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

T. Travis N. Knepp^{1,2}, James J. Szykman^{3,4}, Russell Long³, Rachelle M. Duvall³, Jonathan Krug³, Melinda Beaver³, Kevin Cavender³, Keith Kronmiller⁵, Michael Wheeler⁵, Ruben Delgado⁶, Raymond Hoff⁶, Timothy Berkoff², Erik Olson⁷, Richard Clark⁸, Daniel Wolfe⁹, David Van Gilst¹⁰, and Doreen Neil²

¹Science Systems and Applications Inc., Hampton, Virginia 23666, USA
²NASA Langley Research Center, Hampton, Virginia 23681, USA
³US EPA, Research Triangle Park, Durham, North Carolina 27709, USA
⁴Currently assigned to NASA Langley Research Center, Hampton, Virginia 23681, USA
⁵Jacobs Technology Inc., Tullahoma, Tennessee 37388, USA
⁶Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, Maryland 21250, USA
⁷Space Science and Engineering Center, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA
⁸Department of Earth Sciences, Millersville University, Millersville, Pennsylvania 17551, USA
⁹NOAA/ESRL Physical Sciences Division, Boulder, Colorado 80305, USA

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

Abstract. An assessment of differing Differing boundary/mixed-layer height measurement methods was performed were assessed in moderately-polluted and clean environments, with a focus on the Vaisala CL51 ceilometer. This intercomparison was performed as part of ongoing measurements at the Chemistry And Physics of the Atmospheric Boundary Layer Experiment (CAPABLE) site in Hampton, VA_Virginia and during the 2014 Deriving Information on Surface Conditions from

- 5 Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) field campaign that took place in the Denver, CO area and around Denver, Colorado. We analyzed CL51 data that was were collected via two different methods (i.e. via the BLView software, which applied correction factors, and simple terminal emulation logging) to determine the impact of data collection methodology. Further, we evaluated the STRAT STRucture of the ATmosphere (STRAT) algorithm as an open-source alternative to BLView (NOTE: note that the current work presents an evaluation of the BLView and STRAT
- 10 algorithms and does not intend to act as a validation of either). A common filtering criteria was Filtering criteria were defined according to the <u>AMLHchange in mixed-layer height (MLH)</u> distributions for each instrument and algorithm and were applied throughout the analysis to remove high-frequency fluctuations from the MLH retrievals, and was applied throughout the analysis. Of primary interest was determining how the different data-collection methodologies and algorithms compare to each other and to radiosonde-derived boundary-layer heights when deployed as part of a larger instrument network. We determine
- 15 that data collection_determined that data-collection methodology is not as important as the processing algorithm , and that much of the algorithm differences may might be driven by impacts of local meteorology and precipitation events that pose algorithm difficulties. The results of this study show that for LIDAR-based a common processing algorithm is necessary for Light

Detection And Ranging (LIDAR)-based MLH intercomparisons, and for ceilometer-network operation , a common processing algorithm is necessary and that sonde-derived boundary layer heights are higher (10–15% at mid-day) than LIDAR-derived mixed-layer heights. We show that averaging the retrieved MLH to one-hour resolution (i.e. as necessary-1-hour resolution (an appropriate time scale for a priori data for-model initialization) significantly improved correlation between differing instruments and differing algorithms.

5

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 a timescale of 1 hour or less (?). The ABL can be defined via by a number of criteria depending on the particular interest (e.g. the thermo-

- 10 dynamic boundary layer, chemical boundary layer (CBL), aerosol mixed layer, etc.). Traditionally, the ABL has been). The ABL is typically 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, and coverage is limited. Conversely, mixed-layer heights (MLH), as calculated from backscatter LIDAR instruments Light Detection And Ranging (LIDAR) instruments, provide both excellent vertical and
- 15 temporal resolution. Typical analysis of LIDAR data involves identification of gradients within the aerosol profile (?), which is generally considered to be a marker for the MLH. With respect to air quality, the top of the ABL often acts like a lid within on 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.
- 20 ABL variability complicates quantitative determination of surface trace-gas levels from a remote-sensing platform (???????). Therefore, properly accounting for ABL variability from a continuous measurement system such as LIDAR will provide invaluable information to policy, health, modeling, and remote-sensing communities for applications sensitive to the vertical profiles of tracers (???). In 2009, the United States National Research Council highlighted planetary (atmospheric) boundary layer_ABL height as a high priority observation needed to address current inadequacies at the improve meso-scale for improved
- 25 predictions of air quality, short-range severe-weather forecasting, and regional climate modeling (?). More recently, The the National Plan for Civil Earth Observation (?) called out the need called for improved observation density and sampling of the boundary layer (?). In 2015, as part of the revisions to the ozone (O₃) National Ambient Air Quality Standard (NAAQS), EPA Standards, the U.S. Environmental Protection Agency (EPA) finalized a new requirement under the Photochemical Assessment Monitoring Stations (PAMS) program for the collection of continuous mixing layer height MLH observations. By 2019, the
- 30 PAMS program will involve the implementation of approximately fifty air-quality sites around in the United States , providing measurements of MLH on a continuous basis that provide continuous MLH.

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.

- 5 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 10 (??????). 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 (???). Herein is presented results from ? showed that intercomparison of ceilometer data is not a straight-forward endeavor. An intercomparison of ceilometer instrumentation was carried out in support of upcoming PAMS monitoring requirements. Results from an intercomparison of three backscatter LIDAR 's instruments from
- 15 the 2014 DISCOVER-AQ field campaign in Colorado (low aerosol load) and coincident sonde launches from the Chemistry and Physics of the Atmospheric Boundary Layer Experiment (CAPABLE) site at NASA's Langley Research Center (LaRC, moderate aerosol load) in Hampton, VA Virginia are presented herein.

2 Instrumentation

2.1 CL51

- The Vaisala (Vantaa, Finland) CL51 ceilometer is a single-wavelength (eye safe Class 1M InGaAs diode LASER-laser emitting at 910 \pm 10 nm, pulsed at 6.5 kHz with a 110 ns pulse width , and with average pulse power of 19.5 mW, with and an 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 (?????????). 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), all altitudes are with
- 25 <u>respect to ground level</u>) and up to 2 s temporal resolution (depending on the control software), though profiles are generally averaged over 16–36 s to improve the signal-to-noise (see Sec. ratio (see section 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 bandsthese ceilometers.

30 <u>ceilometers operating at the stated wavelengths</u> experience water vapor interference, thereby <u>mitigating lessening</u> their utility in retrieval of aerosol optical properties. However, the interference on aerosol profile and MLH estimation is negligible (?).

Two <u>CL51's CL51s</u> were deployed as part of the 2014 DISCOVER-AQ mission in Colorado (Golden, and Erie, <u>COColorado</u>). Before and after deployment, these ceilometers were <u>stationed at CAPABLE set up to continually collect data at the CAPABLE</u> site and the EPA Ambient <u>air</u> Air Innovative Research Site (AIRS) in Durham, <u>NC, continually collecting dataNorth Carolina</u>. The ceilometers were collocated with <u>meteorological sonde</u> (met-sonde) launch sites during the DISCOVER-AQ campaign and at <u>CAPABLE the CAPABLE site</u>, allowing a direct intercomparison of the sonde and LIDAR ABL/MLH methodologies.

5 Furthermore, during the DISCOVER-AQ campaign the ceilometers were collocated with other LIDAR instruments. Intercomparisons are presented belowin section 5.

2.1.1 Ceilometer Full-profile Collection

The BLView software not only provides Vaisala standard MLH retrieval is based on a proprietary wavelet/gradient technique built into the logging/analysis software BLView. The BLView software provides not only logging and data analysis (e.g. MLH

- 10 and cloud-height estimates) , but also provides data-logging/but also archiving capability. While the CL51 reports backscatter up to 15.4 km, BLView truncates the profile data collection_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 precluding 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
- 15 the standard data-collection software was developed and implemented.

Data transmission from the ceilometer to the logging computer was achieved over a simple by splitting an RS-232 connection that can be split into two ports on the logging computer; one port logging to BLView , and 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 ealibration coefficients proprietary calibration coefficients that are

20 <u>built into the BLView software to the logged data</u>. However, as shown in subsequent sections, this impacts neither the MLH estimates nor the general profile shape substantially.

2.2 MPLMicropulse LIDAR

Elastic LIDAR observations were performed using a Sigma Space (Lanham, Maryland) Micropulse LIDAR (MPL), previously described in by ? and ?. Briefly, the MPL transmitter consists of an eye-safe Nd:YLF laser emitting at 527 nm, and pulsed at 2.5

- kHz and an average with a pulse power of $6 10 \mu J_{\tau}$. It 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 minutesto 15-min. The receiver consists of a 178 mm telescope that collects the backscattered light, which is then focused onto a photon counting silicon avalanche photodiode (APD). The APD output is recorded by a field programmable gate array (FPGA) data system that enables display and storage of range dependent averaged average count rates on a laptop computer. The raw data are converted to aerosol attenuated
- 30 backscatterby taking into account instrumental factors that include corrections for, correcting for instrumental factors such as detector dead time, geometrical overlap, background subtraction, and range-squared normalization. Recorded LIDAR profiles have temporal and vertical resolution of one minute 1 min and 30 meters, respectively. The m, respectively, as set by the UMBC team for the DISCOVER-AQ campaign. MPL is used for continuous recording of aerosol profiles , and optical properties, and calculating MLH values.



Figure 1. Backscatter curtain plot collected on 10-June -2015 when smoke from a Canadian forest fire was transported over the CAPABLE site. The smoke is observed by increased backscatter in the 2500 - 4000 m range.

2.3 Meteorological/Ozone Sondes

The traditional method of identifying the ABL is using meteorological sondes. A meteorological sonde (herein referred to as sondessonde/radiosonde) radiosonde) is the conventional method for measuring temperature, pressure, and humidity throughout

- 5 the atmosphere, and characterizing the ABL. Radiosondes were used to identify steep gradients within the potential temperature (theta) profile (Fig. 2 A) as identified by the Heffter criteria (??), which is shown in Eqs. (1) and (2) where Θ is potential temperature in Kelvin, Z is altitude in meters, and Θ_{top} and Θ_{base} refer to the potential temperature at the top and bottom of the proposed inversion layer as described in (??). This thermodynamic ABL is a product of atmospheric turbulent kinetic energy and lapse rate. Similar gradients can be seen in chemical and aerosol profiles as well (Fig. 2 B-C). For the current study,
- 10 meteorological sondes radiosondes from International Met Systems (iMet; Grand Rapids, Michigan) and ozone sondes from Droplet Measurement Technologies (DMT, now En-Sci; Boulder, Colorado) 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 (?).

 $\frac{\Delta\theta}{\Delta Z} \ge 0.005 K * m^{-1}$



Figure 2. Potential temperature, ozone, and backscatter profiles recorded on 8-June 2015. The horizontal lines indicate the ABL, CBL, and MLH can be seen by the horizontal lines at 13:00 local time.

$\Theta_{top} - \Theta_{base} \ge 2K$

Numerous Results of numerous analyses have been presented published to illustrate differences between the various chemical and meteorological sensors, and how differing meteorological sensors influence secondary chemical measurements such as ozone (???????). While these influences may impact the can impact the derived CBL, the ABL and MLH remain unperturbed.

(2)

5 Therefore, the remainder of the current work will focus focuses on the MLH and ABL, and CBL variability is considered to be with CBL variability regarded as outside the current scope.

3 Algorithms

3.1 BLView

The Vaisala standard MLH retrieval is based on a proprietary wavelet/gradient technique built within the logging/analysis

10 software BLView .BLView makes use of variable time and altitude averaging when calculating the MLH. Typical averaging time ranges from 14 minutes min at night to 52 minutes min during clear-sky, daytime conditions, and is automatically adjusted within the software according to signal-to-noise ratio. Altitude averaging is variable with altitude , varies with altitude and ranges from 80 m near the surface to 360 m above 1.5 km. Further, BLView selectively removes false-positive MLH iden-

tifications by requiring a minimum number of similar MLH values (± 140 m) be within the last several minutes, and has the ability to discriminate between MLH inversions and changes in backscatter intensity induced by eloudclouds, precipitation, and fog.

An advantage Advantage of the BLView software is are the standardization of the retrieval parameters, retrieval parameters

5 and a user interface that provides flexibility in setting user-specified sensitivities. This comes These come 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.

3.2 STRAT

The STRucture of the ATmosphere (STRAT v1.04) algorithm was developed under a GNU General Public License to analyze

- 10 aerosol vertical profiles , as measured via LIDAR, as measured by LIDAR and estimate cloud heights and aerosol MLH from a variety of LIDAR instruments. It is currently in use by the European Aerosol Research LIdar Lidar NETwork (EARLINET) (????). STRAT utilizes uses a covariance wavelet technique (CWT), of which the full details can be found in ? and ?. STRAT can be run exclusively in MATLAB, or a combination of MATLAB and Python. Due to its wide use throughout the European network it is considered here as a viable open-source alternative to BLView.
- 15 While BLView provides limited user control of the retrieval process, which is beneficial in regards with regard to standardizing the retrieval process across a network, STRAT provides a significantly greater amount of user control. Such control is beneficial desirable since retrieval parameters in a heavily polluted region will likely be different than retrievals done from those 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 can run batch jobs, which is beneficial useful 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 averaging time and vertical resolution 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.

25 3.3 UMBC Algorithm

The UMBC University of Maryland Baltimore County (UMBC) algorithm was developed independently for estimating MLH from LIDAR backscatter profiles using a CWT similar to STRAT. The STRAT software was designed specifically for singlechannel LIDARs (primarily ceilometers) and is not readily customizable to other LIDAR systems, such as the Miero-Pulse Lidar. By using the aerosols as tracers of the atmospheric dynamics, the LIDAR is a powerful tool for visualizing, in real time,

30 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-MPL. The UMBC algorithm was designed to be more flexible than STRAT in that regard and uses a <u>CWT to identify the</u> sharp gradient changes in the LIDAR backscatter profiles to determine the indicative of the MLH (Davis et al. 2000; Brooks 2003).

Detailed A detailed description of the UMBC algorithm has been published elsewhere ?. The first step in the MLH algorithm

5 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. in ?.

4 Locations

10 4.1 CAPABLE Site

The CAPABLE site (37.103° N, 76.387° W, 5 m ASL) was established at LaRC, in the greater Hampton Roads region (collection of cities on a group of cities in 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 di-

- 15 rectly 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 <u>in-situ in situ</u> and remote-sensing instruments, including a CL51 ceilometer and sounding station, that allows. These instruments allow thorough sampling of the atmosphere to provide valuable <u>in-situ in</u> <u>situ</u> and profile information within the lower troposphere in a highly complex (<u>i.e.</u> due to bay-breeze events, ;; see ?) and moderately polluted <u>environment that will provide (NO₃, SO₂, aerosols) environment yielding</u> valuable satellite ground-truthing
- 20 and model a-priori a priori estimates.

CAPABLE (37.103° N, 76.387° W, 5 m ASL) is located on a peninsula between the James River to the southwest, the 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 \approx 15,000 vehicles per day, Commander Shepard Blvd to the south at \approx 20,000 vehicles per day; Commander Shepard and Wythe Creek share much of the same traffic, so it is not reasonable

25 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.

The Hampton Roads region can be described as moderately-polluted moderately polluted. Aerosol statistics ($PM_{2.5}$ and aerosol optical thickness (AOT) as recorded by a sun photometer within the AERosol Robotic NETwork (AERONET) as

30 described by ?) are presented in Table 1. The <u>data show</u> AOT loads at CAPABLE are <u>seen to be</u> significantly higher than <u>at</u> the corresponding Colorado sites, particularly in the <u>lower-size distributions</u>. <u>lower size distributions (i.e., lower wavelengths in Table 1)</u>.

Site	$\lambda \text{ (nm)/PM Size}$	Mean	Q_1	Q_2	Q_3
BAO-Tower	380	0.23	0.13	0.19	0.32
BAO-Tower	500	0.15	0.09	0.13	0.22
BAO-Tower	675	0.09	0.05	0.08	0.13
BAO-Tower	870	0.06	0.04	0.06	0.08
BAO-Tower	1020	0.05	0.03	0.04	0.06
Golden, CO	380	0.20	0.10	0.16	0.27
Golden, CO	500	0.13	0.06	0.10	0.18
Golden, CO	675	0.08	0.04	0.06	0.11
Golden, CO	870	0.05	0.03	0.04	0.07
Golden, CO	1020	0.04	0.02	0.03	0.05
CAPABLE	380	0.34	0.21	0.33	0.45
CAPABLE	500	0.23	0.13	0.23	0.32
CAPABLE	675	0.14	0.08	0.14	0.20
CAPABLE	870	0.09	0.05	0.09	0.13
CAPABLE	1020	0.07	0.03	0.06	0.10
CAPABLE	$PM_{2.5}$	5.80	2.59	5.00	8.44

Table 1. Aerosol optical thickness statistics at the three sites under study. Here, Q_1 , Q_2 , and Q_3 represent the 25^{th} , 50^{th} , and 75^{th} percentiles, respectively. Data have been filtered to show only contain data collected during the DISCOVER-AQ 2014 field campaign period (i.e. July – August 2014).

4.2 Erie, CODISCOVER-AQ/BAO-TowerFRAPPE Sites

From 2011 through 2014 the National Aeronautics and Space Administration (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. A primary objective of DISCOVER-AQ was to investigate the ability of satellite remote sensing to inform surface air quality. Since the ABL limits vertical exchange of primary pollutants and directly influences near-surface pollutant concentrations, the ABL height directly influences 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, as well as the diurnal variability of these distributions in conjunction with the ABL. The final DISCOVER-AQ field mission was conducted over Denver and the Front Range region of Colorado in July and August 2014.

10 and was conducted jointly with the Front Range Air Pollution and Photochemistry Experiment (FRAPPE).

4.2.1 Erie, Colorado/BAO-Tower

5

Data were collected at the Erie, CO site (Colorado site (40.045° N, 105.005° W, 1500 m ASL), which is considered to be a clean environment as compared to CAPABLE , (see Table 1, 40.045° N, 105.005° W, 1500 m ASL) from 14-July–12-August

) from 14-July to 12-August 2014 as part of the DISCOVER-AQ field mission. The Erie , or BAO-Tower, site site (rural community surrounded by agricultural activity) 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 and served as a combined DISCOVER-AO/FRAPPE ground site. The site is often referred to as BAO-Tower because of the site's primary

5 featurewas: a 300 m tower(known as-. BAO-Tower), which provided a unique profiling ability for in-situ 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 2014As part of FRAPPE, the University of Wisconsin's (UW) Space Science and Engineering Center trailer, which housed a High Spectral Resolution LIDAR (HSRL) high spectral resolution lidar and from which regular

10 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 trailerand the CL31, both ceilometers experienced the same chemical, aerosol, and meteorological conditions.

4.3 Golden, CO

Data

15 4.2.1 Golden, Colorado

CL51 data were collected at the Golden, CO site (Colorado site (39.750° N, 105.183° W, 1850 m ASL) (considered to be a clean environment as compared to CAPABLE, see Table 1, 39.750° N, 105.183° W, 1850 m ASL) from 14-July–12-August) from 14-July to12-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 (a flat-topped geographic structure). Due to the site's elevation

- 20 on the mesa, and its 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. This is demonstrated in Fig. 3 by the lack of structure in the diurnal MLH profile. While both the BAO and CAPABLE sites demonstrate the expected nocturnal low/daytime high MLH, the Golden diurnal variability is not as well defined, consistent with ABL development in mountainous terrain (????).
- The Golden site housed the U.S. EPA trailer, the LaRC ozone LIDAR, micro-pulse LIDAR (MPL) MPL and LEOSPHERE
 ALS-450 LIDAR operated by UMBC, a SOund SOnic 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 $\frac{36-60-36}{36}$ to $\frac{60}{50}$ s, while sonde-profile data had average measurement times of 1 s. To harmonize all Due to the nature of sounding data sets, sonde-based

30 <u>ABL's were not averaged to 5-min resolution. To harmonize LIDAR</u> datasets to a common time frame the data were averaged to 5-min resolution unless otherwise specified. Further, it is well known that the atmosphere changes throughout



Figure 3. Diurnal variability of the MLH at the three sites. Data were resampled to 5-min averages and filtered.

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 4-hour segments to remove biases caused by time-of-day influences. Since the primary objective of the assessment is this assessment was to understand how the CL51 MLH compares compared with other instruments/methods, all work is analytical results are 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 data-acquisition methodology is presented first,

10 followed by a comparison of MLH retrieval algorithm on algorithms applied to data collected by a single instrument, and finally then a comparison of the various instrumentation. As MLH variability follows a distinct diurnal cycle as shown in Fig. 3, all dates/times are presented in local standard time.

5.1 Data Acquisition

The objective of the current subsection is Data-acquisition methods were analyzed 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

- 5 and preconditioning the profiles (here referred to as BLView; note: , however, that this refers to the backscatter-profile that is logged by BLView , and not the BLView-calculated MLH), while the Python script logged the raw incoming data stream up to the full profile (FP) height (i.e., 15.4 km, this dataset is referred to as full-profile, or FP). The question being: was, does application of the LIDAR calibration factor influence the MLH estimate? This question will be is addressed in section ??, but first, a viable filtering criteria that removes to remove spurious MLH fluctuations from the dataset must be defined data set
- 10 were developed prior to analysis. Defining this criteria will be the topic of , as discussed in section ??.

5.1.1 Filtering ProcedureCriteria

Regardless of the method of data acquisition data-acquisition method (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

- 15 mannerwere needed for quality control. Since ABL and MLH variations occur 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, frontal systems, window contamination). Therefore, it remains to identify these cutoff criteria for implementing data filtering . Since the effect of the implementation were identified. This portion of the analysis was conducted first because application of these cutoff criteria will influence the data-acquisition 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 with standard deviations significantly increasing over the longer time periods and during rapid transition events. Therefore, the current analysis must be was performed on short-time-series data (e.g. 5 10 minutesi.e., MLH resampled to 5-min resolution)
- 285 to eliminate a bias caused by natural low-frequency changes. Figure ?? shows a series of histogram percentile plots for data collected at LaRC (the largest dataset within the current analysisN > 30E5), where the standard deviation of MLH was calculated over five-minute intervals 5-min intervals and subsequently averaged to provide mean standard deviation every four hours. This figure is elucidative as it shows the distribution elucidates the variability 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,
- 290 excluding and algorithms. Except for the afternoon period (12:00 19:00local time where, local time) when 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 with a 5-min standard deviation greater than 0.20 km were removed from subsequent analysis (labeled "filtered")and data that have. Data with a relative standard deviation greater than - or equal to -20% were also removed. Implementation of these filter criteria removed up to 10% of the data at each site.



Figure 4. Histogram plots showing the distribution of Percentiles for MLH standard deviations deviation throughout the day from the CAPABLE site. Vertical lines represent percentiles. Panels Data in panel A were collected and processed in BLView, data in panel B were collected with the Python script and processed in STRAT; data in panel C was processed were collected in BLView and processed in STRAT. It is observed that variability was maximum during the afternoon regardless of collection method or processing algorithm.

295 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. ??) 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. ??). 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.

300 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 <u>abovein section 3.2</u>, followed by a <u>5-minute-5-min</u> block average. Figure ?? presents the data-

305 <u>The data were replotted</u> 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



Figure 5. Backscatter curtain plot with STRAT-derived MLH values (5-min mean) from the BLView (BLV) and Python (FP) collection methods.

a Gaussian-based kernel-density estimation (??) as supplied in Python's scipy.stats.kde module(, represented mathematically in Eqs. ??-?? where X is the 2 x n vector of the x and 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. ?? 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 w here. It is observed that the majority of the MLH estimates fall along the 1:1 line (center column in Fig. ????), though there is significant scatter along both axes. The source of the 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.-

$$\Delta \mathbf{X} = \mathbf{X} - \mathbf{X}[:,i] \tag{3}$$

$$\mathbf{E} = \sum_{j=1} \Delta \mathbf{X}_j \cdot \frac{cov(\mathbf{X})^{-1}}{f^{-2}} \bullet \Delta \mathbf{X}_j \tag{4}$$

$$\mathbf{w}[i] = \frac{\sum\limits_{k=1}^{n} e^{-\mathbf{E}_k}}{\sqrt{\det\left[2\pi \cdot \cot(\mathbf{X}) \cdot f^2\right]}} \bigg|_i^n \qquad \{i \in \mathbb{N} : i \le n\}$$
(5)



Figure 6. 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 Data density is presented to better understand the distribution within the scatter plots. Data were averaged to five-minute 5-min resolution, without application of filtering criteria (left and center columns), and averaged to one-hour resolution with application of filtering criteria (right column).

Figure ??

- 320 Figure ?? was divided into four-hour 4-hour blocks to identify any time-of-day dependence. It is observed that regardless of the time of day. The figure shows that most of the data continued to fall along the 1:1 line regardless of time of day, as indicated in the density plots, for CAPABLE and BAO-Tower, while the density plots. The Golden site displays some disruption in the 16:00 19:59 panel. The, but the source of this discrepancy is currently unknown. However, it It has become clear, however, that the meteorology at the Golden site is different from that observed at CAPABLE and BAOBAO-Tower. It is suggested that this difference is primarily driven by orographic perturbations and as well as the Golden site being located 's location atop a
- mesa, both of which may can inhibit formation of stable ABL and ML (???).

Figure 7. Same data set as in Figure ??, but with the data resampled to 1-h means after application of filtering criteria. Due to the sparseness of the data, as compared to Fig ??, there is not need to present data density on the z-axis. Here, the z-axis represents relative standard deviation to show the relative variability within each 1-h block after filtering.

Ceilometer-derived MLH values have application as model a-priori inputs that have been For regulatory and modeling applications, 1-hour averages are standard, requiring the data be averaged down to one-hour 1-hour resolution. The impact of the filtering criteria and re-sampling to one-hour 1-hour resolution throughout the day is can be seen in Fig. ?? (panels C, F, I), and ??. Note, the density of data around the 1:1 line is readily apparent in Fig. ??, therefore the z-axis has been converted to relative standard deviation to show the relative variability within each 1-h time block, after application of filtering criteria. The intention is to provide some understanding of how much the MLH will change within the model and regulatory applications' time frame. Table ?? 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 1-hour blocks, the scatter is dramatically reduced (Fig. ?? panels C, F, I)... This is likely due to the scatter being evenly distributed about around 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. ????

	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 -1 <u>-h</u>	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-1-h	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 -1-h	0.77	$y=0.792{\cdot}x+0.35$	-0.14(13.0)

Table 2. Summary of aggregate statistics for the Python-collected (FP)/STRAT-processed and the BLView-collected (BLV)/STRATprocessed MLH estimates (y and x, respectively). Filtering eriteria Data were applied resampled to 5-min resolution followed by application of filtering criteria to both datasets (lines labeled 1-h present statistics after data were filtered and subsequently resampled by a 1-h block average). Values in parentheses indicate percent of the difference value with respect to the BLView-derived MLH.

From It can be concluded 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. In addition, no significant difference is observed between the BLView- and Python-collected data sets 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 section ?? further support this conclusion.

5.1.3 MLH Algorithm Dependence

340

In the previous sectionit was demonstrated that, the data collection method (i.e.raw serial logger, Python vs. BLView) has was shown to have little impact on the derived MLH values when the two datasets are processed on were processed using a common algorithm (here, STRAT). It remains to be seen The question remains of how the two datasets data sets compare when processed in different algorithms. Whereas collection methods were compared in the previous section, the algorithms will be compared here. Data collected by To answer this question, data collected with the Python script were processed using the STRAT algorithm , and are and were compared with data collected , and processed , by and processed with BLView.

- Figure ?? presents scatter plots similar to Fig. ?? those in Fig. ??, but with data collected and processed by differing means.
 350 It is observed that the majority of using the two different methods. Most data continued to fall along the 1:1 line, as attested by shown in the density plots, and that much of the scatter is caused by short-term variability. However, in contrast to Fig. ?????, the scatter is neither as evenly distributed nor as tightly grouped about around the 1:1 line. The STRAT-derived MLHs were generally lower than those calculated in BLView (according to given by the slopes) at all sites, while the aggregate mean difference shows the opposite for the Colorado sites (Table ??), which is likely being driven by outliers.
- 355 It is observed that the The agreement between the two datasets is less than when a common algorithm was employed (Table ??). Despite the increased scatter, there remains a significant subset of data that lies remains along the 1:1 line. As a test for how well the data fit the 1:1 line, the R and LOBF values were re-calculated using Eq. ?? with weights applied according to Eq. (??) (i.e. weighting according to data density)data density. Therefore, points that have more data points surrounding

$$m = \frac{N\sum_{i=1}^{N} w_i x_i y_i - \left(\sum_{i=1}^{N} w_i x_i\right) \left(\sum_{i=1}^{N} w_i y_i\right)}{N\sum_{i=1}^{N} w_i x_i^2 - \left(\sum_{i=1}^{N} w_i x_i\right)^2}$$
(7)

365
$$b = \frac{\sum_{i=1}^{N} w_i y_i - m \sum_{i=1}^{N} w_i x_i}{N}$$
 (8)

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 distributed around the 1:1 line to the exclusion of other regions. As an example, despite weighting, the improvement in the Golden regressionsare, despite weighting, is notably less than the other two sites. This is likely due to more spread in the data, thus mitigating which mitigates the influence of the points along the 1:1 line on-in the regression analyses. Therefore, we can conclude that the preponderance of the data collected at the CAPABLE and BAO-Tower sites fall nearer the 1:1 line when processed through using the different algorithms as compared to the data collected at the Golden site. Further, despite the majority of most data falling nearer the 1:1 line for these two sites, there remain influences influences remain that neither the STRAT configuration nor the current filter methodology can account for, which is likely driving the poor correlation

375 as compared to Table ??. This is likely the possibly a product of how the differing algorithms handle atmospheric interferential events (e.g. precipitation, fog, etc.). Application of a filtering methodology to account <u>for</u> 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 with the BLView-collected/processed product. It was concluded in Sec. ??that As concluded in section ??, the data

- 380 collection method had little influence on the MLH estimation when both datasets were processed on using 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 ??. 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 ??, wherein we see similarity with Table ??. This further supports These findings further support the conclusion that data collection methods (including application)
- 385 of calibration factors) play relatively less much less of a role in identifying a qualitative gradient within the profile as compared to than 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 by a common algorithm.

5.2 Sonde Intercomparison

370

Meteorological soundings have been a staple for profiling the atmosphere and deriving ABL heights for decades. These ABL 390 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 hereinhere. 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.

Intercomparison of sonde-based ABL and ceilometer-based MLH can be complicated due to the fundamentally different

- 395 nature of the two observations. Sondes provide a direct measurement of the atmosphere, while ceilometers provide an indirect (i.e. remotely-sensed) measurement. Therefore, care must be taken when comparing the two sets of observations. Further, the aerosol profile can be impacted by aerosol layers transported aloft, thereby offsetting the MLH estimate. 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 30-min for comparison. Additionally, each measurement
- 400 can be impacted by different atmospheric phenomena which atmospheric phenomena that can affect the measurements in different ways , which in turn can and can in turn affect the comparison of the measurements. A radiosonde Met-sondes can be impacted by local updraftsor down-drafts, /downdrafts and result in ABL estimates that are higher or lower than the true time- or space-averaged MLH. The response time of the sensors is less than one second, thereby minimizing offset in vertical structure. The CL51 MLH is sensitive to the calculated based on identification of a sufficiently steep, vertically-averaged, here a like the sensor is less than one second.
- 405 backscatter gradient, so if there are additional aerosol layers just above the MLH, the contrast between the aerosol layers may might not be strong enough for the CL51 to identify each layer or the correct altitude of the MLH.

Correlation plots for the CL51 MLH <u>calculated via BLView</u> compared to sonde ABL are shown in Fig **??** <u>panels A-C</u>. For all coincidence times, the CAPABLE site showed the best correlations between the CL51 and radiosondessondes. The correlation for the CL51 versus all the radiosondes sondes (N = 25) at the CAPABLE site was R=0.820.79, with a similar correlation R

410 = $0.83 \cdot 0.82$ (N = 22) when the filtering criteria were implemented. For daytime data, the CAPABLE site contained 2 early morning radiosondes two early morning sondes (before 10:00 local time), with all other radiosondes sondes launched between 10:00 and 16:00 local time. By the late morning, $\approx 10:00$ 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 **??** A, with Fig. **??** A where many of the data points falling fall close to the 1:1 line.

415 Radiosonde

<u>Met-sonde</u> 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 <u>2 morning radiosonde two morning sonde</u> launches, which is occurred during a transition period when the boundary layer is was experiencing rapid growth. Upon applying the filtering criteria, the <u>2-two</u> early morning data points

420 were removed, resulting in a much improved correlation (filtered R = 0.74; N = 10) for the Golden site. This appears to These results indicate the CL51 may might have difficultly capturing an accurate MLH during rapidly changing conditions, such as during early morning and late evening transition periods in a clean atmosphere.

It is somewhat surprising that the 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 farther to the east of the Rocky Mountains, at the start of the High Plains, which is are less influenced by very local geographic perturbations, and that a similar relationship is not observed in the CL51 intercomparisons (e.g. Tables ??, ??, and ??). As a check on the radiosonde of the met-sonde potential temperate profiles, the

Figure 9. Correlation plots for CL51 MLH and sonde-derived ABL estimates. Statistics data in black text are for Black error bars represent the entire spread in unfiltered dataset, while the red text represents error bars represent the filtered dataset. MLH values (30-min average, centered on sonde-launch time) were calculated in BLView (panels A-C) and STRAT (panels D-F) and resampled to 30-min resolution. Error bars indicate standard deviation of the CL51-derived MLH within the 30-min period.

Site	Algorithm	Not Filtered	Filtered
CAPABLE	BLView	0.79 (26)	0.82 (23)
CAPABLE	STRAT	0.14 (14)	0.82 (11)
BAO-Tower	BLView	0.63 (16)	0.58 (14)
BAO-Tower	STRAT	0.34 (16)	0.79 (14)
Golden	BLView	-0.28 (12)	0.74 (10)
Golden	STRAT	0.70(13)	0.70 (13)
Composite	BLView	0.55 (54)	0.80 (47)
Composite	STRAT	0.44 (43)	0.72 (38)

Table 5. Summary of statistics for the CL51/sonde MLH/ABL intercomparison, corresponding to Fig. ??. Numbers in parentheses indicate sample size. Composite statistics were generated by looking at all sites as a single dataset. In this table only, the filtering method for the STRAT-based MLH is based on visual identification of false MLH values due to clouds/precipitation events and unusually clean atmospheres as described in text.

potential temperature data from the NASA P-3B aircraft spirals conducted over the Golden and Erie sites is shown in Figs. **??** and **??**. These spirals are coincident with the launch of the radiosondes met-sondes from the sites. Also The coincident CL51 backscatter profiles are also plotted in Figs. **??** and **??** is the coincident CL51 backscatter profile. The agreement between

430 the radiosonde and P-3B aircraft profiles is good, indicating that the potential temperature within the aircraft spiral radius is consistent with that of the radiosonde. 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.

Figure 10. Example plot where STRAT identifies cloud deck as the MLH (12:00).

The STRAT-derived intercomparison with sonde ABL is presented in Fig ?? panels D-F, where it is observed that the agreement is significantly less than when BLView was used to calculate MLH. This disparity is caused by spurious MLH values

- 435 from STRAT that are observed under two conditions: 1. during heavy cloud cover/precipitation events STRAT sometimes falsely identified the cloud deck as the MLH and completely ignored the MLH gradient 1-2 km below the cloud; 2. STRAT failed to identify a valid MLH when the atmosphere was exceptionally clean, and instead identified a stronger, spurious, gradient 2-4 km up. An example of the first type is presented in Fig. ?? where STRAT switches from properly identifying the MLH at \approx 0.5 km to identifying the cloud deck (\approx 2.4 km) as the MLH starting around 12:00 local time and an example of
- 440 the second type is shown in Fig. ??. A corresponding shift was not observed in the BLView-derived MLH for the same day, indicating BLView has been trained to recognize these spurious events and ignore them.

After removing these "false" MLH values the coefficient of correlation between STRAT-derived MLH and sonde ABL (pre-filtering) improved for all sites to 0.82, 0.79, and 0.70 for LaRC, BAO-Tower, and Golden, respectively. The results of these correlations is encouraging and is indicative of the importance of properly training the STRAT algorithm to identify and

445 exclude these false-positive events. The down side is that despite having better correlation (after removing spurious events), the variance of STRAT MLH values larger than that of BLView, indicating that defining an MLH filter criteria is dependent on

Figure 11. Example plot where STRAT fails to identify a reasonable MLH due to unusually clean conditions.

the algorithm in use. However, the positive aspect of this is that the STRAT algorithm, being open source with the source code available, can, in theory, be modified by end users to identify and account for these spurious events.

450

Overall, all 3 sites show a three sites show good correlation between the CL51 and radiosonde met-sonde data, with MLH fand ABL estimates from the radiosondes 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. ??, with the exception being the unfiltered results for Golden.

5.3 **MPL Intercomparison**

455

The MPL instrument was collocated with the CL51 stationed at the NREL site in Golden, COColorado. Being a LIDAR instrument, it profiles the atmosphere similarly to the CL51 with the major differences being difference being their hardware. The two instruments emit different wavelengths (CL51:910 nm, MPL:532 nm), 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 thicknessAOT). However, if the ML is well mixed, then the general particle distribution and gradient will be the same, making the two inter-comparable.

Figure 14. Correlation and data-density plots for the CL51 (BLView processed) and MPL (UMBC algorithm) MLH estimates from Golden, COColorado.

- 460 It is seen in Fig. ?? Figures ?? and ?? shows that the agreement between the two instruments and algorithms (BLViewprocessing CL51,, STRAT for CL51 profiles and UMBC algorithm processing MPL profiles) is poor, even though a significant subset of data fall along the 1:1 line, as indicated by data density (*z*-axis). The low correlation is partly driven by the invariability in one instrument as compared to the other at lower MLH values (i.e. \leq 500 m). Removal of data MLH below 500 m improved the coefficient of correlation to 0.368, 0.512, and 0.390 respectively. coefficients of correlation for the 5-min averaged data to
- 465 0.467, 0.489, and 0.469 for BLView-derived MLH values (Fig. ??, panels A, B, C respectively) and 0.433, 0.471, and 0.368 for STRAT-derived MLH (Fig. ??, panels A, B, C respectively) values. Similar to the algorithm comparison, much of the variability between the two instruments and algorithms occurs during events which that inhibit a reliable MLH estimate being made estimation (e.g., fog, precipitation) of MLH (as seen in Fig. ??).

The most commonly used statistical techniques used for comparing two datasets depended on two key assumptions:

- 470 data being-were normally distributed and homoscedastic. The CL51 and MPL MLH 5-minute-5-min 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 determined using the two-sample Kolmogorov-Smirnov test. It was determined that the 5-minute-5-min averaged MPL and CL51 datasets are were statistically different (p ≪ 0.01), regardless of filtering and averaging. However, when considering 1-hour averaged
- 475 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 bias induced by these low-altitude MLH values it is quite clear that they are significantly influencing significantly influence the intercomparison. Given that this is

the first intercomparison of these two instruments and algorithms, it is not surprising that a significant difference was identified in this regimewas identified.

480 6 Conclusions

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

- Two collection methods and processing algorithms were tested for the CL51 MLH calculation. We demonstrated that the data-collection method plays played an insignificant role in MLH estimation when the datasets are processed on were processed using a common algorithm. Furthermore, the choice of processing algorithm plays played a significant role in MLH estimation. We Therefore, we recommend that, for ceilometer and LIDAR networks, a common MLH processing algorithm be employed. Agreement between the different algorithm products may might be dictated, to a large degree, by local atmospheric fluctuations and interferential events (e.g. fog), and should be the topic of which should be a topic for future investigation.
- 490 A total of 53 potential temperature profiles from radiosondes were used to evaluate the CL51. While the 53 radiosondes were spread across 3 sites, this represents a robust data set of soundings. Overall, the radio-sonde-derived-met-sonde-derived ABL was higher than the CL51 MLH(e. g. Figure ??). Comparison of MLH from the CL51 versus radio sondes show met-sondes shows the CL51 performed best at the CAPABLE research site (non-filtered R = 0.810.79, filtered R = 0.82), a moderately-polluted moderately polluted coastal site primarily influenced by a combination of sulfate and marine aerosols.
- Both the Golden and BAO-Tower sites are located in cleaner environments than CAPABLE and show good correlation between the CL51 and radiosondes met-sondes (Golden filtered R = 0.74, BOA non-filtered R = 0.63, filtered R = 0.38) with 2 early morning radiosondes 0.58) with two early morning sondes at the Golden site strongly influencing the non-filtered correlation (R = -0.28). These 2 radiosondes two sondes measured a very shallow boundary layer, < 500 m, while the CL51-identified CL51 identified the MLH above 2 km, which was likely due to residual aerosol layers aloft. The lower correlations at the Colorado
- 500 sites (Golden and BAOBAO-Tower) 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 (???). With the Golden site being (????). The Golden site likely experienced greater up- and down-slope flows than the BOA-Tower site because of its location 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
- 505 <u>can impact or challenge likely impacted or challenged</u> the well-mixed assumption required to compare thermodynamic ABL measured via potential temperature and MLH measured via aerosol backscatter , and in the current study. These influences should be made a consideration in <u>future</u> intercomparisons.

The results of the CL51 versus the UMBC $\frac{MPL}{MPL}$ algorithm that was run on MPL data 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

510 plots. When data filtering criteria are applied data-filtering criteria were applied, the two data sets were statistically indistin-

guishable (p > 0.8). Additional analysis is planned to further explore the cause of the low correlation. Howeveras can be seen in Figure 15, the MLH from the CL51 and MPL agree well when there is a well-defined MLH.

Acknowledgements. Funding for this work was provided by the NASA Applied Sciences Program am, EPA collaborations under an EPA-LaRC memorandum of agreement, EPA Air, Climate, and Energy Research Program and GEO-CAPE mission studies, and Langley Innovative

515 Partnership Program .mission studies. T. N. Knepp was funded through the STARS-III contract. Although this paper has been reviewed by the EPA and approved for publication, it does not necessarily reflect EPA policies or views. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Banta, R.: Daytime Boundary-Layer Evolution over Mountainous Terrain .1. Observations of the Dry Circulations, Monthly Weather Review,

- **520** 112, 340–356, doi:10.1175/1520-0493(1984)112<0340:DBLEOM>2.0.CO;2, 1984.
 - 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.
- 525 Brooks, I.: Finding boundary layer top: Application of a wavelet covariance transform to lidar backscatter profiles, Journal of Atmospheric and Oceanic Technology, 20, 1092–1105, doi:10.1175/1520-0426(2003)020<1092:FBLTAO>2.0.CO;2, 2003.
 - Coen, M. C., Praz, C., Haefele, A., Ruffieux, D., Kaufmann, P., and Calpini, B.: Determination and climatology of the planetary boundary layer height above the Swiss plateau by in situ and remote sensing measurements as well as by the COSMO-2 model, Atmospheric Chemistry and Physics, 14, 13 205–13 221, doi:10.5194/acp-14-13205-2014, 2014.
- 530 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,
- 535 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
- 540 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
- 545 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
- 550 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 MF-

- 555 DOAS 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 pre-
- 560 cipitation: Finland's new ground-based remote-sensing network, Atmospheric Measurement Techniques, 7, 1351-1375, doi:10.5194/amt-7-1351-2014, 2014.
 - Holben, B., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET - A federated instrument network and data archive for aerosol characterization, Remote Sensing of Environment, 66, 1-16, doi:10.1016/S0034-4257(98)00031-5, 1998.
- 565 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,
- 570 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.
- Kotthaus, S., O'Connor, E., Münkel, C., Charlton-Perez, C., Haeffelin, M., Gabey, A. M., and Grimmond, C. S. B.: Recommendations for processing atmospheric attenuated backscatter profiles from Vaisala CL31 ceilometers, Atmospheric Measurement Techniques, 9, 3769-575 3791, doi:10.5194/amt-9-3769-2016, http://www.atmos-meas-tech.net/9/3769/2016/, 2016.
- 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.,
- 580 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, 11587-11609, 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, Inter-
- national Specialty Conference on Regional Photochemical Measurement and Modeling Studies, SAN DIEGO, CA, NOV 08-12, 1993. 585 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 590 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.
- 595 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 Ouality Radiosonde Systems Yangijang, China, 12 July 3 August 2010, Report WMO/TD-No. 1580, World Meteorol. Organ., Geneva, Switzerland, 2011. 600
- NRC: Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks, in: Observing Weather and Climate from the Ground Up: A Nationwide Networks of Networks, pp. 1-234, Natl Academies Press, 2101 Constitution Ave, Washington, DC 20418 USA, 2009.
 - NSTC: Office National Plan for Civil Earth Observations, Tech. rep., Executive of the President,
- 605 https://www.whitehouse.gov/sites/default/files/microsites/ostp/NSTC/national plan for civil earth observations - july 2014.pdf, 2014.
 - 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.: 610 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.,
- 615 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,
- 620 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.
- 625
 - 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, 2012.

- 630 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, 2013. Schween, J. H., Hirsikko, A., Löhnert, 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, John Wiley & Sons, 1992. 635 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.
- 640 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.

- 645 Stull, R. B.: An Introduction to Boundary Layer Meteorology, Atmospheric Sciences Library, Springer, 1988.
- Sullivan, J. T., McGee, T. J., Langford, A. O., Alvarez, II, R. J., Senff, C. J., Reddy, P. J., Thompson, A. M., Twigg, L. W., Sumnicht, G. K., Lee, P., Weinheimer, A., Knote, C., Long, R. W., and Hoff, R. M.: Quantifying the contribution of thermally driven recirculation to a high-ozone event along the Colorado Front Range using lidar, Journal of Geophysical Research-Atmospheres, 121, 10377-10390, doi:10.1002/2016JD025229, 2016.
- 650 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.
- 655
 - 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.

We appreciate the thorough review from referee #1. The manuscript has been updated to implement these recommendations as described below.

- 1. "... being as clear as possible such as when describing averaging time and vertical resolutions, and exactly which CL51 data processing algorithm is being depicted in each figure..."
- (a) Clarification was made throughout the text to aid the reader in knowing what averaging, resolutions, and algorithms are being used.
 - 2. Page 2, Line 21-23: Do you have a reference for this statement from the US NRC?
 - (a) Reference has been included.
 - 3. Page 2, Line 24: This reference seems incomplete
- 10 (a) Reference corrected

- 4. Page 2, Lines 29-34: These lines are written in passive voice. Please rewrite
 - (a) Paragraph rewritten
- 5. Page 2, Line 34: The way this line is written makes it seem like you are comparing three CL51s from Colorado against sondes from CAPABLE.
- 15 (a) Sentence corrected.
 - 6. Page 3, Line 26: It's never stated why BLView truncates data at 4.5 km. Are there concerns about measurement uncertainties or S/N ratios at higher altitudes? I realize this probably doesn't have an effect on the MLH calculations.
 - (a) That is unknown to us and is one of the challenges in using proprietary software as it remains a black box. Any comment we provide on this would be speculative, so we will not comment.
- 20 7. Page 4, Line 12: Are the 1 min and 30 m resolutions from the MPL what you've chosen to record specifically for this study? Please state
 - (a) Stated
 - 8. Page 5, Line 25: Delete "to be."
 - (a) Deleted
- 25 9. Page 5, Line 26: clouds not cloud.
 - (a) Changed
 - 10. Page 6, Figure 2: The .5s are missing on the y-axis labels
 - (a) y-tick labels corrected
 - 11. Page 7, Line 3: What parameters specifically? Averaging time? Vertical resolution?
- 30 (a) Correct. Now specified in text.
 - 12. Page 7, Line 10: "A detailed description of the UMBC algorithm has been published in Compton et al. (2013)."
 - (a) Recommendation implemented.

- 13. Page 7, Lines 10-14: These lines contain jargon that receives no other mention. You can probably tack the single sentence on line 10 to the end of the previous paragraph and delete the rest.
 - (a) Change implemented.
- 14. Page 8, Line 2: farther not further
 - (a) Change implemented.
- 5 15. Page 9, Line 11-13: Are you saying that there were two CL51s at the BAO-Tower? I'm confused about the instrument set up here.
 - (a) No, they housed the CL51 used in the current study. Text changed to "the CL51" instead of "a CL51" to indicate this.
 - 16. Page 9, Line 15: "CL51 data were collected ..."
- 10 (a) Change implemented.
 - 17. A figure showing average diurnal MLH from each of the three sites would be very helpful here and would give context for the statement that Golden often does not observe a well-developed boundary layer.
 - (a) New figure (Fig. 3) inserted and text added within body.
 - 18. Page 9, Line 25: Only the CL51 and MPL data were averaged to 5 min resolution, correct? There are a lot of time and vertical resolution averaging numbers being thrown around and they should all be clear.
 - (a) That is correct. Due to the nature of sonde data we cannot resample to a longer time period. Clarification is made within the text.
 - 19. Page 10, Line 23: "... the standard deviation of MLH was calculated ..."
 - (a) Recommended change implemented
- 20 20. Page 11, Figure 3: Somewhere in the text it would be useful to state that all times presented are in local standard time
 - (a) Statement added in analysis section
 - 21. Page 11, Lines 4-9, Figure 5: I found Figure 5 to be confusing and in need of some clarification. How should this figure be interpreted? That variability within the 5 min measurement period is generally very low when the methods agree, and peaks when the difference between the two methods is between .5 and 1km? Shouldn't relative standard deviation (σ /xbar) be unitless? It has units of km on Figure 5. Please clarify.
 - (a) You are correct, σ / xbar should be unitless. While this figure is interesting, it is only mentioned in the text once and we do not feel that it adds significantly to the manuscript. Rather, inclusion only distracts the reader and may cause unnecessary confusion. The intention of including this figure was to further support the selection of filter criteria, though we feel these criteria are adequately supported without this figure. The figure was removed from the manuscript.
 - 22. Page 15, Figure 7: The color bar and what's plotted on the z-axis are not the same as Figure 6. Did you mean to plot data density rather than relative standard deviation? The current Figure 7 seems to present similar data as Figure 5 in a different way.
 - (a) We appreciate the reviewer's sharp eye to catch this. Figure 7 was properly labeled, but the caption needed updated and supporting text within the manuscript's body needed clarified. The caption was updated and descriptive text was added within the paragraph beginning with "For regulatory and modeling applications...".

30

25

15

- 23. age 19, Line 8: According to Figure 9, the correlations are actually 0.81, and 0.82, not 0.82 and 0.83.
 - (a) Statistics corrected.
- 24. Page 19, Figure 9: Do these statistics significantly change based on processing method? What do the error bars represent? In general, many of the figures would benefit from more detailed captions.
- 5 Page 19, Figure 9: Can you add additional plots to Figure 9 showing the STRAT and sonde comparison?
 - (a) Yes and no. Text was added to explain this, as were two additional figures. The STRAT algorithm gets tricked in places (as discussed in the added text) and will need further refining before it is capable of operating to lesser degrees of human intervention. However, this may be a strength of the open source software paradigm in that the end user can adjust the algorithm to train it for specific purposes, if desired.
- 10 25. Page 19, Figure 9: Please adjust the axes to less than 7 km so spread in the data can be better visualized.
 - (a) Figure changed.
 - 26. Page 19, Figure 9: I'm curious what the correlation of MLH with all sondes is. Better or worse than the individual sites?
 - (a) That is an interesting thought. The composite correlations are not much different than the weighted average of the individual statistics. A "composite" dataset has been added to Table 5.
- 15 27. Page 19, Line 20: "It is somewhat surprising that the filtered..."
 - (a) Change implemented

20

- 28. Page 19, Line 20: It's difficult to definitively say that correlations at one site are "better" than another given the small sample size. What are the 90 or 95% confidence interval limits on these correlations?
 - (a) You are correct that marking one set as "better" is challenging due to the small sample sizes. However, calculation of a Pearson's coefficient of correlation confidence interval is highly unreliable due to the size of the data sets. We do not feel this would be representative of the true population statistics, so we will forbear including this statistic here. This may be a beneficial statistic to include in future work that involves larger data sets.
- 29. Page 20, Line 20: Yes, there is similar behavior at CAPABLE in the comparisons on Figure 8. This is worth future exploration for the BLView output. Did you look at STRAT processing vs. the MPL? Does this invariance feature disappear? Can you add additional plots to Figure 12 showing the MPL vs. STRAT?
 - (a) Similar behavior is seen with the STRAT and BLView algorithms. An additional figure has been added to show CL51 comparison with MPL via the two algorithms.
- 30. Page 20, Line 21: "Removal of MLH below 500 m..."
 - (a) Suggested correction implemented.
- 30 31. Page 21, Figure 10: Why do the CL51 profiles only go up to 3 km here? Same with Figure 11.
 - (a) The focus of the manuscript is on the mixed layer or boundary layer, which is well below 3 km throughout the study. As nothing of relevance is within the 3 km+ profile the profile was truncated to prevent the figure from becoming overly crowded and allow inclusion of text within the upper-left corners of each figure.
 - 32. Page 23, Figure 13: Please adjust the y-axis on plot C so we can better observe the variability in MLH differences
- 35 (a) Axis changed to show full-scale variability.
 - 33. Page 23, Figure 13: Please adjust the y-axis on plot C so we can better observe the variability in MLH differences.

- (a) Change implemented.
- 34. Page 24, Line 20: sites' not sites

(a)

- 35. Page 24, Line 22: A good up-to-date reference from DISCOVER-AQ Colorado on these types of circulations and how they affect pollution distribution is Sullivan et al. (2016, JGR...
 - (a) Reference included

We appreciate the reviewer's comments, suggestions, and taking the time to review the manuscript. We address the comments below.

- Aerosols are used as a tracer for the vertical structure of the atmospheric boundary layer when evaluating MLH from aerosol backscatter intensities. It should be kept in mind that atmospheric particles need some time to adapt to a changing vertical structure of the atmospheric boundary layer (see, e.g., the lower right frame in Fig. 1 in Emeis and Schäfer 2006). Therefore, it might be advisable to compare radiosonde results to ceilometer results obtained in the hour after (or even in the two hours after) the radiosonde ascent.
 - (a) We agree that time is required for particle distribution to adapt to changing atmospheric thermodynamics. However, these changes will be most noticeable during transition times (e.g. dawn and dusk). The number of data points from our dataset during these transition times is too sparse to generate a statistically meaningful analysis. The bulk of our data were collected outside transition events, when the MLH/ABL is comparatively stable. Therefore we consider the analysis, as presented, to be correct and would implement the reviewer's suggestion for data collected during transition events.
- 2. Horizontal advection of atmospheric particles can deteriorate the relation between the vertical structure of the boundary layer and the vertical profile of aerosol backscatter intensity.
 - (a) Now addressed.
 - 3. Radiosonde data usually have some sort of a hysteresis. The sensors need some time to adapt to the environmental conditions during the ascent. This could lead to a small bias towards higher MLH.
 - (a) Now addressed
- 20 4. A minor point is that the Spanish word "mesa" should be explained to readers not acquainted to the topography of the surroundings of Boulder, Colorado.
 - (a) Mesa is the proper term for a geographic structure. A very brief description was added to the text.

15

5