 1.12$	0.20 (16.1)

Table 2. Summary of statistics for the Python-collected/STRAT-processed and the BLView-collected/BLView-processed MLH estimates. Values in parentheses indicate percent of the difference value with respect to the BLView-derived MLH, and the w subscript indicates a weighting function was applied. Filtering criteria were applied to all datasets.

	R	Line of best fit	$\langle FP - BL \rangle$ (km)	R_w	$LOBF_w$	$\langle FP - BL \rangle_w$ (km)
CAPABLE 5-min	0.54	$y = 0.553 \cdot x + 0.61$	-0.19 (21.2)	0.91	$y = 0.975 \cdot x + 0.61$	-0.03 (4.2)
CAPABLE 1-Hr	0.54	$y = 0.519 \cdot x + 0.64$	-0.18 (18.5)	0.87	$y = 0.973 \cdot x + 0.64$	-0.06 (9.8)
BAO 5-min	0.41	$y = 0.326 \cdot x + 0.58$	0.14 (13.1)	0.78	$y = 0.843 \cdot x + 0.58$	-0.01 (1.4)
BAO 1-Hr	0.31	$y = 0.232 \cdot x + 0.65$	0.32 (25.5)	0.58	$y = 0.573 \cdot x + 0.65$	-0.09 (13.0)
Golden 5-min	0.25	$y = 0.184 \cdot x + 0.83$	0.36 (24.5)	0.49	$y = 0.484 \cdot x + 0.83$	0.16 (16.6)
Golden 1-Hr	0.14	$y = 0.101 \cdot x + 0.96$	0.55 (32.8)	0.27	$y = 0.229 \cdot x + 0.96$	0.37 (29.1)

Table 3. Summary of statistics for the BLView-collected/STRAT-processed and the BLView-collected/BLView-processed MLH estimates. Values in parentheses indicate percent of the difference value with respect to the BLView-derived MLH, and the w subscript indicates a weighting function was applied. Filtering criteria were applied to all datasets. Herein, the comparison is limited strictly to the MLH algorithms.

$$b = \frac{\sum\limits_{i=1}^{N} w_i y_i - m \sum\limits_{i=1}^{N} w_i x_i}{N} \qquad (6)$$

These weighted statistics are not included to suggest that the agreement has actually improved (R), nor do they suggest improved predictability (LOBF). Rather, the improved R values and slopes reflect the degree to which the data are predominantly centered about the 1:1 line to the exclusion of other regions. As an example, despite weighting, the improvement in the Golden regressions are notably less than the other two sites. This is likely due to more spread in the data, thus mitigating the influence of the points along the 1:1 line on the regression analyses. Therefore, we can conclude that the preponderance of data collected at CAPABLE and BAO-Tower fall nearer the 1:1 line when processed through the different algorithms as compared to the data collected at the Golden site. Further, despite the majority of data falling nearer the 1:1 line for these two sites, there remain influences that neither the STRAT configuration nor the current filter methodology can account for, which is likely driving the poor correlation as compared to Table 1. This is likely the product of how the differing algorithms handle atmospheric interferential events (e.g. precipitation, fog, etc.). Application of a filtering methodology to account and remove these events will be the subject of future study.

Finally, the analysis was repeated by using STRAT to process backscatter data that was collected by BLView for

comparison against the BLView-collected/processed product. It was concluded in Sec. 5.1.2 that the data collection method had little influence on the MLH estimation when both datasets were processed on a common algorithm (here, STRAT). Based on that conclusion, it would be expected that the current comparison would be similar to the previous comparison as summarized in Table 2. This is, in fact, what was observed. The aggregate statistics for the BLView-collected, STRAT-processed vs. BLView-collected/processed intercomparison are presented in Table 3, wherein we see similarity with Table 2. This further supports the conclusion that data collection methods (including application of calibration factors) play relatively less role in identifying a qualitative gradient within the profile as compared to the choice of MLH algorithm. Indeed, it can be concluded that choice and configuration of the algorithm is critical and that, for network intercomparisons, all networked LIDAR systems should have their data processed on a common algorithm.

5.2 Sonde Intercomparison

Meteorological soundings have been a staple for profiling the atmosphere and deriving ABL heights for decades. These ABL heights are typically derived using potential temperature (e.g. using the Heffter criteria) or through analyzing skew-t-log-p plots that implement potential temperature, both of which are different from the gradient-based MLH algorithms implemented herein. As ABL data are typically

used in chemical transport models, it is necessary to determine how these MLH data compare to the sonde-derived ABL data collected at the three measurement locations.

Since the sondes capture an ephemeral snapshot of the atmosphere's current conditions, and traverse several kilometers in the horizontal direction due to winds, the ceilometer data were averaged over thirty minutes for comparison. Additionally, each measurement can be impacted by different atmospheric phenomena which can affect the measurements in different ways, which in turn can affect the comparison of the measurements. A radio sonde can be impacted by local updrafts or down-drafts, and result in ABL estimates higher or lower than the true MLH. The CL51 MLH is sensitive to the backscatter gradient, so if there are additional aerosol layers just above the MLH, the contrast between the aerosol layers may not be strong enough for the CL51 to identify each layer or the correct altitude of the MLH.

Correlation plots for the CL51 MLH compared to sonde ABL are shown in Fig 7. For all coincidence times, the CAPABLE site showed the best correlations between the CL51 and radio sondes. The correlation for the CL51 versus all the radio sondes (N=25) at the CAPABLE site was R=0.82, with a similar correlation R=0.83 (N=22) when the filtering criteria were implemented. For daytime data, the CAPABLE site contained 2 early morning radio sondes (before 10:00 local time), with all other radio sondes launched between 10:00 and 16:00 local time. By the late morning, ≈10:00 local time, the vertical dispersion of aerosols due to turbulent mixing has likely resulted in a well-mixed boundary layer, so the ABL and MLH coincide in elevation, which is evident in 7 A, with many of the data points falling close to the 1:1 line.

Radio sonde data collected at the BAO-Tower site showed lower correlations than the CAPABLE site (unfiltered R=0.63; N=16 and filtered R=0.58; N=14), while the Golden site correlations (unfiltered R=-0.28; N=12) appear to be strongly impacted by 2 morning radio sonde launches, which is a transition period when the boundary layer is experiencing rapid growth. Upon applying the filtering criteria, the 2 early morning data points were removed, resulting in a much improved correlation (filtered R=0.74; N=10) for the Golden site. This appears to indicate the CL51 may have difficulty capturing an accurate MLH during rapidly changing conditions, such as early morning and late evening transition periods in a clean atmosphere.

It is somewhat surprising that filtered correlation for the Golden site is better than the filtered result for the BAO-Tower site, given the BAO-Tower site is situated further to the east of the mountain range, at the start of the High Plains, which is less influenced by very local geographic perturbations, and that a similar relationship is not observed in the CL51 intercomparisons (e.g. Tables 1, 2, and 3). As a check on the radio sonde potential temperate profiles, we plotted the potential temperature data from the NASA P-3B aircraft spirals conducted over the Golden and Erie sites, Figs. 8 and 9. These spirals are coincident with the launch of the radio

sondes from the sites. Also plotted in Figs. 8 and 9 is the coincident CL51 backscatter profile. The agreement between the radio sonde and P-3B aircraft profiles is good, indicating that the potential temperature within the aircraft spiral radius is consistent with that of the radio sonde. These figures show agreement between the potential temperature ABL and CL51 MLH by identifying the same first major gradient in the MLH data on certain days.

Overall, all 3 sites show a good correlation between the CL51 and radio sonde data, with MLH/ABL estimates from the radio sondes being, on average, higher than the CL51 MLH (200 m (13%), 390 m (15%), -240 m (9%) for CAPABLE, BAO-Tower, and Golden respectively) as indicated in the linear regression lines plotted in Fig. 7, with the exception being the unfiltered results for Golden.

5.3 MPL Intercomparison

The MPL instrument was collocated with the CL51 stationed at the NREL site in Golden, CO. Being a LIDAR instrument, it profiles the atmosphere similar to the CL51 with the major differences being hardware. The MPL emits on the 532 nm line (6 – 10 μJ pulsed at 2.5 kHz, with 150 m overlap and 30 m vertical resolution), while the CL51 emits on the 910 line (2.5 nJ pulsed at 6.5 kHz, with 10 m overlap and 10 m vertical resolution). The primary difference between the two instruments is the emission wavelength, causing the instruments to differ in sensitivity with respect to particle size and geometry. Therefore, it is feasible that the two instruments observed "different" atmospheres in a quantitative manner (e.g. aerosol optical thickness). However, if the ML is well mixed, then the general particle distribution and gradient will be the same, making the two inter-comparable.

It is seen in Fig. 10 that the agreement between the two instruments and algorithms (BLView processing CL51, UMBC algorithm processing MPL profiles) is promising, with most data points falling along the 1:1 line. The low correlation is partly driven by the invariability in one instrument as compared to the other at lower MLH values (i.e. ≤ 500 m). Removal of data below 500 m improves the coefficient of correlation to 0.368, 0.512, and 0.390 respectively. Similar to the algorithm comparison, much of the variability between the two instruments and algorithms occurs during events which inhibit a reliable MLH estimate being made (as seen in Fig. 11). This is somewhat surprising considering the low coefficients of correlation initially observed.

The most commonly used statistical techniques used for comparing two datasets depend on two key assumptions: data being normally distributed and homoscedastic. The CL51 and MPL MLH 5-minute averaged datasets were confirmed to be non-normal via the Kolmogorov-Smirnov test and passed Levene's test for homoscedasticity (p-value 0.39). Therefore, determination of similarity between the two corresponding probability distributions was performed using the two-sample Kolmogorov-Smirnov test. It was determined
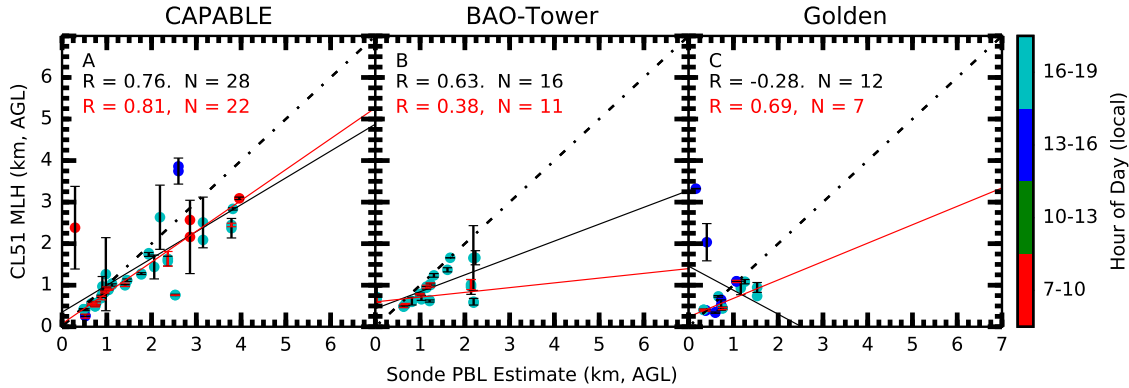
Figure 7. Correlation plots for CL51 MLH and sonde-derived ABL estimates. Statistics data in black text are for the entire data set, while the red text represents the filtered dataset. MLH values were calculated in BLView.

that the 5-minute averaged MPL and CL51 datasets are statistically different ($p \ll 0.01$), regardless of filtering and averaging. However, when considering 1-hour averaged data that were filtered to remove data with large relative standard deviations (i.e. ≥ 0.20) and MLH ≤ 0.5 km, the two datasets were statistically indistinguishable (p 0.8). While we cannot account for the variability induced by these low-altitude MLH values it is quite clear that they are significantly influencing the intercomparison. Given that this is the first intercomparison of these two instruments and algorithms it is not surprising that a significant difference in this regime was identified.

6 Conclusions

A CL51-focused intercomparison of different ABL/MLH methodologies was performed at three different sites, which experience different meteorological, aerosol, and emission conditions. The CL51 MLH results were compared with ABL from radio sondes at all three locations; as well as an MPL at the Golden, CO site.

Two collection methods and processing algorithms were tested for the CL51 MLH calculation. It was determined that the data-collection method plays an insignificant role in MLH estimation when the datasets are processed on a common algorithm. Further, it was shown that the choice of processing algorithm plays a significant role in MLH estimation. We recommend that, for ceilometer and lidar networks, a common MLH processing algorithm be employed.

A total of 53 potential temperature profiles from radio sondes were used to evaluate the CL51. While the 53 radio sondes were spread across 3 sites, this represents a robust data set of soundings. Overall, the radio-sonde-derived ABL was higher than the CL51 MLH (e.g. Figure 7). Comparison of MLH from the CL51 versus radio sondes show the CL51 performed best at the CAPABLE research site (non-filtered R=0.82, filtered R=0.83), a coastal site primarily influenced by a combination of sulfate and marine aerosols. Both the Golden and BAO-Tower site show a good correlation between the CL51 and radio sondes (Golden filtered R=0.74, BOA non-filtered R=0.63, filtered R=0.58) with 2 early morning radio sondes at the Golden site strongly influencing the non-filtered correlation (R=-0.28). These 2 radio sondes measured a very shallow boundary layer, < 500 m, while the CL51-identified the MLH above 2 km, which was likely due to residual aerosol layers aloft. The lower correlations at the Colorado sites (Golden and BAO) were likely due to the sites proximity to the Rocky Mountains. Complex atmospheric flow patterns, which are driven by the Rocky Mountains to the west of the Front Range area, can induce the formation of distinctive dynamic features such as up and downslope flows Bossert et al. (1989); Bossert and Cotton (1994); Tripoli and Cotton (1989). With the Golden site being along the slope of the mountains and on a mesa, the Golden site would likely experience up and downslope flows versus BOA. Such local orographic influences can impact or challenge the well-mixed assumption required to compare thermodynamic ABL measured via potential temperature and MLH measured via aerosol backscatter.

The results of the CL51 versus the UMBC MPL showed low correlation (R=0.3).However the majority of coincident MLH observations from both instruments were clustered around the 1:1 line in the regression plots. When data filtering criteria are applied the two data sets were statistically indistinguishable (p > 0.8). Additional analysis is planned to further explore the cause of the low correlation. However as can be seen in Figure 15, the MLH from the CL51 and MPL agree well when there is a well-defined MLH.

Finally, the analysis indicates the STRAT software can serve as a viable alternative to BLView and offers an advantage of use of a consistent MLH method with EARLINET.
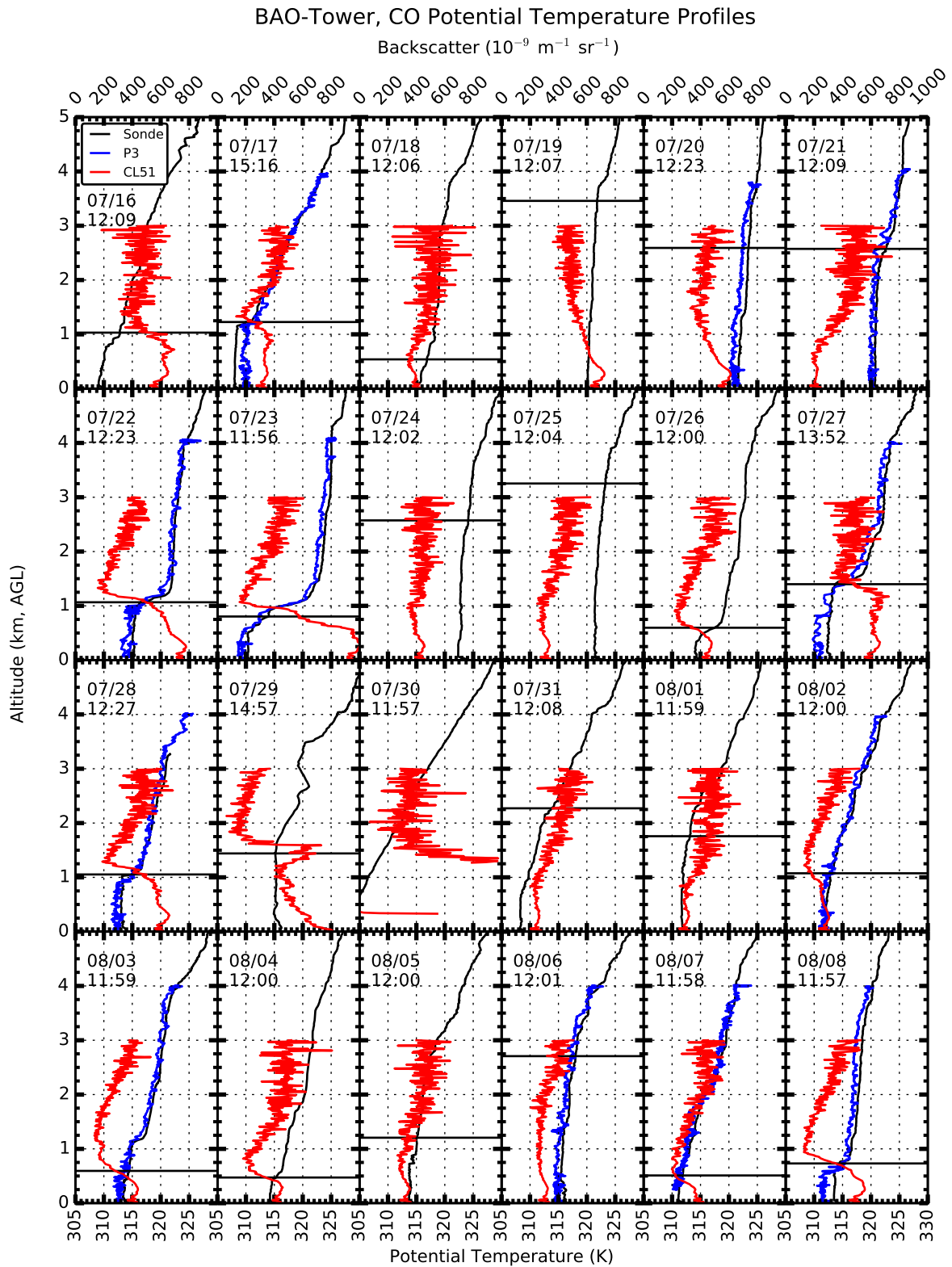
Figure 8. Potential temperature and Cl51 backscatter profiles collected at the BAO site. Horizontal lines indicate MLH as determined via BLView.
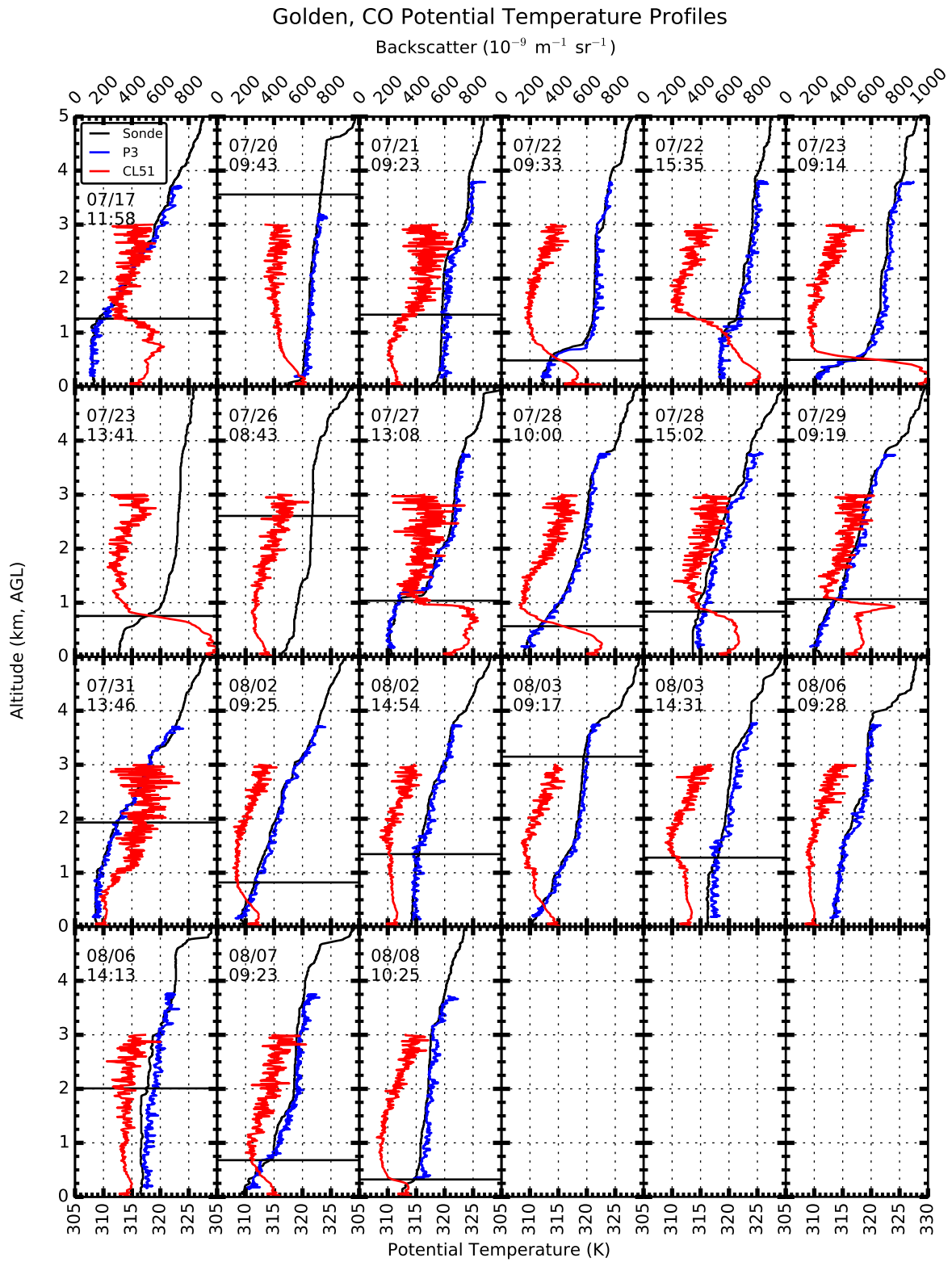
Figure 9. Potential temperature and Cl51 backscatter profiles collected at the Golden NREL site. Horizontal lines indicate MLH as determined via BLView.
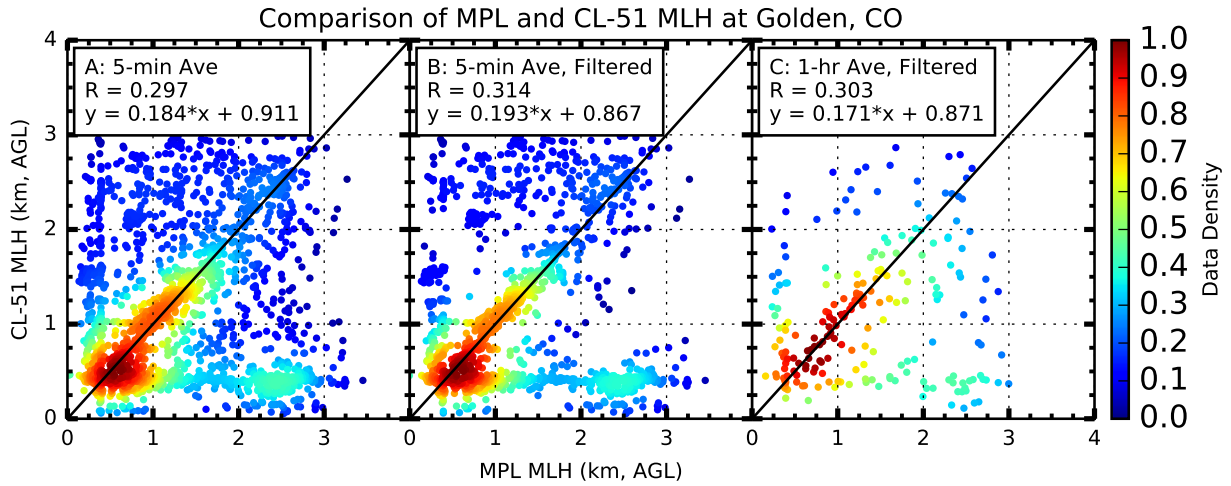
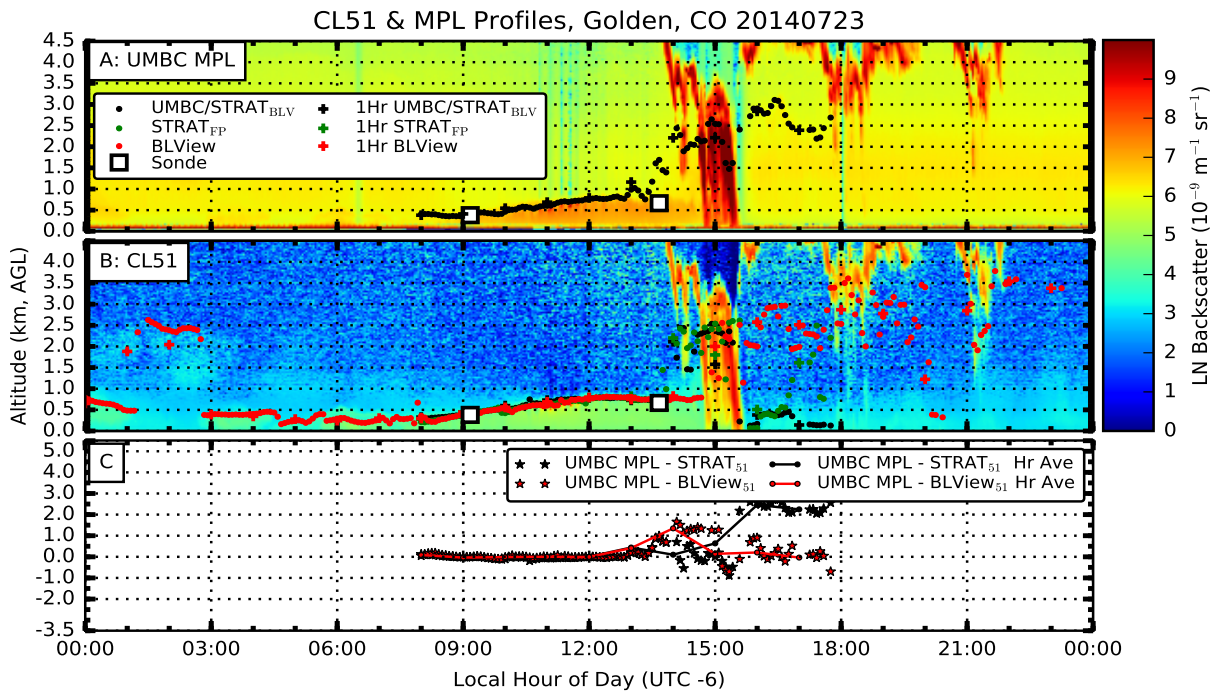**Figure 10.** Correlation and data-density plots for the CL51 and MPL MLH estimates from Golden, CO.

Figure 11. Comparison of the CL51 and UMBC MPL profiles for 23/27-July 2014 at the Golden, CO site.

References

Bossert, J. and Cotton, W.: Regional-Scale Flows in Mountainous Terrain 1. A Numerical and Observational Comparison, Monthly Weather Review, 122, 1449–1471, doi:10.1175/1520-0493(1994)122<1449:RSFIMT>2.0.CO;2, 1994.

Bossert, J., Sheaffer, J., and Reiter, E.: Aspects of Regional-Scale Flows in Mountainous Terrain, Journal of Applied Meteorology, 28, 590–601, doi:10.1175/1520-0450(1989)028<0590:AORSFI>2.0.CO;2, 1989.

Compton, J. C., Delgado, R., Berkoff, T. A., and Hoff, R. M.: Determination of Planetary Boundary Layer Height on Short Spatial and Temporal Scales: A Demonstration of the Covariance Wavelet Transform in Ground-Based Wind Profiler and Lidar Measurements, Journal of Atmospheric and Oceanic Technology, 30, 1566–1575, doi:10.1175/JTECH-D-12-00116.1, 2013.

Deshler, T., Mercer, J. L., Smit, H. G. J., Stubi, R., Levrat, G., Johnson, B. J., Oltmans, S. J., Kivi, R., Thompson, A. M., Witte, J., Davies, J., Schmidlin, F. J., Brothers, G., and Sasaki, T.: Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes, Journal of Geophysical Research-Atmospheres, 113, doi:10.1029/2007JD008975, 2008.

Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, Atmospheric Measurement Techniques, 7, 4463–4490, doi:10.5194/amt-7-4463-2014, 2014.

Emeis, S. and Schäfer, K.: Remote sensing methods to investigate boundary-layer structures relevant to air pollution in cities, Boundary-Layer Meteorology, 121, 377–385, doi:10.1007/s10546-006-9068-2, 2006.

Emeis, S., Schäfer, K., and Muenkel, C.: Long-term observations of the urban mixing-layer height with ceilometers, in: 14TH International Symposium for the Advancement of Boundary Layer Remote Sensing, edited by Mann, J and Bingol, F and Courtney, M and Jorgensen, HE and Lindelow, P and Mikkelsen, T and Pena, A and Sjoholm, M and Wagner, R, vol. 1 of *IOP Conference Series-Earth and Environmental Science*, doi:10.1088/1755-1307/1/1/012027, 14th International Symposium for the Advancement of Boundary Layer Remote Sensing, Tech Univ Denmark, Copenhagen, DENMARK, JUN 23-25, 2008, 2008a.

Emeis, S., Schäfer, K., and Muenkel, C.: Surface-based remote sensing of the mixing-layer height - a review, Meteorologische Zeitschrift, 17, 621–630, doi:10.1127/0941-2948/2008/0312, DACH 2007 Conference, Hamburg, GERMANY, SEP 10-14, 2007, 2008b.

Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G. P., Lolli, S., O'Dowd, C. D., Sauvage, L., Xueref-Remy, I., Wastine, B., and Feist, D. G.: Evaluation of Mixing-Height Retrievals from Automatic Profiling Lidars and Ceilometers in View of Future Integrated Networks in Europe, Boundary-Layer Meteorology, 143, 49–75, doi:10.1007/s10546-011-9643-z, 15th International Symposium for the Advancement of Boundary-Layer Remote Sensing (ISARS), Univ Versailles St-Quentin-en-Yvelines, Inst Pierre Simon Laplace (IPSL), Paris, FRANCE, JUN 28-30, 2010, 2012.

Heffter, J.: Air Resources Laboratories atmospheric transport and dispersion model (ARL-ATAD), NOAA, 1980.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., and Abuhassan, N.: NO2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, Journal of Geophysical Research-Atmospheres, 114, doi:10.1029/2009JD011848, 2009.

Hirsikko, A., O'Connor, E. J., Komppula, M., Korhonen, K., Pfuller, A., Giannakaki, E., Wood, C. R., Bauer-Pfundstein, M., Poikonen, A., Karppinen, T., Lonka, H., Kurri, M., Heinonen, J., Moisseev, D., Asmi, E., Aaltonen, V., Nordbo, A., Rodriguez, E., Lihavainen, H., Laaksonen, A., Lehtinen, K. E. J., Laurila, T., Petaja, T., Kulmala, M., and Viisanen, Y.: Observing wind, aerosol particles, cloud and precipitation: Finland's new ground-based remote-sensing network, Atmospheric Measurement Techniques, 7, 1351–1375, doi:10.5194/amt-7-1351-2014, 2014.

Johnson, B., Oltmans, S., Vomel, H., Smit, H., Deshler, T., and Kroger, C.: Electrochemical concentration cell (ECC) ozonesonde pump efficiency measurements and tests on the sensitivity to ozone of buffered and unbuffered ECC sensor cathode solutions, Journal of Geophysical Research-Atmospheres, 107, doi:10.1029/2001JD000557, 2002.

Knepp, T., Pippin, M., Crawford, J., Chen, G., Szykman, J., Long, R., Cowen, L., Cede, A., Abuhassan, N., Herman, J., Delgado, R., Compton, J., Berkoff, T., Fishman, J., Martins, D., Stauffer, R., Thompson, A., Weinheimer, A., Knapp, D., Montzka, D., Lenschow, D., and Neil, D.: Estimating surface NO2 and SO2 mixing ratios from fast-response total column observations and potential application to geostationary missions, Journal of Atmospheric Chemistry, pp. 1–26, doi:10.1007/s10874-013-9257-6, http://dx.doi.org/10.1007/s10874-013-9257-6, 2015.

Lamsal, L. N., Martin, R. V., van Donkelaar, A., M., S., Celarier, E. A., Bucsela, E., Dunlea, E. J., and Pinto, J. P.: Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument, Journal of Geophysical Research-Atmospheres, 113, doi:10.1029/2007JD009235, 2008.

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., and Knepp, T. N.: Evaluation of OMI operational standard NO2 column retrievals using in situ and surface-based NO2 observations, Atmospheric Chemistry and Physics, 14, 11 587–11 609, doi:10.5194/acp-14-11587-2014, 2014.

Marsik, F., Fischer, K., McDonald, T., and Samson, P.: Comparison of Methods for Estimating Mixing Height used during the 1992 Atlanta Field Intensive, Journal of Applied Meteorology, 34, 1802–1814, doi:10.1175/1520-0450(1995)034<1802:COMFEM>2.0.CO;2, International Specialty Conference on Regional Photochemical Measurement and Modeling Studies, SAN DIEGO, CA, NOV 08-12, 1993, 1995.

Martin, R. V.: Satellite remote sensing of surface air quality, Atmospheric Environment, 42, 7823–7843, doi:10.1016/j.atmosenv.2008.07.018, 2008.

Martins, D. K., Stauffer, R. M., Thompson, A. M., Knepp, T. N., and Pippin, M.: Surface ozone at a coastal suburban site in 2009 and 2010: Relationships to chemical and meteorological processes, Journal of Geophysical Research-Atmospheres, 117, doi:10.1029/2011JD016828, 2012.

Miloshevich, L., Paukkunen, A., Vomel, H., and Oltmans, S.: Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements, Journal of Atmospheric and Oceanic Technology, 21, 1305–1327, doi:10.1175/1520-0426(2004)021<1305:DAVOAT>2.0.CO;2, 2004.

Morille, Y., Haeffelin, M., Drobinski, P., and Pelon, J.: STRAT: An automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data, Journal of Atmospheric and Oceanic Technology, 24, 761–775, doi:10.1175/JTECH2008.1, 2007.

Nash, J., Smout, R., Oakley, T., Pathack, B., and Kurnosenko, S.: WMO Intercomparison of Radiosonde Systems Vacoas, Mauritius, 2-25 February 2005, Report WMO/TD-No. 1303, World Meteorol. Organ., Geneva, Switzerland, 2006.

Nash, J., Oakley, T., Vömel, H., and Wei, L.: WMO Intercomparison of High Quality Radiosonde Systems Yangjiang, China, 12 July - 3 August 2010, Report WMO/TD-No. 1580, World Meteorol. Organ., Geneva, Switzerland, 2011.

Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linne, H., Ansmann, A., Boesenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmospheric Measurement Techniques, 7, 2389–2409, doi:10.5194/amt-7-2389-2014, 2014.

Petritoli, A., Bonasoni, P., Giovanelli, G., Ravegnani, F., Kostadinov, I., Bortoli, D., Weiss, A., Schaub, D., Richter, A., and Fortezza, F.: First comparison between ground-based and satellite-borne measurements of tropospheric nitrogen dioxide in the Po basin, Journal of Geophysical Research-Atmospheres, 109, doi:10.1029/2004JD004547, 2004.

Piters, A. J. M., Boersma, K. F., Kroon, M., Hains, J. C., Van Roozendael, M., Wittrock, F., Abuhassan, N., Adams, C., Akrami, M., Allaart, M. A. F., Apituley, A., Beirle, S., Bergwerff, J. B., Berkhout, A. J. C., Brunner, D., Cede, A., Chong, J., Clemer, K., Fayt, C., Frieß, U., Gast, L. F. L., Gil-Ojeda, M., Goutail, F., Graves, R., Griesfeller, A., Grossmann, K., Hemerijckx, G., Hendrick, F., Henzing, B., Herman, J., Hermans, C., Hoexum, M., van der Hoff, G. R., Irie, H., Johnston, P. V., Kanaya, Y., Kim, Y. J., Baltink, H. K., Kreher, K., de Leeuw, G., Leigh, R., Merlaud, A., Moerman, M. M., Monks, P. S., Mount, G. H., Navarro-Comas, M., Oetjen, H., Pazmino, A., Perez-Camacho, M., Peters, E., du Piesanie, A., Pinardi, G., Puentedura, O., Richter, A., Roscoe, H. K., Schönhardt, A., Schwarzenbach, B., Shaiganfar, R., Sluis, W., Spinei, E., Stolk, A. P., Strong, K., Swart, D. P. J., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Whyte, C., Wilson, K. M., Yela, M., Yilmaz, S., Zieger, P., and Zhou, Y.: The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results, Atmospheric Measurement Techniques, 5, 457–485, doi:10.5194/amt-5-457-2012, 2012.

Scarino, A. J., Obland, M. D., Fast, J. D., Burton, S. P., Ferrare, R. A., Hostetler, C. A., Berg, L. K., Lefer, B., Haman, C., Hair, J. W., Rogers, R. R., Butler, C., Cook, A. L., and Harper, D. B.: Comparison of mixed layer heights from airborne high spectral resolution lidar, ground-based measurements, and the WRF-Chem model during CalNex and CARES, ATMOSPHERIC CHEMISTRY AND PHYSICS, 14, 5547–5560, doi:10.5194/acp-14-5547-2014, 2014.

Schäfer, K., Wagner, P., Emeis, S., Jahn, C., Münkel, C., and Suppan, P.: Mixing layer height and air pollution levels in urban area, in: Remote Sensing of Clouds and the Atmosphere XVII; and LIDAR Technologies, Techniques, and Measurements for Atmospheric Remote Sensing VIII, edited by Kassianov, EI and Comeron, A and Picard, RH and Schafer, K and Singh, UN and Pappalardo, G, vol. 8534 of *Proceedings of SPIE*, SPIE; SELEX GALILEO; THALES, doi:10.1117/12.974328, Conference on Remote Sensing of Clouds and the Atmosphere XVII - and Lidar Technologies, Techniques and Measurements for Atmospheric Remote Sensing VIII, Edinburgh, SCOTLAND, SEP 24-27, 2012, 2012.

Schäfer, K., Ling, H., Münkel, C., and Emeis, S.: Long-term study of air urban quality together with mixing layer height, in: Remote Sensing of Clouds and the Atmosphere XVIII; and Optics in Atmospheric Propagation and Adaptive Systems XVI, edited by Comeron, A and Kassianov, EI and Schafer, K and Stein, K and Gonglewski, JD, vol. 8890 of *Proceedings of SPIE*, SPIE, doi:10.1117/12.2027045, Conference on Remote Sensing of Clouds and the Atmosphere XVIII; and Optics in Atmospheric Propagation and Adaptive Systems XVI, Dresden, GERMANY, SEP 23-26, 2013, 2013.

Schween, J. H., Hirsikko, A., Loehnert, U., and Crewell, S.: Mixing-layer height retrieval with ceilometer and Doppler lidar: from case studies to long-term assessment, Atmospheric Measurement Techniques, 7, 3685–3704, doi:10.5194/amt-7-3685-2014, 2014.

Scott, D.: Multivariate Density Estimation: Theory, Practice, and Visualization, A Wiley-interscience publication, Wiley, 1992.

Silverman, B.: Density Estimation for Statistics and Data Analysis, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis, 1986.

Smit, H.: Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, Report, World Meteorol. Organ., Geneva, Switzerland, 2013.

Sokol, P., Stachlewska, I. S., Ungureanu, I., and Stefan, S.: Evaluation of the boundary layer morning transition using the CL-31 ceilometer signals, Acta Geophysica, 62, 367–380, doi:10.2478/s11600-013-0158-5, 2014.

Spinhirne, J.: Micropulse LIDAR, IEEE Transactions on Geoscience and Remote Sensing, 31, 48–55, doi:10.1109/36.210443, 1993.

Stauffer, R. M., Morris, G. A., Thompson, A. M., Joseph, E., Coetzee, G. J. R., and Nalli, N. R.: Propagation of radiosonde pressure sensor errors to ozonesonde measurements, Atmospheric Measurement Techniques, 7, 65–79, doi:10.5194/amt-7-65-2014, 2014.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Atmospheric Sciences Library, Springer, 1988.

Tripoli, G. and Cotton, W.: Numerical Study of an Observed Orogenic Mesoscale Convective System 2. Analysis of Governing Dynamics, Monthly Weather Review, 117, 305–328,

doi:10.1175/1520-0493(1989)117<0305:NSOAOO>2.0.CO;2, 1989.

Welton, E., Voss, K., Gordon, H., Maring, H., Smirnov, A., Holben, B., Schmid, B., Livingston, J., Russell, P., Durkee, P., Formenti, P., and Andreae, M.: Ground-based lidar measurements of aerosols during ACE-2: instrument description, results, and comparisons with other ground-based and airborne measurements, Tellus Series B-Chemical and Physical Meteorology, 52, 636–651, doi:10.1034/j.1600-0889.2000.00025.x, 2000.

Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiss, A., Pappalardo, G., Schäfer, K., and Thomas, W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, Atmospheric Measurement Techniques, 7, 1979–1997, doi:10.5194/amt-7-1979-2014, 2014.