JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1002/,

Aerosol Lidar Observations of Atmospheric Mixing in Los Angeles: Climatology and Implications for Greenhouse Gas Observations

John Ware,[1,2] Eric A. Kort,[2] Phil DeCola,[3] and Riley Duren[4]

Corresponding author: J. F. Ware, Department of Climate and Space Sciences and Engineering, University of Michigan, Space Research Building, 2455 Hayward St., Ann Arbor, MI 48109, USA. (johnware@umich.edu)

[1]Department of Physics, University of Michigan, Ann Arbor, Michigan, USA.

[2]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan, USA.

[3]Sigma Space Corporation, Lanham, Maryland, USA.

[4]NASA Jet Propulsion Laboratory, Pasadena, California, USA.

Key Points.

- Aerosol lidar maps LA mixing depth in space (pilot mobile study) and time (2 years data)
- Automatic mixing depth retrieval system finds daily variability far exceeds seasonal difference
- PBL heights in models used for GHG monitoring show biases that will carry over to flux estimates

Abstract. Atmospheric observations of greenhouse gases provide essential information on sources and sinks of these key atmospheric constituents. To quantify fluxes from atmospheric observations, representation of transport – especially vertical mixing – is a necessity and often a source of error. We report on remotely sensed profiles of vertical aerosol distribution taken over a two-year period in Pasadena, California. Using an automated analysis system, we estimate daytime mixing layer depth, achieving high confidence in the afternoon maximum on 51% of days with profiles from a Sigma Space Mini Micropulse LiDAR (MiniMPL) and on 36% of days with a Vaisala CL51 ceilometer. We note that considering ceilometer data on a logarithmic scale, a standard method, introduces an offset in mixing height retrievals. The mean afternoon maximum mixing height is 770 m AGL in summer and 670 m in winter, with significant day-to-day variance (within-season $\sigma = 220\,\mathrm{m} \approx 30\%$). Taking advantage of the MiniMPL's portability, we demonstrate the feasibility of measuring the detailed horizontal structure of the mixing layer by automobile. We compare our observations to PBL heights from sonde launches, NARR reanalysis, and a custom WRF model developed for

GHG monitoring in Los Angeles. NARR and WRF PBL heights at Pasadena are both systematically higher than measured, NARR by 2.5 times; these biases will cause proportional errors in GHG flux estimates using modeled transport. We discuss how sustained lidar observations can be used to reduce flux inversion error by selecting suitable analysis periods, calibrating models, or characterizing bias for correction in post-processing.

1. Introduction

Improved understanding of sources, sinks, and controlling processes of CO_2 and other greenhouse gases (GHGs) will require robust methods for estimating surface fluxes. Observations of GHG concentrations capture the influence of known and unknown sources and sinks alike, making these observations an important complement to models and inventories. Top-down GHG inversions have been used for some time to estimate fluxes on global [*Tans et al.*, 1990], continental [*Bousquet et al.*, 2000], and regional [*Lauvaux et al.*, 2015; *Peters et al.*, 2007; *Schuh et al.*, 2010] scales, and there is increasing focus on bringing a similar approach to individual cities [*McKain et al.*, 2012; *Lauvaux et al.*, 2013; *Breon et al.*, 2014; *Turnbull et al.*, 2015]. However, relating observed concentrations to surface fluxes requires a representation of atmospheric transport. On the regional and urban scales, the extent and variablity of vertical mixing is a dominant source of uncertainty [*McKain et al.*, 2012] that can easily overwhelm the effects of instrument error. It is therefore critical to represent vertical mixing accurately.

The spatiotemporal structure of vertical mixing and diffusion can be complex. However, it can be useful to approximate gases recently emitted from the surface as being confined to and uniformly distributed throughout a near-surface layer. A cluster of related concepts – atmospheric or planetary boundary layer (PBL), convective boundary layer, mixed layer – are commonly used to describe the part of the atmosphere which "responds to surface forcings with a timescale of about an hour or less." [*Stull*, 1988] Various specific definitions of these layers are in use [*Seibert et al.*, 2000], some referring to thermodynamic variables and others directly to mixing or turbulence. Layers identified by different definitions can be conceptually distinct and therefore need to be considered differently.

The layer relevant to the dilution of GHGs is that within which substantial vertical mixing takes place. The time scale of mixing under turbulent conditions has been estimated at tens of minutes [*Stull*, 1988; *van Stratum et al.*, 2012; *Janssen and Pozzer*, 2015]. Nonetheless, a fully well-mixed equilibrium may not exist; we therefore follow *Seibert et al.* [2000] in referring to the *mixing layer*. When we refer to the *mixing height* or *mixing depth*, we mean the altitude of the top of the mixing layer. In addition to GHGs, the mixing depth also controls the dilution of aerosols and of other trace gases produced primarily within the mixing layer, including those that contribute to poor air quality. It is well known that shallow mixing contributes to air quality exceedances as these species are trapped near the surface, and observations such as those presented here can help in defining the presence of these conditions.

While it is difficult to measure the vertical distribution of GHGs directly, especially on an ongoing basis, we can measure the mixing height by observing the distribution of aerosol. Lidar systems measure the backscatter of a laser from particulate matter in the atmosphere, providing a vertical (or skew) profile of the concentration of scattering particles. We make use of such an instrument, the Sigma Space Mini-Micropulse LiDAR (MiniMPL), as well as a Vaisala CL51 ceilometer. These and other remote sensing instruments benefit from continuous operation, making observations at a rate of once per minute or more. New models like the MiniMPL are smaller and more portable than earlier research lidars and have better signal-to-noise performance than ceilometers.

Note that the mixing layer may not always coincide with the boundary layer commonly diagnosed by applying thermodynamic criteria to data from radiosondes (e.g. using the parcel method). In Pasadena, comparison to results from a series of sonde launches suggests that the mixing depth is related to, though not identical with, the depth of the boundary layer as defined using thermodynamic criteria. This finding is consistent with past results. Working in Indiana and the Amazon basin, respectively, *Coulter* [1979] and *Martin et al.* [1988] found that mixing depths determined using lidar observations were similar to and well-correlated with, though generally somewhat higher than, those determined from temperature profiles. *Marsik et al.* [1995] found that mixing depths from lidar in Atlanta were slightly *lower* that those measured using sondes. We discuss the comparison to sonde data in greater detail in section 3.3.

Given the high frequency of observations, operational use of lidar to measure the mixing height benefits from an at least partially automated method of analysis. A variety of schemes have been used. The simplest, the gradient method [*Endlich et al.*, 1979], searches for the maximum (most negative) vertical gradient of the backscatter signal, indicating a sudden decrease in density of scatterers. Related is the inflection point method [*Menut et al.*, 1999], which searches for zeros of the second spatial derivative of the backscatter. The wavelet method [*Ehret et al.*, 1996; *Davis et al.*, 1997, 2000; *Baars et al.*, 2008], which we use, is a refinement of the gradient method that takes into account the typical spatial scale of the boundary region at the top of the mixing layer. The variance method [*Hooper and Eloranta*, 1986; *Menut et al.*, 1999] identifies the entrainment zone at the top of the mixing layer by detecting a maximum in the temporal variance of backscatter, indicating the presence of turbulent vertical mixing. The idealized-profile method [*Steyn et al.*, 1999;

Eresmaa et al., 2006; *Münkel et al.*, 2006] attempts to fit the vertical backscatter profile to an ideal representation of aerosol density in and above the mixing layer – typically an error function. Some studies have applied a combination of methods: for example, using the gradient to refine a spatially [*Lammert and Bösenberg*, 2006] or temporally [*Hennemuth and Lammert*, 2006] coarse estimate generated by the variance method, or using gradient methods to select a number of candidate heights, then selecting between them by minimizing disagreement with a physical model [*Di Giuseppe et al.*, 2012].

In any method, the most serious challenge in automated mixing layer detection is to distinguish between the mixing layer top and other similar boundaries in the atmosphere, such as fog, low clouds, or residual layers of scatterers remaining aloft from previous days [*Haeffelin et al.*, 2012; *Lewis et al.*, 2013]. One approach to this challenge is to use the automated system only to generate a set of candidate heights and then rely on a human expert to distinguish between them. A person with some knowledge of atmospheric physics can often, though not always, identify the top of the mixing layer by visual inspection of a whole day's backscatter data. We take a different approach, aiming to automate the entire process in order to allow for long-term continuous operation. Following recent work [*Gan et al.*, 2011; *Lewis et al.*, 2013], we apply criteria that constrain the detected boundary to behavior that is physically reasonable, and we automatically detect and exclude conditions in which the instrument beam is blocked by fog or clouds. Finally, modifying a method introduced by *Lewis et al.* [2013], we implement a voting scheme, processing the day's data in several different ways and interpreting the degree of concurrence as a measure of confidence that our algorithm has selected the correct boundary.

In section 2, we describe the backscatter data, the instrument used to obtain them, and our automated method for extracting the mixing height. Section 3 presents our findings as to the climatological mixing state in the LA area and its temporal and spatial variation. We compare the results obtained with the MiniMPL to mixing depth estimates from a ceilometer, to a one-day sonde intensive, and to PBL heights from models and reanalysis. Finally, in section 4, we discuss the implications of our work for GHG flux estimation and suggest possible future applications.

2. Method

2.1. Instrumentation

We collected aerosol backscatter data using a Sigma Space Mini-Micropulse LiDAR (MiniMPL) operating at the Caltech campus in Pasadena, California. The MiniMPL is a compact version of the standard MPL, also manufactured by Sigma Space, that populates the NASA MPLNET lidar network. The MiniMPL inherits many of the design features of the MPL, such as a fiber coupled detector and robust optical train. Compared to the MPL, the MiniMPL reduces the power-aperture product to minimize cost, size, weight, and power requirements. As a result, detection range is limited to the troposphere while the MPL measures into the stratosphere. For tropospheric applications such as GHG flux estimation and air quality monitoring, however, the MiniMPL is designed to match the data quality of a standard MPL.

The MiniMPL transceiver shown in Fig. 1 weighs 13 kg and measures 380 x 305 x 480 mm in width, depth and height. The system consists of a laptop and the lidar transceiver, which are connected by a USB cable and consume 100W during normal operation. The whole system fits in a storm case with a telescopic handle and wheels that can be checked in as regular luggage during a domestic or international flight. The system's portability allows for applications that would not be possible with the standard MPL. In section 3.6, we demonstrate the feasibility of operating the MiniMPL out of a moving car, enabling us to observe the spatial structure of the mixing layer without the use of aircraft.

The MiniMPL's Nd:YAG laser emits polarized 532 nm light at a 4 KHz repetition rate and 3.5 uJ nominal pulse energy. The laser beam is expanded to the size of the telescope aperture (80 mm) to satisfy the eye safe requirements in ANSI Z136.1.2000 and IEC 60825 standards. Laser light is scattered back toward the instrument by particles and molecules in the atmosphere and collected by an 80mm diameter receiver. Distance to the scattering event is calculated from the time of flight. The instrument reports the number of scattering events recorded during a user-defined accumulation time (in our case, 30 s) originating in each vertical bin. We use a vertical range resolution of 30 m. Although this study does not make use of it, the MiniMPL also measures the depolarization [*Flynn et al.*, 2007] of the scattered light with a contrast ratio greater than 100:1.

The receiver uses a pair of narrowband filters with bandwidth less than 180 pm to reject the majority of solar background noise. The filtered light is then collected by a 100 um multimode fiber and fed into a Silicon Avalanche Photodetector (Si APD) operating in photon-counting mode (Geiger mode). Photon-counting detection enables the MiniMPL design to be lightweight and compact with high signal-to-noise ratio (SNR) throughout the troposphere.

To further maximize the SNR, MiniMPL uses a coaxial design; the transmitter and receiver Field of View (FOV) overlap with each other from range zero. This design eliminates the need for a wide FOV in order to minimize the overlap distance as in some biaxial lidar systems [*Kuze et al.*, 1998]. A wide FOV can result in measuring multiple scattering from aerosol [*Spinhirne*, 1982] and can distort depolarization measurements [*Tatarov et al.*, 2000]. On the other hand, a very narrow receiver FOV could make the lidar system sensitive to external factors like shock, vibration and temperature, making the system unsuitable for field deployment. The design of MiniMPL balances the above requirements and constraints, with an FOV of 240 urad.

Additional technical specifications of the MiniMPL, along with those of the standard MPL for comparison, are given in table 1.

2.2. Calibration

The raw event count reported by the MiniMPL must be calibrated and normalized in order to arrive at the quantity of interest, Normalized Relative Backscatter (NRB), which is approximately proportional to the concentration of scatterers at a given distance above the instrument. First, the event count is corrected for the deadtime of the detector, a period after each photon incidence during which no additional photons can be detected. The likely number of missed incidences can be extrapolated from the rate of detected photons. After the deadtime correction, the background (no laser light) value is subtracted. The event rate is then scaled by the laser pulse energy, which prevents changes in pulse energy from appearing as variation of the measured backscatter. Next, a correction is applied to account for laser light, called afterpulse, that strikes the inside of the instrument and returns to the detector without interacting with the atmosphere.

Finally, two corrections account for the fraction of scattered photons that are intercepted by the detector. The solid angle subtended by the collecting lens is inversely proportional to the square of the distance to the scattering event, so the event rate is multiplied by r^2.

Since the MiniMPL laser beam overlaps with the receiver field of view from range zero, there is no need for an overlap correction in the sense required by a biaxial instrument. However, because not all of the light incident on the collecting lens is focused onto the photon counter, a geometric factor calibration is still required. For historical reasons, this factor is also referred to as an overlap correction.

These steps are summarized in the following calibration equation:

$$B_{nr} = \left[\frac{x(z)C(x(z)) - bC(b)}{E} - \frac{x_{ap}(z)C(x_{ap}(z)) - b_{ap}C(b_{ap})}{E_{ap}} \right] \frac{z^2}{O(z)} \tag{1}$$

where $x(z)$ is the raw event rate signal at distance z from the instrument, $C(x)$ is the deadtime correction factor for event rate x, b is the background, E is the laser pulse energy, $x_{ap}(z)$ is the afterpulse signal at the time corresponding to distance z, b_{ap} is the background of the afterpulse signal, E_{ap} is the energy of the afterpulse, $O(z)$ is the overlap correction factor at distance z, and B_{nr} is the Normalized Relative Backscatter (NRB). An example vertical profile of NRB can be seen in Figure 3. In order to reduce the impact of short time scale fluctuations on our mixing depth retrieval, we apply a additional two-minute sliding average to the NRB values already aggregated to a thirty-second accumulation time by the instrument.

2.3. Observations

The MiniMPL collected backscatter data at Caltech on 530 days between August 1, 2012 and October 23, 2014, operating between dawn and dusk. Of those, 54 included data gaps of longer than one hour, including late starts to data collection, persistent midday rain or fog, or obstruction of the beam by obstacles. We exclude those days from the analysis. The remaining 476 days are distributed across all months other than July.

Table 2 shows the number of days of data by month as well as the concurrence scores of the mixing depth estimates (see section 2.4).

We analyze backscatter data in daily increments. Over the course of the day, the changing backscatter profile gives a picture of the distribution of scatterers in the lower atmosphere (Figure 2a). In the Los Angeles area, the scattering signal is typically quite strong due to the high levels of anthropogenic aerosols. Since they are produced primarily within the mixing layer, aerosols are concentrated near the surface. During the day, solar heating of the surface drives vertical mixing, causing the mixing layer to deepen and carrying aerosols to higher altitudes. As surface heating decreases in the late afternoon, the region of active vertical mixing shrinks, but the aerosols may remain aloft for some time. Frequently, aerosols carried aloft by one day's mixing can still be observed the following day in a residual layer disconnected from the surface. In the coastal mountain environment of Los Angeles, aerosols can also be carried above the mixing layer by the dominant circulation pattern, resulting in a sometimes complex stratification structure with thin, lofted aerosol layers [*Lu and Turco*, 1994, 1995].

2.4. Analysis

We use a Haar wavelet covariance method to identify boundaries between layers with high and low aerosol density. At a given height z, the wavelet covariance w is given by integrating the product of the backscatter profile with a Haar wavelet H centered at z:

$$w(z) = \frac{1}{d}\int \mathrm{d}z'\; B_{nr}(z')H(z,z',d) \tag{2}$$

$$H(z,z',d) = \begin{cases} 1 & z - d/2 < z' < z \\ -1 & z < z' < z + d/2 \\ 0 & |z' - z| > d/2 \end{cases} \tag{3}$$

The width d, or dilation, of the wavelet is chosen to correspond to the typical size of the transition zone at the top of the mixing layer, 200 m. As illustrated in Figure 3, the covariance is highest where the backscatter decreases rapidly with height. Because aerosols are concentrated within the mixing layer, such a rapid decrease in backscatter occurs at the top of the mixing layer. We therefore use high wavelet covariance values to identify the mixing layer top.

In order to increase the likelihood of detecting the mixing layer top rather than some other boundary – for example, a structure within the mixing layer or a residual layer of aerosols further aloft – each day's data is considered as a whole. Call the set of times during a single day at which backscatter data is available $t_1, \ldots, t_{max}$. First, designate a single time t_k and compute the altitude $z_k(t_k)$ at which the Haar covariance is maximum. Any later timepoints are then considered in order, beginning with t_{k+1}. The altitude z_k of the detected boundary is constrained to vary at a rate no faster than v. For the MiniMPL, we set $v = 100$ m/min, a conservative upper bound on typical rates of change of the mixing layer height [*Stull*, 1988]. This is equivalent to setting the Haar covariance to zero outside the range $(z_k(t_{i-1}) - v(t_i - t_{i-1}), z_k(t_{i-1}) + v(t_i - t_{i-1}))$. In addition, a multiplicative bias factor is applied to suppress the Haar covariance for unlikely but possible rates of change, decreasing linearly from one at $(2/3)v$ to zero at v. Similarly, any timepoints earlier than t_k are considered in reverse order, beginning with t_{k-1}.

For $t_k = t_1$, an additional physical constraint is applied: the mixing layer top must begin each day within 500 m of the ground. This aids in selecting a boundary that is continuous with the top of the nocturnal boundary layer, as the mixing layer should be.

This process is repeated for five values of t_k distributed evenly throughout the day, including the earliest time t_1 and the latest time t_{max}. The result is a set of estimates $z_{k_1}(t), \ldots, z_{k_n}(t)$ of the mixing height as a function of time.

A voting procedure is then used to select one estimate from the set. First, estimates are checked for pairwise agreement according to one of several criteria. In this study, we consider two estimates to be in agreement if they differ by no more than one unit of instrument vertical resolution (30 m for the MiniMPL as we operate it) as to the maximum depth of the mixing layer during the midday period. This criterion is optimized for determining that maximum; other criteria, such as agreement to within a tolerance over a specified fraction of the data period, might be better suited for other purposes. Next, this pairwise agreement is used to calculate a concurrence score for each estimate. An estimate E has a concurrence score equal to the fraction of all estimates that agree with E according to the selected criterion – see Figure 2 for an example. A 3/10 penalty is applied to the concurrence score of any estimate that violates the start-of-day condition, i.e. that puts the mixing layer top above 500 m at the start of the day. This was already forbidden during processing for the estimate beginning at t_1, but it may occur in other cases, and it generally indicates that the estimate has been fooled by a residual layer. After applying the penalty, the estimate with the highest score is selected for reporting, and the concurrence score can be used as a measure of confidence. Concurrence ties are broken by selecting the estimate with the earliest start time t_k; note that for concurrence scores of better than one-half, the tied estimates necessarily agree as to the chosen criterion. We recommend excluding estimates with scores less than one-half.

Because fog or clouds can completely extinguish the instrument beam, preventing any information from being returned from higher altitudes, it is important that our processing algorithm be able to detect this circumstance. Under foggy conditions, the altitude of highest Haar wavelet covariance does not represent the top of the mixing layer – in fact, there likely is no mixing layer – but only the maximum altitude to which the beam was able to pierce the fog before being extinguished. This situation is common in Pasadena in the early morning. We detect fog by checking directly for beam extinction, i.e. a layer of very high backscatter values with close to zero signal from above, and do not report any mixing height while fog is present.

Although the altitude of maximum Haar covariance on a foggy morning does not represent the mixing layer top, it remains useful, since that altitude transitions smoothly into the mixing layer top as the fog burns off. Fog, clouds, or rain that occur in the middle of the day are more problematic, since they often produce discontinuous changes in signal. We treat such occurrences as data gaps, and we exclude days on which gaps, including fog or rain, persist for too much of the total data period. In any case, we report for each day the maximum length of any gap in data, including instrument malfunction, a late start to data gathering, or beam extinction. It is important to check the maximum gap length before making use of the data, and to establish a standard for maximum allowable gap length, since long gaps can produce nonsensical results.

3. Results

3.1. Climatology and Variation

On the basis of our estimates, we emphasize the very large daily variability in the mixing height in the LA basin. The maximum depth of the mixing layer in afternoon may differ by a factor of two from one day to the next. On average, the greater degree of insolation does produce deeper afternoon mixing layers in summer than in winter. Using backscatter data from the MiniMPL on days with concurrence scores of 4/5 or higher and without gaps longer than 1 hr, we find the mean afternoon maximum mixing depth to be 770 m AGL in summer (June and August) and 670 m AGL in winter (December-February).

However, as illustrated in Figure 4, this seasonal difference is overwhelmed by the very large day- to week-scale variability. Within-season standard deviations in afternoon maximum mixing height are about 220 m in both summer and winter, representing 29% and 32% of the means, respectively. Similarly, a given day's mixing height cannot reliably be extrapolated from measurements made on previous days. Across 105 cases across all seasons in which we achieve concurrence scores of 4/5 or higher on both of two consecutive days, the root-mean-square difference in afternoon maximum mixing depth at Caltech is 230 m.

The high variability reinforces that applications of climatological mixing depth values are subject to large uncertainties; sustained observations like those we present here can quantify those uncertainties. Such observations can also be used to calibrate models or to choose between parameterization schemes in meteorological models, as we discuss in sections 3.4 and 3.5. Comprehensive comparisons to a model and/or to other meteorological observations over a long period could also provide a more granular understanding of the mixing dynamics. A robust explanation is needed for the variation we observe, which takes place too consistently and on too short a time scale to be attributed solely to unusual events such as forest fires or the LA basin's periodic Santa Ana winds.

3.2. Ceilometer

Alongside the MiniMPL, we also operated a Vaisala CL51 ceilometer (a successor instrument to the CL31, for details see e.g. *Münkel and Rasanen* [2004]; *McKendry et al.* [2009]; *Münkel et al.* [2011]) at the same site. The measurement principle of the ceilometer is similar to that of the lidar, but the overlap correction and other calibration steps are performed by proprietary software not visible to or modifiable by the user [*Wiegner et al.*, 2014]. The resulting quantity is referred to simply as the backscatter profile. The CL51 operates at 910 nm, in the near-infrared; it uses a 16 s temporal bin and a 10 m vertical range resolution.

We apply a version of the same algorithm to estimate mixing depths based on ceilometer backscatter data as we use with the MiniMPL . An example is shown in Figure 2b for comparison to the MiniMPL results on the same day. As is visible in the figure, especially in regions of low backscatter signal, the ceilometer's signal-to-noise performance is not as good as that of the MiniMPL. As a result, some adjustments are necessary. First, the maximum allowed rate of change v in the mixing layer height must be relaxed; for the ceilometer we set it to 150 m/min. This change is necessary because noise can temporarily disguise a change in the boundary location; the algorithm must be able to "snap back" to the true location of the boundary even after it has moved some distance away.

Second, the ceilometer tends to show an unrealistically large signal in the near field. *He et al.* [2006] note a similar artifact, which they attribute to an imperfectly corrected overlap error [see also *Wiegner et al.*, 2014]. Such errors are caused by differences in the optical geometry of the outgoing beam aperture and the detector that collects scattered photons. Because the erroneous backscatter signal associated with the artifact decays very rapidly, it has a high wavelet covariance. The algorithm therefore tends to detect the artifact in place of the real boundary, estimating the mixing depth at the lowest possible altitude.

To solve this problem, and to dampen noise in general, it is standard to take the logarithm of the ceilometer backscatter data prior to processing. Although physically unmotivated, this preprocessing step flattens out large signals, decreasing the influence of the low-altitude artifact. Figure 5 shows an example of a case in which the log transform allows the algorithm to detect the correct boundary. However, the log transform also introduces an offset. It suppresses the magnitude of the gradient of the backscatter more where backscatter values are higher:

$$\frac{\mathrm{d}}{\mathrm{d}z}\log(b(z)) = \frac{1}{b(z)}\frac{\mathrm{d}b(z)}{\mathrm{d}z} \tag{4}$$

Backscatter decreases with height in the transition from the mixing layer to the free troposphere above, so the strongest gradient in $\log(b(z))$ generally occurs at a higher altitude than the strongest gradient in $b(z)$. This effect carries over to the wavelet covariance method, causing a positive offset of about 50 m. The offset is due to a methodological choice to identify the altitude of greatest relative change in scattering, not a difference in physical reality. It should therefore be noted and compensated for in comparisons with estimates that identify the altitude of greatest absolute change, i.e. those that do not employ a log transform.

The effect of the log transform on the whole dataset is shown in Figure 5. There are two distinct populations. On some days, the low-altitude artifact traps the maximum mixing depth at the bottom of the instrument range. Applying the log transform removes the effect of the artifact, allowing the true mixing depth (which is variable) to be detected. On days on which the algorithm is not fooled by the artifact, the offset introduced by the log transform is visible: applying the transform results in an average increase of about 50 m in the estimated mixing depth.

Even with adjustments, our confidence in mixing height estimates derived from the ceilometer is not as high as in those derived from the MiniMPL. One proxy for confidence in a given day's results is the degree of concurrence among estimates in the voting procedure (see section 2). As can be seen in Figure 6, the MiniMPL achieves unanimity (concurrence score of 5/5) or near-unanimity (score of 4/5) on 51% days for which data is available. By contrast, the ceilometer achieves a score of 4/5 or better on only 36% of days. It is for this reason that we focus our results on estimates derived from MiniMPL observations.

3.3. Sonde Comparison

In September 2012, a one-day intensive campaign of sonde launches was conducted for comparison to mixing layer information from the MiniMPL. Sondes were launched every three hours between 7:00am and 7:00pm local time. The results are displayed in Figure 7. In each case, the PBL height is extracted from the sonde using the method of *Heffter* [1980]. At 7:00am, morning fog is still present and the mixing layer has not yet developed. At 10:00am, 1:00pm, and 4:00pm, the mixing height identified using the backscatter data coincides with the sonde-derived PBL height to within 150 m. Since the top of the mixing layer is in fact a transitional zone of 100 to 200 m thickness, it should not be considered to have a well-defined exact location. Some discrepancy should therefore be expected even between methods that detect substantially the same layer. In this one-day comparison, the backscatter method displays no identifiable systematic bias with respect to the sonde method; of course, the comparison presented is too limited to conclude that no bias exists.

We emphasize that, while the sonde comparison provides some confidence that, at least during the day, the layer in which elevated aerosol levels are present does correspond to the thermodynamic boundary layer, it is in any case the former that is of most interest for interpreting atmospheric concentrations of trace gases. For the purpose of linking atmospheric measurements to emissions rates, the important question is what part of the atmosphere should be considered in contact with the surface. In other words, through what volume are species emitted from the surface dispersed?

By 7:00pm, the mixing layer has begun to collapse and the structure is becoming more complicated. Two distinct boundaries are visible in both the potential temperature profile and the backscatter distribution, and both methods select the higher of these. Indeed, the day's aerosol emissions are distributed up to the higher boundary at 920 m. However, with the decrease in solar heating to drive vertical motion, the upper part of the identified layer (above about 500 m) is probably no longer interacting with the surface. Our method has therefore failed to detect a region of substantial, active vertical mixing. This case serves as a reminder that the mixing layer concept is not always straightforwardly applicable, particularly in the evening as vertical mixing tapers off. Care should be taken in interpreting and applying our or any other mixing depth estimates around sundown, even on days – like this one – with otherwise robust retrievals.

3.4. North American Regional Reanalysis (NARR) Comparison

GHG flux inversion studies typically make use of PBL heights derived from meteorological models or reanalysis products. We compare afternoon maximum mixing depth estimates based on MiniMPL data to PBL height estimates from the Weather Research and Forecasting model (WRF) and the North American Regional Reanalysis (NARR).

NARR is a reanalysis product providing a variety of atmospheric and surface variables over North America at 32 km spatial resolution and at 3 hour intervals [*Mesinger et al.*, 2005]. We find a large and persistent difference between afternoon maxima of MiniMPL-derived mixing depths at Caltech and PBL height estimates at the nearest NARR grid location. Figure 8 shows the distributions of these quantities over 227 days on which the MiniMPL estimate achieves a concurrence score of at least 4/5 and without data gaps longer than one hour. The maximum NARR PBL height exceeds the maximum MiniMPL-derived mixing depth on all but one day, differing by a factor of two or more on 63% of days. Summary statistics are in Table 3.

Interestingly, although maximum NARR PBL heights are an average of 2.5 times MiniMPL derived mixing depths, the two quantities are similarly distributed. Both show substantial variability, with standard deviations about 32% of the respective means, and both are skewed toward high values, with skewness 0.84 (NARR) and 0.88 (MiniMPL). However, NARR does not reproduce the detailed timing of this variability. Even after scaling maximum NARR PBL heights down by a factor of 2.5 to account for the mean difference, a root-mean-square difference of 360 m remains between scaled NARR estimates and MiniMPL estimates on the same days. This is almost as large as the RMS difference of 370 m in a sample of 10^6 random pairs of MiniMPL estimates and scaled NARR estimates.

We can attribute NARR's failure to accurately represent the boundary layer in Pasadena at least in part to its coarse spatial grid. The meteorology of the Los Angeles basin is strongly influenced by the coastal mountain topography (see Figure 9), resulting in a complex pattern of circulation [*Lu and Turco*, 1994, 1995]. It comes as no surprise that a product unable to resolve the rapid changes in elevation will struggle to predict PBL heights in this environment. If NARR is used to drive a transport model for GHG flux estimation in Los Angeles or in other areas with meteorology strongly influenced by the detailed topography, careful evaluation and correction of mixing depth biases will be critical for avoiding large errors. Since a biased mixing depth results in a proportional bias in flux estimates (see the general argument in section 4), we would expect a 250% bias in an LA flux inversion using NARR.

3.5. Weather Research and Forecasting (WRF) Comparison

Given the difficulty posed by the rapidly-varying topography of the LA basin, one might expect a high-resolution model to better represent the mixing dynamics. We compare mixing depth estimates from MiniMPL data taken during a deployment of the instrument in October-November 2015 to PBL heights from such a high-resolution model, a WRF setup developed specifically for the Los Angeles environment by *Feng et al.* [2016] to simulate CO_2 concentrations. The model is initialized with NARR and with sea surface temperatures from NCEP and uses three nested domains, with the innermost domain covering the LA basin at a resolution of 1.3 km. Using observations from the intensive Calnex campaign in 2010, including aircraft and ceilometer PBL measurements, *Feng et al.* [2016] tested a variety of WRF configurations. We employ only the MYNN_UCM_d03 configuration, which they found to minimize errors.

We redeployed the MiniMPL to Caltech for the three-week period of October 21 to November 9, 2015. Of these twenty days of observations, the mixing depth estimation algorithm achieves a concurrence score of 4/5 or better on six days and a score of 3/5 on another nine days. Although this comparison period is too short to allow robust statistical conclusions, we make some preliminary observations. Like NARR, WRF PBL heights show variability that is similar in relative terms to that of Mini-MPL derived mixing heights. Over the three-week comparison period, the standard deviation of maximum afternoon WRF PBL heights is 540 m, about 37% of the mean. However, WRF estimates PBL heights that are greater than MiniMPL-derived mixing depths on all but one afternoon. On average, afternoon maximum WRF PBL height exceeds afternoon maximum MiniMPL mixing depth by 730 m. Considering only days with high concurrence scores reduces the discrepancy considerably. The mean difference on days with scores of 4/5 or better is 380 m, suggesting that the concurrence voting scheme effectively identifies days that are easier to analyze.

The discrepancy we find between modeled PBL height and MiniMPL-derived mixing depth is surprising given the excellent agreement reported by *Feng et al.* [2016]. During the 2010 Calnex campaign period, they report a mean WRF-derived daytime PBL height (using the same MYNN_UCM_d03 configuration we use here) of 828.8 m, in good agreement with a mean mixing depth of 835.7 m obtained from ceilometer measurements using the gradient method. They also report substantially less variability in modeled PBL height than in measured mixing depth. Further work, including a model-data comparison covering a longer period, is clearly needed to resolve this perplexing difference. While such a comparison is beyond the scope of this study, we do note that NARR PBL height estimates for May-June 2010 are generally similar to those from our comparison period in October-November 2015, with a mean daily maximum of 2.1 km.

Our analysis here cannot distinguish between differences due to errors in mixing depth estimation, errors in modeled PBL depth, or conditions under which the mixing layer fails to correspond to the thermodynamic PBL. Still, it is prudent to expect that the same complex stratification which can cause the mixing depth estimation algorithm to fail might also indicate challenging conditions for the model. By selecting days with high concurrence scores, MiniMPL observations can be used to choose "golden days" for model analysis. Alternately, if a model is run over a long period, days with good agreement between the model and lidar estimates can be selected for flux estimation. For example, Figure 10 shows a pair of days which would not be readily distinguished on the basis of model results alone. The additional information provided by the lidar estimates lets us assign greater confidence to modeling on the day with good agreement (panel a) than that with poor agreement (panel b).

Sustained lidar can also inform the choice of model configurations or parameters, as *Feng et al.* [2016] and others [e.g., *Nehrkorn et al.*, 2013] have done with PBL observations from limited campaigns. In addition to increasing confidence in that choice simply by virtue of a larger volume of data, long-term observations can provide more detailed information about how model errors depend on season or on other meteorological conditions. For example, *Lewis et al.* [2013] found that PBL height as estimated by the general circulation model GEOS-5 differs most from that measured by the lidar network MPLNET in winter. Unlike sondes, lidar data can validate not only the depth of the mixing layer but also the timing of its development and collapse. That timing can be critical; for example, in an urban setting, a difference of one hour may determine whether the mixing layer begins to develop before, during, or after the emissions peak associated with the morning rush hour.

An alternate method for integrating mixing depth observations into flux estimation is to characterize a known model bias and correct for it after the modeling stage. *Zhao et al.* [2009] use three months of wind profiler measurements to derive an empirical relationship between observed and modeled PBL heights. They apply that relationship to scale down modeled PBL height before computing fluxes, reducing the residual error by a factor of 1.5. Among the advantages of postprocessing corrections of this kind are that they are simple to apply, allowing accuracy to be improved even in less detailed inversions, and that they can be combined with the strategies discussed above to further control any errors remaining after tuning model parameterization and/or selecting out "golden days."

3.6. Spatial Variation

Taking advantage of the MiniMPL's portability, we also conducted a one-time pilot mobile study in which backscatter data was collected over a period of about twenty minutes as the MiniMPL was transported due west toward the Pacific coast in the back of a passenger car. This observing strategy, which could not have been implemented with a full-size research lidar, is made possible by the compact size and low power requirements of the MiniMPL. Now that we have demonstrated its feasibility, we hope that this new approach will allow for both more regular mapping of the spatial structure of the mixing layer and more nimble mobile deployment of lidar in response to irregular events like fires and gas leaks.

The spatial profile of aerosol backscatter near the Pacific coast is shown in Figure 11. The transition between the shallow marine layer, which extends two to three kilometers onto land, and the convective regime that dominates further inland is clearly visible. The vertical structure in this case is simple, with a well-defined mixing layer of high backscatter adjacent to the ground and a sharp decrease in backscatter at the top of that layer. The mixing depth as estimated by the minimum backscatter gradient is indicated in the figure by the black circles (our retrieval algorithm is not suitable, since it relies on the temporal evolution of the boundary at a fixed location).

Figure 11 also shows the PBL height as predicted by WRF. The WRF prediction agrees well with the MiniMPL-derived mixing depth near the coast, but does not increase as sharply further inland. Unlike at Caltech, in this case the WRF PBL height is lower than the observed mixing depth. Repeated measurements of this kind could reveal whether the difference is consistent with time and at locations elsewhere along the coast, both in the immediate Los Angeles area and elsewhere, which could contribute to model development in the challenging coastal environment. Further work characterizing the coastal transition could also aid in understanding the fate of GHG emissions from sources like ports and marine industry.

4. Conclusions

Researchers have recognized that the representation of mixing dynamics is both critical for the interpretation of top-down emissions estimates and also a major source of uncertainty [e.g., *Newman et al.*, 2013; *Zimnoch et al.*, 2010]. *McKain et al.* [2012] advocate the use of column-integrated concentration measurements in urban studies, among other reasons in order to avoid the impact of mixing height errors. A common strategy [*Breon et al.*, 2014] is to rely only on observations made during midafternoon, when the mixing layer is at or near its maximum depth and the detailed timing of its dynamics are less important. But we observe even afternoon maximum mixing height in Los Angeles to vary substantially from day to day, typically differing from the seasonal mean by 30%.

A simple dimensional argument demonstrates the impact of such variations. Suppose that an instrument measures the in situ concentration of a trace gas at a particular location. This concentration is expressed as a molar fraction, or, equivalently given the local density of dry air, as a volume concentration C in moles of gas per unit volume, i.e. in n/L^3. The goal is to use the measurement to infer a surface flux F, expressed in moles of gas emitted or absorbed per unit area per unit time, i.e. as $n/(L^2T)$. On dimensional grounds, any method for relating the concentration to the flux must incorporate some temporal information, such as the time τ during which the sampled air mass was exposed to the flux, and also some vertical length scale.

The relevant vertical length scale is the mixing height h, which controls the height of the space into which the emitted gas is diluted. We therefore expect

$$F \propto hC/\tau \tag{5}$$

in which case an error in the mixing height h will result in a proportional error in the flux estimate. In detailed models, this picture is complicated to some degree by higher-order effects, e.g. the coupling between vertical motion and horizontal wind shear, but the essential proportionality remains. Applying sustained observations to control mixing depth errors, whether by validating models, choosing suitable periods for analysis, or characterizing and correcting for errors in postprocessing, is critical for accurate GHG flux estimation.

We have focused above on determining the depth of the mixing layer, especially at its afternoon maximum. But the mixing layer concept is not always applicable. Even when the mixing height is applicable, it does not fully describe the complex structure of the lower troposphere. The potential exists to extract much more information about that structure from lidar backscatter data. Among other applications, a more complete picture of the mixing state could contribute to our understanding of the transport of species emitted from the surface. Here we suggest one direction in particular for future work.

The mixing layer itself may exhibit internal structure. For example, in Los Angeles, the sea breeze circulation pushes near-surface air inland during the day. As a result, the air mass within the mixing layer over Pasadena in the afternoon has traveled over downtown in the preceding hours. The time scale of this horizontal motion, and the varying emissions rates and compositions from the traversed areas, may create a stratification, in which fresh emissions from Pasadena are concentrated in the lowest part of the mixing layer while those from downtown are more thoroughly mixed throughout. If we were able to observe and understand within-layer dynamics of this kind, we could much more precisely link trace gases observed in the atmosphere to their points of emission, allowing us to answer more specific questions about the sources and composition of emissions in the urban environment.

Since the lidar is primarily sensitive to aerosols and not to trace gases, the distribution of aerosol would need to be used as a proxy for the distribution of co-emitted trace gases, assuming that the two are transported within the mixing layer in a similar way, at least on short time scales and over small distances. That assumption would need to be tested before it could form the basis of any future work. Challenges notwithstanding, this is an exciting possibility for future applications, including more detailed validation of transport models and finer-scale attribution of emissions sources within complex urban environments like that of Los Angeles.

Acknowledgments. This work was supported by NASA under grant #NNN12AA01C. Portions of this work were performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. We thank Taylor Jones for assistance in setting up and operating the MiniMPL instrument and Athena Sparks for help with data preprocessing. We thank Vineet Yadav for generating and providing WRF model output. NARR data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. The authors would also like to thank the Megacities Carbon Project team for useful discussion and feedback. Mixing depth data will be available through the Megacities Carbon Project portal at https://megacities.jpl.nasa.gov. To obtain a copy of the analysis system used to generate the estimates, please contact the authors at johnware@umich.edu.

References

Baars, H., A. Ansmann, R. Engelmann, and D. Althausen (2008), Continuous monitoring of the boundary-layer top with lidar, *Atmospheric Chemistry and Physics*, *8*(23), 7281, doi:10.5194/acp-8-7281-2008.

Bousquet, P., P. Peylin, P. Ciais, C. Le Quéré, P. Friedlingstein, and P. P. Tans (2000), Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science*, *290*, 1342–1346.

Breon, F. M., G. Broquet, Puygrenier, F. Chevallier, I. Xueref-Rémy, M. Ramonet, E. Dieudonne, M. Lopez, M. Schmidt, O. Perrussel, and P. Ciais (2014), An attempt at estimating Paris area CO 2 emissions from atmospheric concentration measurements, *Atmospheric Chemistry and Physics*, *14*, 9647–9703, doi:10.5194/acpd-14-9647-2014.

Coulter, R. L. (1979), A comparison of three methods for measuring mixing-layer height, *Journal of applied meteorology*, *18*, 1495–1499, doi:10.1175/1520-0450(1979)018¡1495:ACOTMF¿2.0.CO;2.

Davis, K. J., D. H. Lenschow, S. P. Oncley, C. Kiemle, G. Ehret, A. Giez, and J. Mann (1997), Role of entrainment in surfaceatmosphere interactions over the boreal forest, *Journal of Geophysical Research*, *102*, 29,219–29,230.

Davis, K. J., N. Gamage, C. R. Hagelberg, C. Kiemle, D. H. Lenschow, and P. P. Sullivan (2000), An Objective Method for Deriving Atmospheric Structure from Airborne Lidar Observations, *Journal of Atmospheric and Oceanic Technology*, *17*(11), 1455, doi:10.1175/1520-0426(2000)017.

Di Giuseppe, F., A. Riccio, L. Caporaso, G. Bonafé, G. P. Gobbi, and F. Angelini (2012), Automatic detection of atmospheric boundary layer height using ceilometer backscatter data assisted by a boundary layer model, *Quarterly Journal of the Royal Meteorological Society*, *138*, 649–663, doi:10.1002/qj.964.

Ehret, G., A. Giez, C. Kiemle, K. J. Davis, D. H. Lenschow, S. P. Oncley, and R. D. Kelly (1996), Airborne water vapor DIAL and in situ observations of a sea-land interface, *Contributions to Atmospheric Physics*, *69*, 215–228.

Endlich, R. M., F. L. Ludwig, and E. E. Ludwig (1979), An automatic method for determining the mixing depth from lidar observations, *Atmospheric Environment*, *13*, 1051–1056.

Eresmaa, N., A. Karppinen, S. M. Joffre, J. Räsänen, and H. Talvitie (2006), Mixing height determination by ceilometer, *Atmospheric Chemistry and Physics*, *6*(6), 1485–1493, doi:10.5194/acp-6-1485-2006.

Feng, S., T. Lauvaux, S. Newman, P. Rao, R. Ahmadov, A. Deng, L. I. Diaz-Isaac, R. M. Duren, M. L. Fischer, C. Gerbig, K. R. Gurney, J. Huang, S. Jeong, Z. Li, C. E. Miller, D. O'Keeffe, R. Patarasuk, S. P. Sander, Y. Song, K. W. Wong, and Y. L. Yung (2016), Los Angeles megacity: a high-resolution land-atmosphere modelling system for urban CO2 emissions, *Atmospheric Chemistry and Physics*, *16*, 9019-9045, doi:10.5194/acp-16-9019-2016.

Flynn, C. J., A. Mendoza, Y. Zheng, and S. Mathur (2007), Novel polarization-sensitive micropulse lidar measurement technique, *Optics Express*, *15*(6), 2785–2790, doi:10.1364/oe.15.002785.

Gan, C. M., Y. Wu, B. L. Madhavan, B. Gross, and F. Moshary (2011), Application of active optical sensors to probe the vertical structure of the urban boundary layer and assess anomalies in air quality model PM 2.5 forecasts, *Atmospheric Environment*, *45*, 6613–6621, doi:10.1016/j.atmosenv.2011.09.013.

Haeffelin, M., F. Angelini, G. P. Gobbi, Y. Morille, G. Martucci, C. D. O'Dowd, S. Frey, S. Lolli, L. Sauvage, I. Xueref-Rémy, B. Wastine, and D. G. Feist (2012), 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.

He, Q. S., J. T. Mao, J. Y. Chen, and Y. Y. Hu (2006), Observational and modeling studies of urban atmospheric boundary-layer height and its evolution mechanisms, *Atmospheric Environment*, *40*, 1064–1077, doi:10.1016/j.atmosenv.2005.11.016.

Heffter, J. L. (1980), Air resources laboratories atmospheric transport and dispersion model, *Tech. Rep. ERL ARL-81*, NOAA.

Hennemuth, B., and A. Lammert (2006), Determination of the atmospheric boundary layer height from radiosonde and lidar backscatter, *Boundary-Layer Meteorology*, *120*, 181–200, doi:10.1007/s10546-005-9035-3.

Hooper, W. P., and E. W. Eloranta (1986), Lidar measurements of wind in the planetary boundary layer: the method, accuracy and results from joint measurements with radiosonde and kytoon, *Journal of Climate and Applied Meteorology*, *25*, 990–1001.

Janssen, R. H. H., and A. Pozzer (2015), Description and implementation of a MiXed Layer model (MXL, v1. 0) for the dynamics of the atmospheric boundary layer in the Modular Earth Submodel System (MESSy), *Geoscientific Model Development*, *8*(3), 453–471, doi:10.5194/gmd-8-453-2015.

Kuze, H., H. Kinjo, Y. Sakurada, and N. Takeuchi (1998), Field-of-view dependence of lidar signals by use of Newtonian and Cassegrainian telescopes, *Applied Optics*, *37*(15), 3128–3132, doi:10.1364/ao.37.003128.

Lammert, A., and J. Bösenberg (2006), Determination of the convective boundary-layer height with laser remote sensing, *Boundary-Layer Meteorology*, *119*(1), 159–170, doi:10.1007/s10546-005-9020-x.

Lauvaux, T., N. L. Miles, S. J. Richardson, A. Deng, D. R. Stauffer, K. J. Davis, G. Jacobson, C. Rella, G. Calonder, and P. L. DeCola (2013), Urban emissions of CO2 from Davos, Switzerland: the first real-time monitoring system using an atmospheric inversion technique, *Journal of Applied Meteorology and Climatology*, *52*(12), 2654–2668, doi:10.1175/jamc-d-13-038.1.

Lewis, J. R., E. J. Welton, A. M. Molod, and E. Joseph (2013), Improved boundary layer depth retrievals from MPLNET, *Journal of Geophysical Research: Atmospheres*, *118*, 9870–9879, doi:10.1002/jgrd.50570.

Lu, R., and R. P. Turco (1994), Air pollutant transport in a coastal environment. Part I: Two-dimensional simulations of sea-breeze and mountain effects, *Journal of the Atmospheric Sciences*, *51*, 2285–2308, doi:10.1175/1520-0469(1994)051¡2285:APTIAC¿2.0.CO;2.

Lu, R., and R. P. Turco (1995), Air pollutant transport in a coastal environment—II. Three-dimensional simulations over Los Angeles basin, *Atmospheric Environment*, *29*, 1499–1518, doi:10.1016/1352-2310(95)00015-Q.

Marsik, F. J., K. W. Fischer, T. D. McDonald, and P. J. Samson (1995), Comparison of methods for estimating mixing height used during the 1992 Atlanta Field Intensive, *Journal of Applied Meteorology*, *34*, 1802–1814.

Martin, C. L., D. Fitzjarrald, M. Garstang, A. P. Oliveira, S. Greco, and E. Browell (1988), Structure and growth of the mixing layer over the Amazonian rain forest, *Journal of Geophysical Research: Atmospheres*, *93*(D2), 1361–1375, doi:10.1029/JD093iD02p01361.

McKain, K., S. C. Wofsy, T. Nehrkorn, J. Eluszkiewicz, J. R. Ehleringer, and B. B. Stephens (2012), Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, *Proceedings of the National Academy of Sciences*, *109*, 8423–8428, doi:10.1073/pnas.1116645109.

McKendry, I. G., D. Van der Kamp, K. B. Strawbridge, A. Christen, and B. Crawford (2009), Simultaneous observations of boundary-layer aerosol layers with CL31 ceilometer and 1064/532 nm lidar, *Atmospheric Environment*, *43*(36), 5847–5852, doi:10.1016/j.atmosenv.2009.07.063.

Menut, L., C. Flamant, J. Pelon, and P. H. Flamant (1999), Urban boundary-layer height determination from lidar measurements over the Paris area, *Applied Optics*, *38*, 945–954.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2005), North American Regional Reanalysis, *Bulletin of the American Meteorological Society*, *87*(3), 343–360, doi:10.1175/BAMS-87-3-343.

Münkel, C., and J. Rasanen (2004), New optical concept for commercial lidar ceilometers scanning the boundary layer, *Proceedings of SPIE*, *5571, Remote Sensing of Clouds and the Atmosphere IX*, 364–374, doi:10.1117/12.565540.

Münkel, C., N. Eresmaa, J. Räsänen, and A. Karppinen (2006), Retrieval of mixing height and dust concentration with lidar ceilometer, *Boundary-Layer Meteorology*, *124*(1), 117–128, doi:10.1007/s10546-006-9103-3.

Münkel, C., K. Schäfer, and S. Emeis (2011), Adding confidence levels and error bars to mixing layer heights detected by ceilometer, *Proceedings of SPIE*, *8177, Remote Sensing of Clouds and the Atmosphere XVI*, doi:10.1117/12.898122.

Nehrkorn, T., J. Henderson, M. Leidner, M. Mountain, J. Eluszkiewicz, K. McKain, and S. Wofsy (2013), WRF simulations of the urban circulation in the Salt Lake City area for CO2 modeling, *Journal of Applied Meteorology and Climatology*, *52*, 323–340, doi:10.1175/JAMC-D-12-061.1.

Newman, S., S. Jeong, M. L. Fischer, X. Xu, C. L. Haman, B. Lefer, S. Alvarez, B. Rappenglueck, E. A. Kort, A. E. Andrews, J. Peischl, K. R. Gurney, C. E. Miller, and Y. L. Yung (2013), Diurnal tracking of anthropogenic CO 2 emissions in the Los Angeles basin megacity during spring 2010, *Atmospheric Chemistry and Physics*, *13*, 4359–4372, doi: 10.5194/acp-13-4359-2013.

Peters, W., A. R. Jacobson, C. Sweeney, A. E. Andrews, T. J. Conway, K. Masarie, J. B. Miller, L. M. P. Bruhwiler, G. Petron, A. I. Hirsch, D. E. J. Worthy, G. R. var der Werf, J. T. Randerson, P. O. Wennberg, M. C. Krol, and P. P. Tans (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proceedings of the National Academy of Sciences*, *104*, 18,925–18,930, doi:10.1073/pnas.0708986104.

Schuh, A. E., A. S. Denning, K. D. Corbin, I. T. Baker, M. Uliasz, N. Parazoo, A. E. Andrews, and D. E. J. Worthy (2010), A regional high-resolution carbon flux inversion of North America for 2004, *Biogeosciences*, *7*, 1625–1644, doi:10.5194/bg-7-1625-2010.

Seibert, P., F. Beyrich, S. E. Gryning, S. Joffre, A. Rasmussen, and P. Tercier (2000), Review and intercomparison of operational methods for the determination of the mixing height, *Atmospheric Environment*, *34*, 1001–1027, doi:10.1016/S1352-2310(99)00349-0.

Spinhirne, J. D. (1982), Lidar clear atmosphere multiple scattering dependence on receiver range, *Applied Optics*, *21*(14), 2467–2468, doi:10.1364/ao.21.002467.

Steyn, D. G., M. Baldi, and R. M. Hoff (1999), The Detection of Mixed Layer Depth and Entrainment Zone Thickness from Lidar Backscatter Profiles, *Journal of Atmospheric and Oceanic Technology*, *16*(7), 953–959, doi:10.1175/1520-0426(1999)016.

Stull, R. (1988), *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers.

Tans, P. P., I. Y. Fung, and T. Takahashi (1990), Observational contrains on the global atmospheric CO2 budget, *Science*, *247*, 1431–1438, doi:10.1126/science.247.4949.1431.

Tatarov, B., T. Trifonov, B. Kaprielov, and I. Kolev (2000), Dependence of the lidar signal depolarization on the receiver's field of view in the sounding of fog and clouds, *Applied Physics B: Lasers and Optics*, *71*(4), 593–600, doi:10.1007/s003400000265.

Turnbull, J. C., C. Sweeney, A. Karion, T. Newberger, S. J. Lehman, P. P. Tans, K. J. Davis, T. Lauvaux, N. L. Miles, S. J. Richardson, M. O. Cambaliza, P. B. Shepson, K. Gurney, R. Patarasuk, and I. Razlivanov (2015), Toward quantification and source sector identification of fossil fuel CO 2 emissions from an urban area: Results from the INFLUX experiment , *Journal of Geophysical Research: Atmospheres*, *120*(1), 292–312, doi:10.1002/2014jd022555.

van Stratum, B. J. H., J. V.-G. de Arellano, H. G. Ouwersloot, K. van den Dries, T. W. van Laar, M. Martinez, J. Lelieveld, J.-M. Diesch, F. Drewnick, H. Fischer, Z. Hosaynali Beygi, H. Harder, E. Regelin, Sinha, J. A. Adame, M. Sorgel, R. Sander, H. Bozem, W. Song, J. Williams, and N. Yassaa (2012), Case study of the diurnal variability of chemically active species with respect to boundary layer dynamics during DOMINO, *Atmospheric Chemistry and Physics*, *12*(12), 5329–5341, doi:10.5194/acp-12-5329-2012.

Wiegner, M., F. Madonna, I. Binietoglou, R. Forkel, J. Gasteiger, A. Geiss, G. Pappalardo, K. Schafer, and W. Thomas (2014), 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.

U.S. Geological Survey, The National Map (2015), 3DEP products and services: The National Map, *3D Elevation Program Web page*, accessed 4 December 2015 at http://nationalmap.gov/3dep_prodserv.html

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, *Journal of Geophysical Research*, *114*, D16,302, doi: 10.1029/2008JD011671.

Zimnoch, M., J. Godlowska, J. M. Necki, and K. Rozanski (2010), Assessing surface fluxes of CO2 and CH4 in urban environment: a reconnaissance study in Krakow, Southern Poland, *Tellus B*, *62*, 573–580, doi:10.1111/j.1600-0889.2010.00489.x.

Figure 1. A complete MiniMPL lidar system consists of an optical transceiver (shown) and a laptop running data acquisition and post-processing software.

Figure 2. A sample day of backscatter data (heatmap) from the MiniMPL (panel a) and ceilometer (panel b, see section 3.2) with mixing heights as estimated by our algorithm (black symbols: majority opinion; green symbols: estimates initialized at other times of day). Prior to 8am, both instrument beams are completely extinguished near the surface; the algorithm recognizes the presence of fog and does not attempt to make an estimate. In the late afternoon – and in the morning in the case of the ceilometer – the various estimates disagree as to the mixing height, identifying two different boundaries. We report the majority opinion together with the degree of concurrence (4/5 for the MiniMPL, 2/5 for the ceilometer). Note that MiniMPL NRB values and ceilometer backscatter values do not use comparable scales.

Figure 3. An illustration of the wavelet method. The instrument returns a vertical profile of normalized relative backscatter (NRB, left). To compute the wavelet covariance at a given altitude z, the backscatter profile is integrated against a Haar wavelet centered at z (middle). In this example, the covariance is given by the difference in area between the orange (upper) shaded region and the blue (lower) region, which indicates the decrease in backscatter over the scale of the wavelet. The resulting Haar wavelet covariance is shown at right.

Figure 4. Solid curves: average diurnal cycles of mixing height in June-August (orange) and December-February (blue). Shaded regions: one standard deviation of between-days variability. Estimates according to the MiniMPL, retaining only days on which the concurrence score was at least 4/5.

Figure 5. (a) Representative backscatter profiles from the ceilometer, with (orange, triangular symbols) and without (blue, round symbols) applying a log transform. The horizontal lines show the corresponding mixing height as estimated by the algorithms: solid blue line, without transform; dashed orange line, with transform. Note the very high backscatter values at low altitudes in the untransformed data, which fool the algorithm into selecting an unrealistically low mixing height.

(b) Maximum afternoon mixing depths as estimated using ceilometer data with (vertical axis) or without (horizontal axis) applying the log transform. Days on which the untransformed data is affected by the low-altitude artifact are indicated by the dashed green ellipse; taking the log transform removes the effect of the artifact. On other days (indicated by the solid pink ellipse), the bias introduced by the transform is visible. The solid black line is the 1-1 line. Only days with concurrence scores of at least 3/5 are shown.

Figure 6. Degree of concurrence achieved by the algorithm using backscatter data from the MiniMPL (blue, solid) or from the ceilometer (orange, dashed), shown as a fraction of days on which both instruments were operating.

Figure 7. Orange: potential temperature profiles from sonde launches, with the corresponding PBL height as calculated using Heffter's method (horizontal dashed line). Blue with triangles: contemporaneous MiniMPL backscatter profiles, with the mean (center horizontal line) and range (shaded area) of the algorithmically-estimated mixing height over the 30-minute period surrounding the sonde launch.

Table 1. Technical specifications for the MiniMPL (used in this study) and the standard MPL.

Figure 8. Gaussian kernel density (smoothed relative frequency) of maximum afternoon mixing depth according to the MiniMPL (pink, left peak) and according to NARR (green, right peak) over 227 days with MiniMPL concurrence score at least 4/5. Solid vertical lines: median; dashed vertical lines: quartiles.

Figure 9. (a) Elevation map of the Los Angeles Basin [*U.S. Geological Survey*, 2015]. The labeled diamonds indicates the location of the measurement site at Caltech (in Pasadena). The solid line shows the route taken in the mobile study; the dashed line corresponds to the cross section in panel (b).
(b) Elevation cross section along the dashed line in panel (a); the longitude scale is the same for both panels.

Figure 10. Examples of days with good (panel a) and poor (panel b) agreement between MiniMPL-derived mixing depths (small circles) and PBL heights as estimated by WRF (large diamonds). NARR PBL heights (large triangles) show large discrepancies in both cases.

Figure 11. Heatmap: MiniMPL backscatter intensity near the Pacific coast (located at longitude −118.41). Small black circles: mixing depth as estimated by the gradient method using MiniMPL data. Large black diamonds: PBL height as estimated by WRF. Black curve at bottom: topography (same vertical scale).

Table 2. Number of days of MiniMPL data collection without gaps longer than one hour, by month and by concurrence score of the mixing depth estimation algorithm.

Month	Total	1/5	2/5	3/5	4/5	5/5
January	52	3	4	11	10	24
February	26	4	7	4	5	6
March	29	1	5	8	7	8
April	31	1	3	12	10	5
May	26	0	7	5	8	6
June	14	1	2	1	3	7
July	0	0	0	0	0	0
August	49	1	5	15	14	14
September	85	1	18	24	16	26
October	64	5	13	16	13	17
November	48	4	13	15	11	5
December	52	4	14	16	7	11
All Months	476	25	91	127	104	129

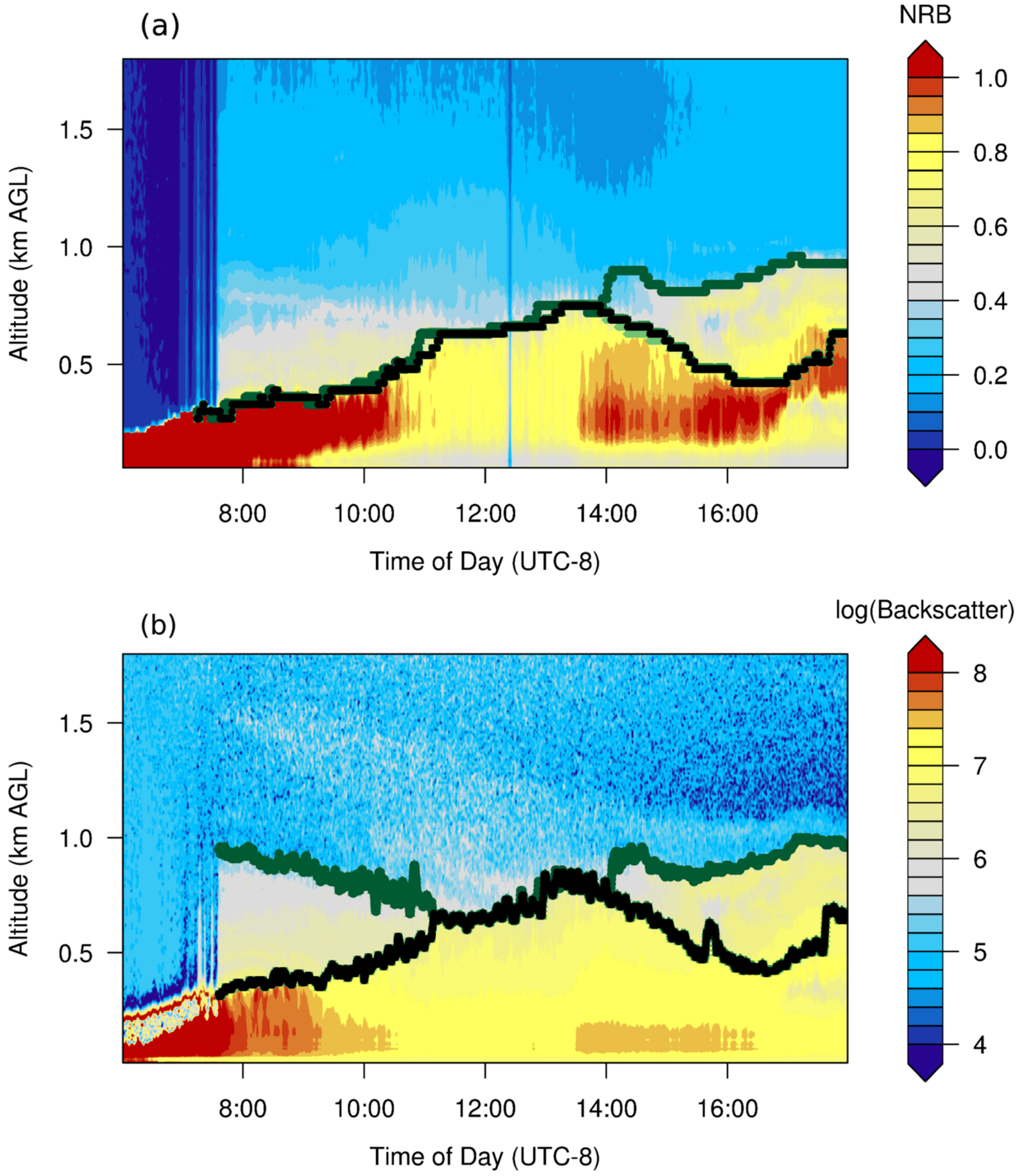
Table 3. Mean, median, 1st and 3rd quartiles, and standard deviation of afternoon maximum PBL height (NARR) or mixing depth (MiniMPL), in km AGL, over 227 days with concurrence score at least 4/5 and without data gaps longer than on hour.

Method	Q_1	Median	Mean	Q_3	σ	σ/Mean
MiniMPL	0.63	0.75	0.84	0.98	0.27	32%
NARR	1.46	1.84	1.92	2.20	0.62	32%

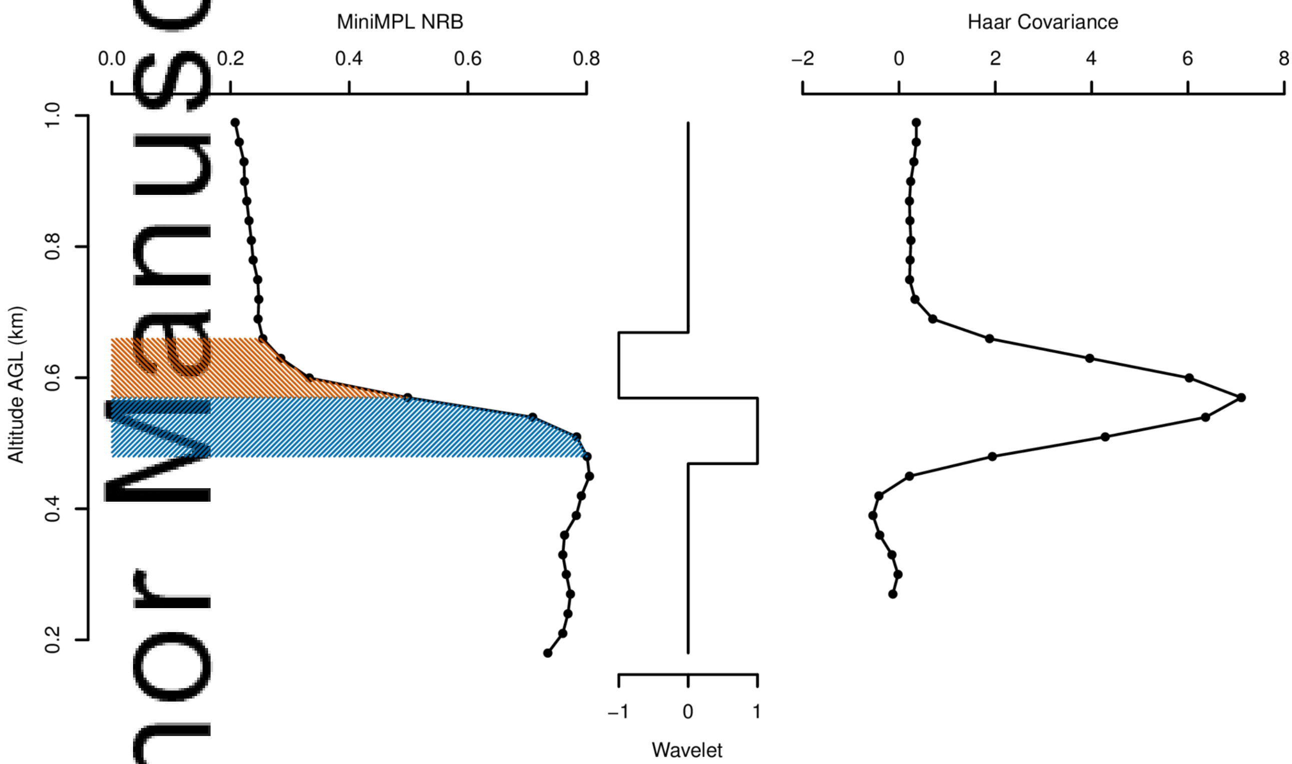
2016jd024953-f01-z-.eps

		MPL	MiniMPL
PERFORMANCE	range resolution	5/15/30/75 m (software programmable)	5/15/30/75 m (software programmable)
	minimum range	150 m	150 m
	accumulation time	1 sec - 15 min	1 sec - 15 min
	detection range	up to 25 km	up to 25 km
	polarization	standard	standard
	scanning	optional	optional
OPTICS	laser wavelength	532 nm	532 nm
	laser pulse energy	6 - 10 µJ @ 2500 Hz	3 - 4 µJ @ 2500 Hz
	eye-safety	ANSI Z136.1 2000 IEC 60825	ANSI Z136.1 2000 IEC 60825
	receiver diameter	178 mm	80 mm
	pump laser diode	fiber coupled user replaceable	
	detector	fiber coupled user replaceable	fiber coupled user replaceable
DIMENSIONS	size	300 x 350 x 850 mm	240 x 305 x 480 mm
	weight (portability)	25 + 2 kg	13 kg
DATA	operating system	Windows 7/10	Windows 7/10
	computer interface	USB	USB
	data transfer	LAN ethernet	LAN ethernet
ENVIRONMENT	temperature	NEMA-4 enclosure	NEMA-4 enclosure
	humidity	enclosure	enclosure
POWER	supply	100/240 V AC 50 - 6- Hz	100/240 V AC 50 - 6- Hz
	consumption	500 W	100 W

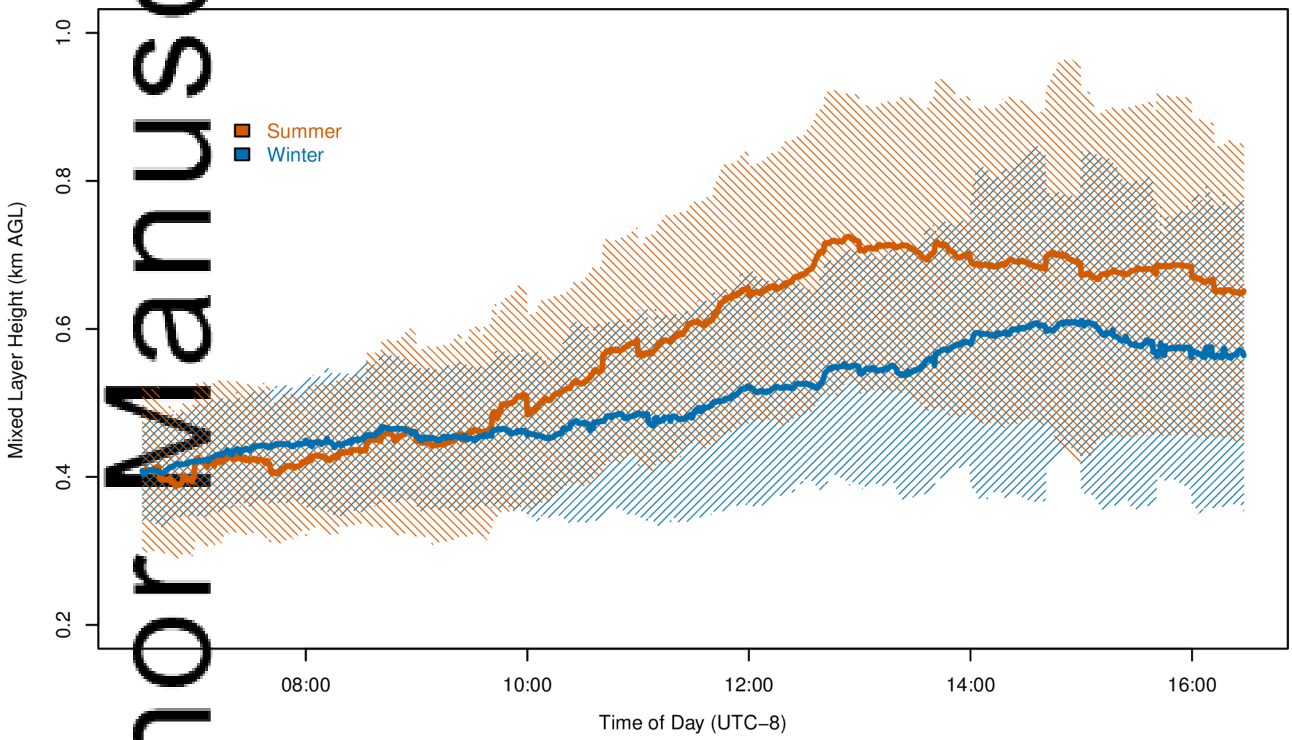
2016jd024953-f02-z-.eps



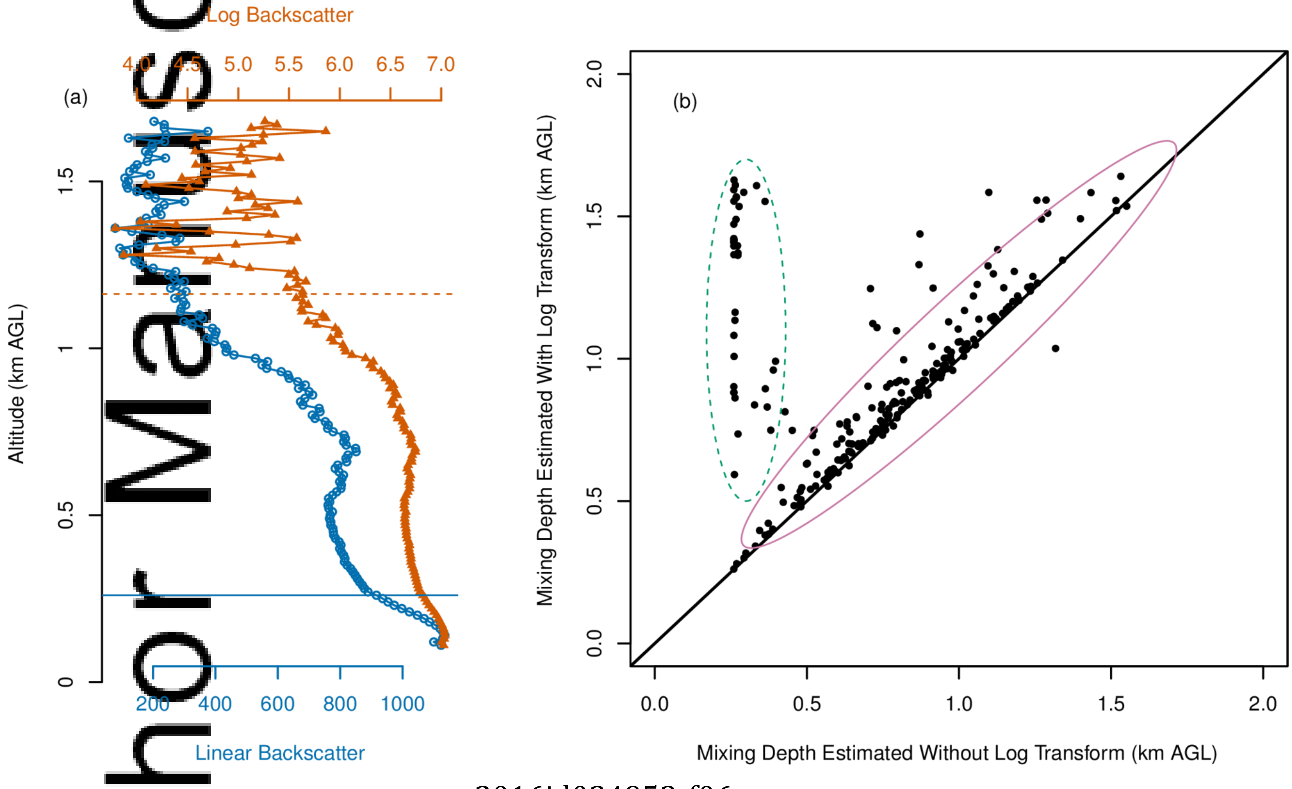
2016jd024953-f03-z-.eps



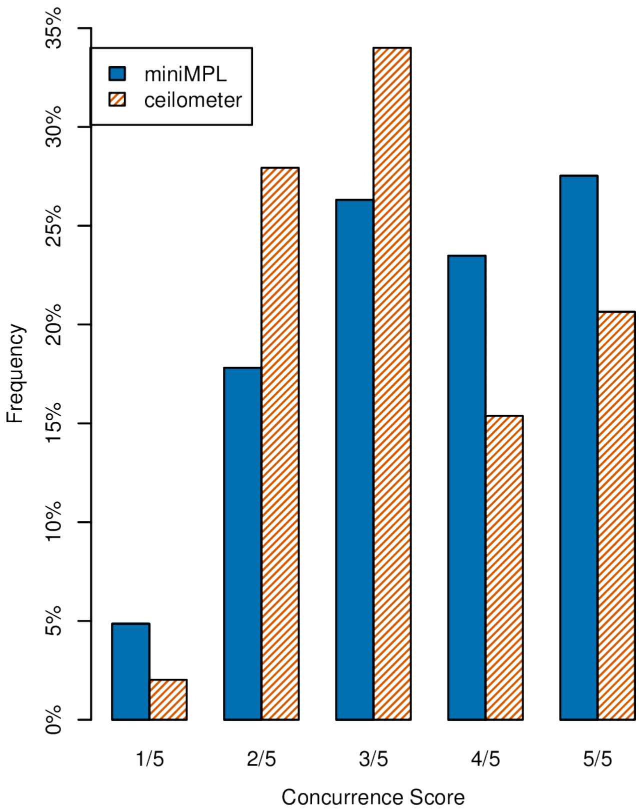
2016jd024953-f04-z-.eps



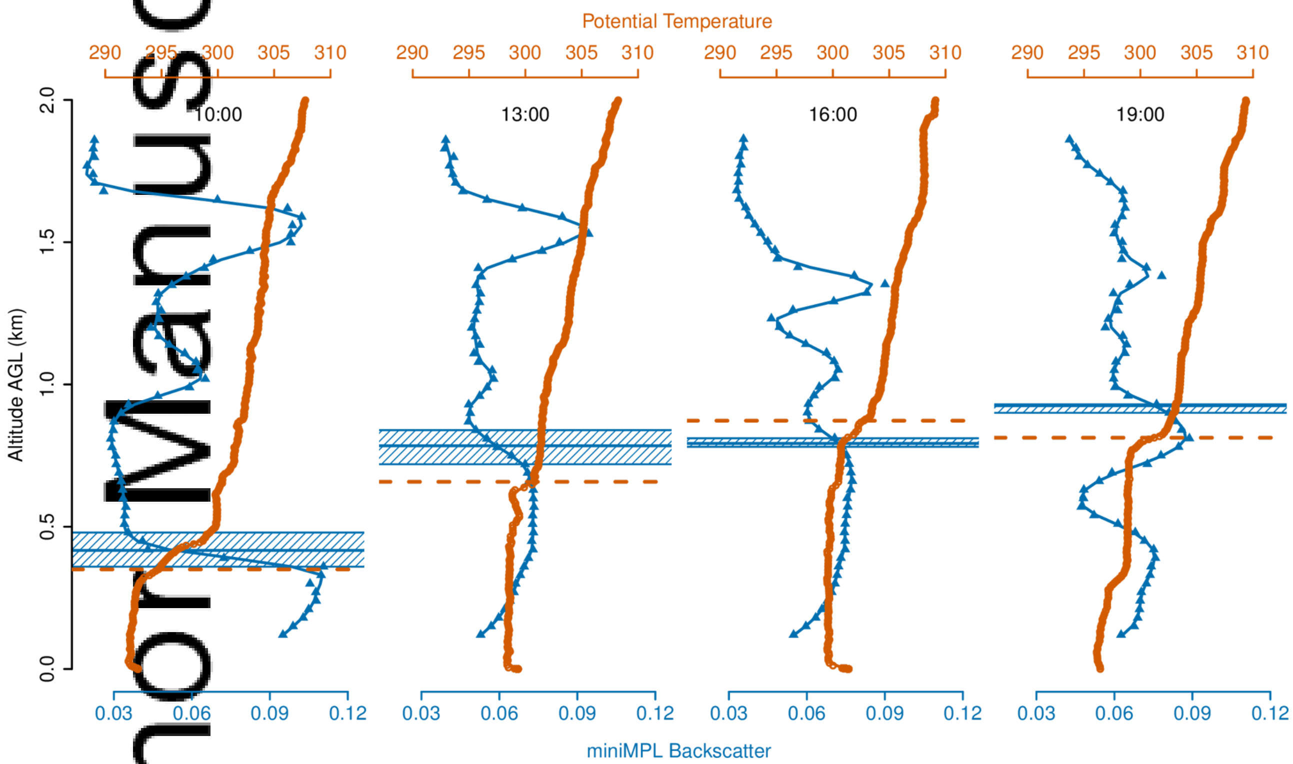
2016jd024953-f05-z-.eps



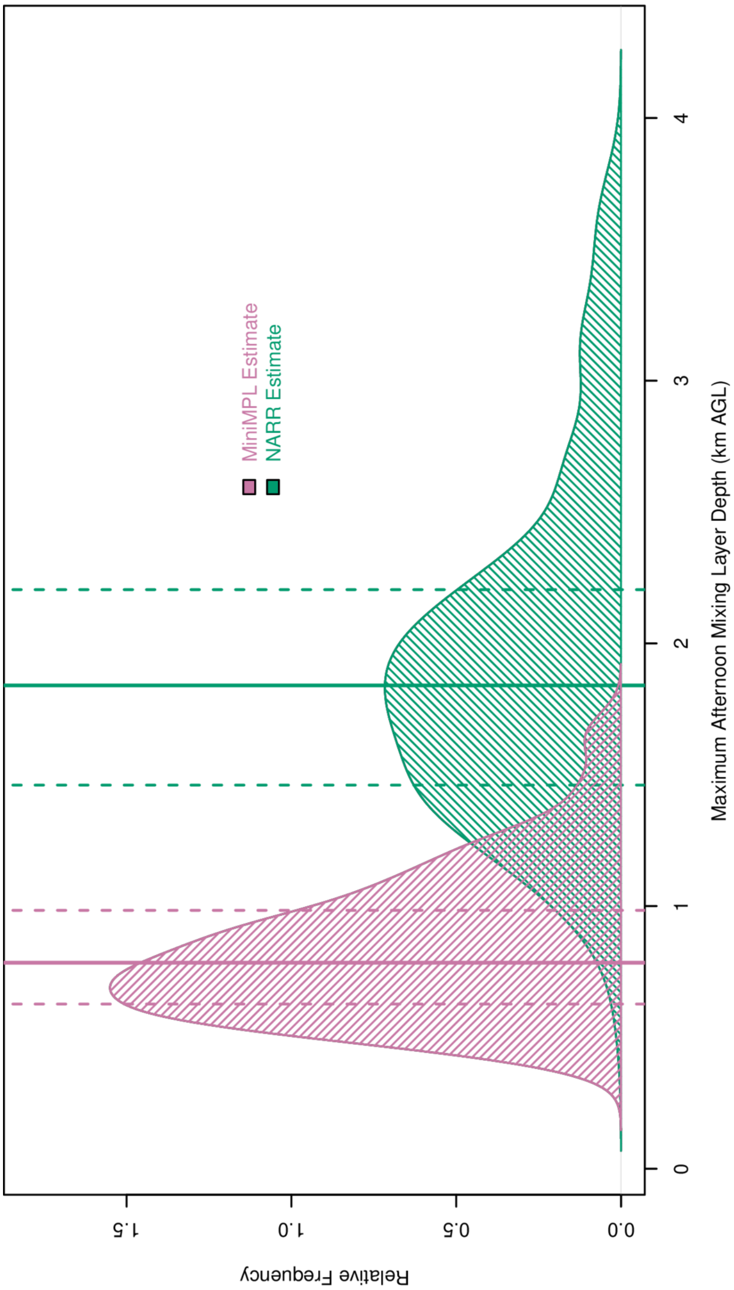
2016jd024953-f06-z-.eps



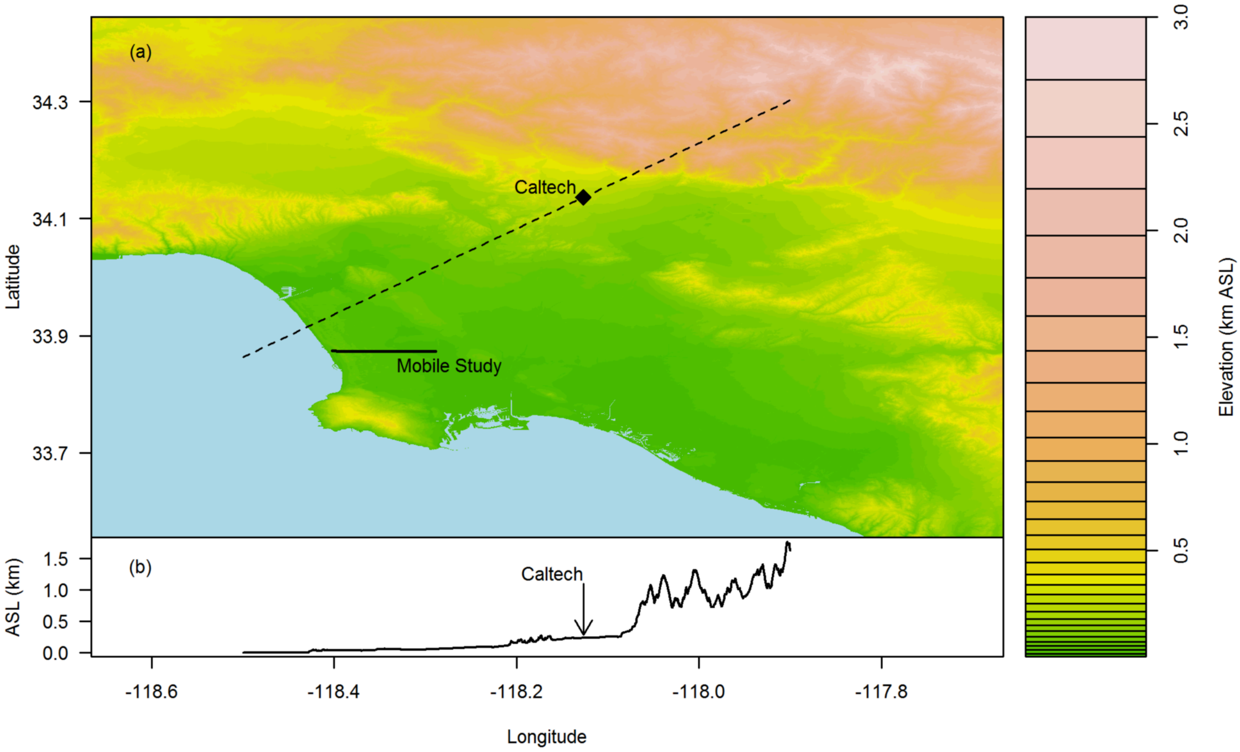
2016jd024953-f07-z-.eps



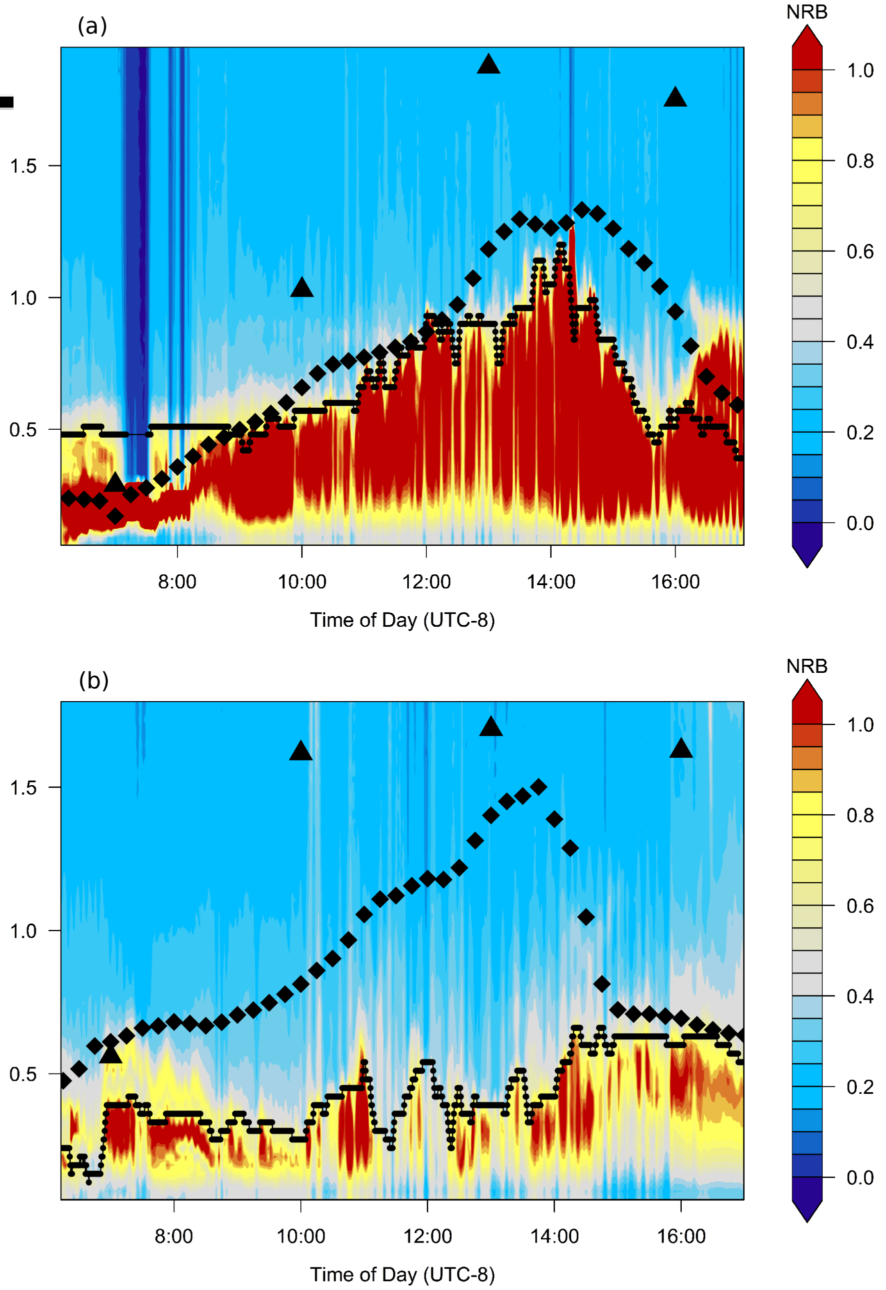
2016jd024953-f08-z-.eps



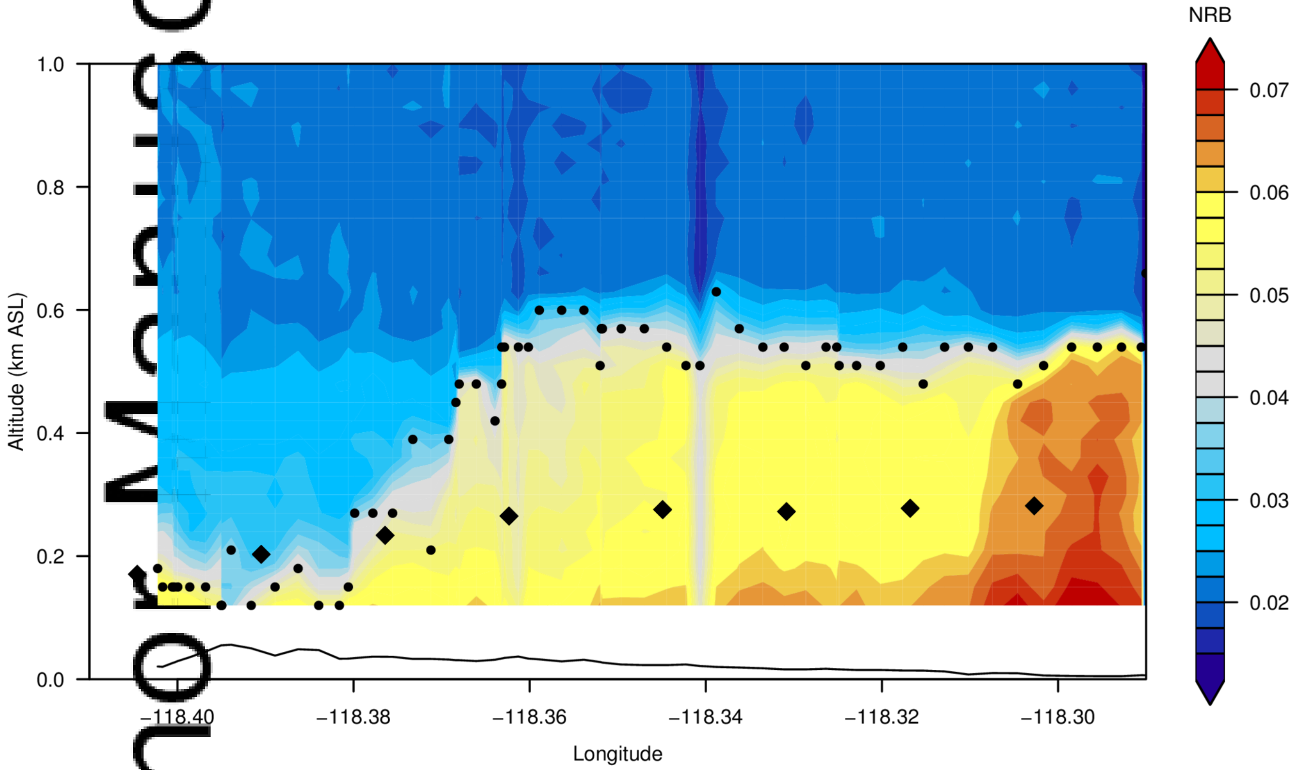
2016jd024953-f09-z-.eps



2016jd024953-f10-z-.eps



2016jd024953-f11-z-.eps



2016jd024953-f12-z-.eps