Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2016-340-AC1, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Comparison of aerosol LIDAR retrieval methods for boundary layer height detection using ceilometer backscatter data" *by* Vanessa Caicedo et al.

Vanessa Caicedo et al.

caicedo.vanessa@gmail.com

Received and published: 10 January 2017

We thank Referee 1 for carefully reading our manuscript and for the suggestions for revising and improving our work. Below we provide the Referee's review (in bold) followed by our response to individual comments. For reference and help to find the modifications we have made, we appended a revised version of the manuscript to our responses.

The wavelet-algorithm, as implemented in this study, seems not complex or robust enough to be able to deal with cases where the strongest gradient in the aerosol backscatter profile does not correspond to the BLH. Also, no climatology of retrievals is presented in this study, such that it is difficult to prognose a

generalization of the performance of the algorithm. We have made a small addition to show the performance of all methods over a clear day. This has been added as Figure 8 -a time series of a day in which the performance of all algorithms is seen and can be compared to results of two radiosonde. All algorithms are able to capture well the NSL, the growth of the BL and the peak BLH with the cluster method showing the most variability due to the detection of lofted aerosol layer signals incorrectly identified as the BLH. The aerosol gradient method and the wavelet method BLHs are very comparable after the manual selection of the aerosol gradient method BLHs.

The authors claim that retrievals are very difficult in case of clouds. I am unsure about it being very difficult, and would be interested to see in their response a Figure showing the performance of the algorithms on the cloudy days, which represent more than 40% of the dataset. A comparison study, not with the same but with other algorithms that I judge being more robust, using the CL31 ceilometer as well, has been published in 2012 (Haeffelin et al.). We have also added Figure 13, which shows the performance of all algorithms with the presence of cloud signals. This figure shows the decrease in correlation due to cloud signals on the overall performance for all algorithms. These results are discussed in Section 4.4.

The Haeffeling et al. (2012) study does an excellent comparison of five BLH algorithms. Most of the methods used in this study are algorithms, all which require careful pre and post manual selection or analysis of data. Our study wanted to arrive at the most automated yet reliable method to apply to long-term backscatter data such as the one available in the Houston site (2009-2015) and other similar data sets, which increasingly become available.

Also, in another study (de Haij et al., Proc., 2010) the wavelet method was applied on the older generation Vaisala ceilometer LD40. The de Haij et al. (2010) study uses a very promising threshold method to prevent the mistaken detection of the BLH due to lofted aerosol layers, the residual layer and so forth. However, this threshold or quality index is arbitrary chosen and it is independent of the absolute value of the aerosol backscatter in each profile used. The authors advise to use this quality index as a first step to check the reliability of the resulting BLHs. They also point out that the quality index has problems with multiple aerosol layers (lofted and/or residual) when combined with cloud layers where this quality index is suppressed and does not prevent the incorrect identification of the BLH, and is not reliable in profiles with low backscatter as the quality index is identified using a well pronounced BL under fully convective conditions, only. We believe that although this method works in some conditions, it does not systematically tackle the issue.

p.1, l.2 be more specific: "Over 40 cloud-free daytime radiosonde profiles" Since the three methods have different number of radiosonde profiles used, we do not specify this in the Introduction, instead this is explained in Results Section 4 along with the reasons for the number of radiosonde profiles used.

p.1, l.6 replace: "The Haar Wavelet method demonstrated to be the most robust" This has been replaced in p.1, l.6.

p.1, l.7-8 shorten the sentence, i.e. replace "only showing limitations (22.5% of all observations) due to the basic assumptions used to derive BLH from aerosol backscatter concentrations rather than errors with the algorithm itself." with "only showing limitations in 22.5% of all observations." This has been edited as suggested in p.1, l.6-7.

p.1, l.8 "Disagreement between thermodynamically and aerosol derived boundary layer heights" You claim "Overall good agreement" in p.1, l.3 but disagreement here. Replace with: "Occasional differences between thermodynamically and aerosol derived boundary layer heights" This has been replaced as suggested in p.1, l.7-8.

p.2, l.12 ", facilitating the monitoring of the nocturnal stable layer, internal aerosol layers and the nighttime residual layer" What about the ML? Here, we wanted to emphasize the ability of ceilometer to measure the BLH not limited to the ML but further measuring the NSL, internal aerosol layer and the residual layer which are seldom continuously monitored and studied. We have clarified this statement in p.2, l.11-14.

p.2, l.14 replace: "in the need for determining the most reliable and accurate method" This has been replaced as suggested in p.2, l.14-15.

p.2, l. 25-26 "Over 80 radiosonde launches and CL31 backscatter profiles beginning in January 2011 through March2015 are used." Why not say 85 (see Sect. 2.2) directly? This has been replaced as suggested in p.2, l.27.

Also, this gives me a launch every 19 days during this time period. Please me more specific (regular launches (i.e. every 19 days)? concentrated on some periods, some seasons?) We have added a figure (now Figure 1) to address this question and show the seasonal distribution as well as the time slots of the launches. Further explanation in p.2, l.27 - p.3, l.1 and also in Section 2.2 p.3, l.31 – p.4, l.2 has also been added.

p.2, l.27 replace: "The effect of cloud signals is analyzed separately for each method" This has been replaced as suggested in p.3, l.2.

p.2, l.27-28 "this data set includes data from the NASA DISCOVERAQ Texas campaign" What kind of data? Radiosonde measurements? Please specify. This has been corrected to specify ceilometer and radiosonde data from the NASA DISCOVER-AQ Texas campaign p.3, l.3-4.

p.3, l.4 Please be more precise: "The backscatter coefficient β or the backscattering cross section per unit volume of air is related to the received power with the following formula:" I prefer the formulation "is related to the received power" instead of "can be calculated from", because calculating β from eq (1) is not straightforward. The CL31 gives you a (scaled) attenuated backscatter coefficient (i.e. C Vaisala β $\tau 2$) in $10-9$ m -1sr-1 units, with C Vaisala around 1. We have replaced as suggested in p.3, l.11-12.

I accept that you call this aerosol backscatter or simply backscatter throughout the paper, as long as you remain consistent. Side comment: the liquid cloud calibration method of O'Connor et al. (2004) is a method that is used to calculate C Vaisala. We have made this consistent throughout the paper by always using "aerosol backscatter".

p.3, l.6 I am missing the background term (i.e. P(x,λ) = . . . + B) in eq. (1) and in its description in lines 7-10. Please add. This has been added in p.3, l.13.

p.3, l.15 Please specify which version of the firmware of the CL31 you use and what noise_h2 setting do you use. With the noise_h2 setting turned off, the CL31 backscatter signal is proportional to P(x)*2400^2 instead of P(x)*x^2 for all ranges>=2400m (in clear sky), because it is assumed that all data further consists of noise. See Kotthaus et al. (2016), Section 3.2 in particular. This could have an influence on the retrieval of BLHs higher than 2400m. We use the firmware version 1.7 with noise_h2 setting on. We have added this information on p.3, l.23.

p.3, l.19 "Radiosonde had a response time of 1s."âĂÍRemove, because redundant with p.3, l.20. This has been removed as suggested in p.3, l.27.

p.3, l.22 replace: "and a response time of 2s." This has been edited as suggested in .3, l.29.

p.3, l.25 "with most launches happening around 13:00 CST" Please be more specific. Add how many launches happened e.g. between 12:30 and 13:30 (or some other period that you define around 13:00 CST), and how many after that midday period. This has been clarified in p.3, l.32 - p.4, l.2 and new Figure 1 shows the hourly time slots of radiosonde launch times.

p.3, l.28 remove "therefore a constant concentration of backscatter" Indeed,

“concentration of backscatter” does not make sense. Also, the growth of aerosols due to swelling does not change their concentration, but changes the backscattered signal (i.e. constant concentration does not mean necessarily constant backscatter). This is a very good point. We have edited as suggested in p.4, l.4-5.

p.4, l.1-2 “within 10 minutes before or after radiosonde launch.” Please be more specific. Do you choose the closest in time in the 20min time-window? the mean aerosol-BLH in the 20min time-window? the closest in altitude in the 20min time-window? Or else? This has been explained in p.4, l.9.

p.4, l.5 replace: “an unstable boundary layer is identified by having a dry adiabatic lapse rate greater than “ This has been edited as suggested in p.4, l.12.

p.4, l.8 replace: “where the temperature profile” This has been edited as suggested in p.4, l.15.

p.4, l.9 replace: “where relative humidity or dew point temperature profiles sharply decrease as seen in the skew-T log-P diagram in Fig. 1b” This has been edited as suggested in p.4, l.16.

p.4, l.11 be more specific: “when comparing ceilometer and radiosonde derived BLHs (both manually) using “ Did you retrieve the BLH from the skew-T log-P diagram manually for each day? Yes, BLHs were retrieved manually. This has been edited as suggested in p.4, l.19.

Haman et al. (2012) found with this method better correlation coefficients (0.96/0.91 in unstable/stable conditions) than in this study, but it seems to be the same instrument. They use the Vaisala v3.5 algorithm, whereas you use the v1.3 algorithm. This was a mistake (BL-Matlab Control v1.3 was mistakenly referred to). Version 3.7 was used for this study. This has been edited in p.4, l.21.

Also, they use some “quality check” criterions (minimum gradient strength, rela-

tive backscatter change around the gradient of 15 %, minimum gradient height). Did you use these quality checks as well (especially this relative backscatter change around the gradient of 15 %)? Yes we use the same 15% sensitivity option and a 30m minimum gradient height. These settings have been added in p.4, l.29-30.

Why using here what is seems like an earlier version of the algorithm? Do you think these aspects explain the better correlation in Haman et al. (2012)? We are sorry for the confusion and have corrected to list the correct algorithm version p.4, l.21. **Or is it due to aging of the optics?** This cannot be ruled out. Haman et al. (2012) used data from 2009 – 2011 and in this study we continued the use of the same instrument for our data set (2011-2015).

Or is it due to the fact that they had measurements mostly during spring and summer, with supposedly stronger convective activity Please comment on this quantitative difference in the paper, in Sect. 4.1 for example. I actually do not expect a correlation coefficient of much more than 0.9 when comparing aerosol gradient BLH with radiosonde BLH, especially if the ML is not fully developed, since aerosol gradients are only a proxy for the thermodynamic BLH, and are not expected to correlate perfectly with it. As can be seen in Figure 1our radiosondes were mainly launched in May, June, September, and October, which also have strong convective activity. We believe the difference in correlation could be due to the manual analysis used by Haman et al. (2012) since they do not report a BLH "if the height of the [BL] is not clear" (p.705 in Haman et al. (2012)) rather than always reporting a gradient found by this algorithm (as long as the algorithm is able to calculate a gradient) as we do in this study. Haman et al. (2013) do not specify the selection criteria, therefore we cannot expand on the specifics and the effect the manual selection might be creating. An addition was made in Section 4.1 p.8, l.25-27.

p.4, l.18 replace: "The temperature correction of -10 is an algorithm setting that adjusts the shape" This has been edited as suggested in p.4, l.25-26.

p.4, l.24 "Therefore, a manual analysis of the algorithm's three resulting layers (Fig. 2) is required in order to prevent the incorrect identification of other aerosol layers". Did you do this manual analysis in this study? Please specify, since not entirely clear here. This has been re-written as "Therefore, a manual analysis of the algorithm's three resulting layers (Fig. 3) is required in order to prevent the incorrect identification of other aerosol layers. The algorithm gives three maximum negative gradients every 1-minute of which one is manually chosen as the BLH. These are then averaged to 10 minutes for radiosonde comparison" in p.5, l.1-3.

p.4, l.25 "The algorithm gives three maximum negative gradients every 1-minute, these are averaged to 10 minutes for radiosonde comparison". How are they averaged? Layer by layer? Or by clustering all points inside the 10 minutes into 3 clusters/layers, and then taking the mean of these clusters? Or else? We have clarified this in p.5, l.3-4. The Vaisala BL Algorithm outputs three calculated gradients for each 1minute output as Layer1, Layer2 and Layer3, for instance. However, Layer1, Layer2 and Layer3 are not necessarily a continuous measurement of the same gradient. A gradient at height x found in Layer1 can be found as a gradient in Layer2 in the following 1-minute output. Averaging layer by layer may include averaging from multiple gradients unrelated to each other and averaging all layers by time could also average multiple gradients. Choosing the BLH manually before averaging prevents this happening and leaves only one output for each one 1-minute estimate. These are then averaged to 10 minutes.

p.4, l.30-31 be more specific "The BLH is typically identified as the (temporal) variance local maximum" This has been edited as suggested in p.5, l.9.

p.5, l.8 replace: "the greater the EZ depth the greater the overestimation of the BLH" This has been edited as suggested in p.5, l.17.

p.5, l.9 punctuation, missing point at the end: "whereas aerosol gradient methods can give multiple results." Comment: the cluster method gives a unique

BLH only if you choose the number of clusters to be k=2. This has been edited as suggested in p.5, l.18.

p.5, l.14 replace "Due to increasing noise in backscatter profiles with height" with "Because the range correction needed to invert Eq. 1 increases noise in backscatter profiles with height," This has been edited as suggested in p.5, l.23-24.

p.5, l.17 be more specific: "The moving time average" This has been edited as suggested in p.5, l.27.

p.5, l.21 In Eq. (2), the power 2 should be applied on the brackets, not on the averaged profile. Also, you already used P as the received power in Eq. (1). Use an another letter for this variable. This has been edited as suggested in p.5, l.31.

p.5, l.22 try to add space and correct: "where P(z,spaceti) is the averaged LIDAR backscattered signal at time ti and height z, and P is the averaged profile from" This has been edited as suggested in p.6, l.1.

p.5, l.22-23 "from N number of profiles"âĂÍHow much is N here? The number of profiles corresponding to 10min? Please specify. This is correct. We have specified this in p.6, l.1-2.

p.5, l.24 "K-means is a data-partitioning algorithm that assigns observations" Explain that the observations are 3D-points where the first dimension is the range, the second the backscattered signal and the third the variance, and that these 3D-points are standardized (Toledo et al. (2014)). This has been specified in p.6, l.4.

p.5, l.30 punctuation, missing point at the end: "Step 4. Compute the average of the observations in each cluster to obtain new centroid locations." This has been edited as suggested in p.6, l.10.

p.6, l.3 replace "Cluster analysis will typically divide a well-mixed boundary layer into two clusters, one below a peak in variance corresponding the center of the

EZ," with "By choosing k=2, cluster analysis will typically divide a well-mixed boundary layer into two clusters, one below a peak in variance corresponding to the center of the EZ," Indeed, the cluster method gives two clusters only if you choose the number of clusters to be k=2. This has been edited as suggested in p.5, l.15-16.

p.6, l.5 "will cause the cluster analysis to assign clusters using other criteria" I do not understand what you mean with "using other criteria". Replace with "will cause the cluster analysis to assign clusters somewhere else" We have clarified with "however profiles with increasing noise and/or lofted aerosol layers will cause the cluster analysis to assign clusters elsewhere (for detailed description of criteria see Results Section 4)" p.6, l.16-17.

p.6, l.13 replace: "top of the BL by calculating the wavelet transform" This has been edited as suggested in p.6, l. 26.

p.6, l.20 be more specific: "Defining the dilation factor a and the range of centers b of the Haar wavelet function" This has been edited as suggested in p.7, l.3.

p.6, l.21-22 "b ranges from the lowest ceilometer recorded backscatter altitude of 10m to a maximum BLH of 2800m" How do you treat cases where the dilation is more than two times the altitude? Do you use a smaller dilation? Do you append zeros to your signal at "negative altitudes"? Please specify. We have edited this section as follows in order to further explain and specify the use of smaller dilations in p.6, l.29 – p.7, l.7-20.

p.6, l.22 "This limit was set as no radiosonde derived BLHs were found above 2800m" Comment: Here you could also justify this "upper climatological value" from the study made by Haman et al. (2012). It is better to avoid, if possible, setting algorithm parameters based on the reference data you compare with. This is an excellent point. We have used Haman et al. (2012) and Rappenglück et al. (2008) here to also justify this upper limit p.7, l.5-6.

p.6, l.26 put in the plural form: "at the heights of the biggest aerosol gradients in the backscatter profile" Indeed, there may be more than on big aerosol gradient (e.g. top of BL and top of lofted aerosol layer). This has been edited as suggested in p.7, l.10.

p.6, l. 24-26 change the order of the 2 sentences, i.e. change: "Lower dilation values create numerous CWTC local minimums (Fig. 4b) at heights of smaller aerosol gradients in the measured profiles. Larger values create large local minimums (Fig. 4c and 4d) at the height of the biggest aerosol gradients in the backscatter profile (Fig. 4a)." with "Larger values create large local minimums (Fig. 4c and 4d) at the heights of the biggest aerosol gradients in the backscatter profile (Fig. 4a). Lower dilation values create numerous CWTC local minimums (Fig. 4b) at heights of also smaller aerosol gradients in the measured profiles." The reason is that at the location of a strong negative gradient, the CWTC with a small dilation factor will also have a (strong) local minimum there. This has been edited as suggested along with the corresponding figure p.7, l.8-11.

This is why, in some studies, for robustness reasons, the local minima over the averaged CWTC profile (averaged over multiple dilations, say here 30m to 300m) are searched for. Indeed, the gradient you seek at the top of the BL, under the assumptions you stated in the beginning of Section 3, has a peak in the Wavelet transform at multiple dilations. See for e.g. Cohn and Angevine (2000) Fig. 2 b). As you already mentioned, going to high dilations however does not apply when seeking the top of the NSL, where only small dilations should be used. We use the mean of the averaged CWT coefficients such as Cohn and Angevine (2000), Compton et al. (2013), Scarino et al. (2014) as well, and have improved the description of the method in Section 3.4. p.7, l.7-20.

p.6, l.29-30 be more specific, i.e. replace "A higher value of 300m (Fig. 4d) for the dilution factor a is applied for daytime BLHs to identify the sharp transition between ML and FT" with "A higher value of 300m (Fig. 4d) for the dilution factor

a is applied for daytime BLHs and the strongest CWTC local minimum is used to identify the sharp transition between ML and FT" This has been reworded as suggested and replaced in p.7, l.14-15.

p.7, l.7 replace "t test" with "t-test" This has been edited as suggested in p.7, l.25.

p.7, l.10 "where X is the mean of the aerosol BLH samples, μ is the radiosonde BLHs mean, S is the standard deviation of samples," This has been edited as suggested in p.7, l.28.

p.7, l.13 replace "p value" with "p-value" This has been edited as suggested in p.7, l.31.

p.7, l.13 replace "or 95% confidence" with "or 5% significance level" This has been edited as suggested in p.7, l.31.

p.7, l.20 replace: "following precipitation or during periods of high wind speeds." This has been edited as suggested in p.8, l.15.

p.8, l.1 "(not statistically significant; $p < 0.05$)." Not statistically significant means $p > 0.05$, please recheck this and all other statements you make about statistically significance. This has been edited as suggested in p.8, l.30.

p.8, l.5-6 replace: "when the algorithm did not find strong enough gradients in the backscatter profile" This has been edited as suggested in p.8, l.34.

p.8, l.8-9 reformulate, i.e. replace "This is due to the assumption in the methodology of using aerosol gradients to detect the BLHs and thermal parameters to detect radiosonde BLHs." with "This is due to the difference of assumptions in the methodologies, using aerosol gradients to detect the BLHs on one side and thermal parameters to detect radiosonde BLHs on the other side." This has been replaced as follows "This is due to the different assumptions in the methodologies when using aerosol gradients to detect LIDAR BLHs or thermal parameters to detect radiosonde BLHs" p.9, l.3-4.

p.8, l.12 What about the statistical significance for this method? An additional column in Table 2 has been created showing p-values for all methods.

p.8, l.13-14 "From the 45 comparisons performed, 13.3% showed the algorithm mistakenly finding maximum variance peaks not corresponding not corresponding to the BL but to noise or other aerosol layers" You claim later in this paragraph 16 cases with noise and 5 cases with lofted aerosol layers, which makes (16+5)/45=46.7% and not 13.3%. Please change accordingly. We have specified as to what the percentages correspond to and have clarified these in p.9, l.23-30. 16 cases (35.5%) where the cluster analysis divided a cluster by noise i.e. smaller variance in lower altitudes and higher variance in higher altitudes. The 13.3% corresponds to instances were a single clear peak in variance (such as that in the ideal case in Fig. 4) was calculated but this peak corresponded to either noise or lofted aerosol layers (5 lofted aerosol layer cases plus 1 case where noise created one long peak in variance).

p.8 l.19, replace "CL31 displays a significant increase in noise with increasing altitude." Note that all lidars in general, not only the CL31, display an increase in noise with height, due to the range-correction. This has been edited as suggested in p.9, l.28.

p.8, l.21-22 "These were not algorithm errors but instead due to the implicit assumptions in using aerosol backscatter for BLH detection (constant concentration of backscatter within the BL and a negative gradient in backscatter corresponds to the top of the BL)". Change "constant concentration of backscatter" by "constant backscatter" (see my comment on p.3, l.28) This has been edited as suggested in p.9, 1.23-33.

I am not satisfied with the formulation: "these were not algorithm errors but instead due to the implicit assumptions". I trust that you did not make a programming error. Still, these errors are due to limitations of the cluster algorithm, and not due to violations of your basic assumptions. Indeed, assume you have

constant backscatter within the BL and a negative gradient in backscatter at the top of the BL. Now suppose there is a lofted aerosol layer on top of the BL and that your measured backscatter profile is affected by noise. This does not violate your basic assumptions of constant backscatter in the BL and gradient at the top, but still perturbs the algorithm. This is correct, cluster analysis noise errors such as the example in Figure 9 do not violate our basic assumptions when we are dealing with uncertainties introduced by noise. These are errors from the "limitations of the cluster algorithm" (not programing errors), which we call "algorithm errors". Errors made by the algorithms which violate basic assumptions of aerosol backscatter algorithms such as BLHs identified from lofted aerosol layers or the residual layers, are considered separately since these assumption errors will occur across all aerosol backscatter algorithms. These would not be called algorithm errors but rather errors arising from our assumptions. Here we sought to differentiate and separate errors in this cluster analysis algorithm from those independent errors not seen in the other two algorithms (the noise errors).

The example of a noise affected lofted aerosol layer could certainly be an error due to the cluster analysis noise sensitivity or the basic assumptions we use. We identify the assumption or algorithm errors based on how the BLH was identified. Therefore, if the profile is affected by noise the identification of the BLH could happened in the separation of clusters into similar variance intensity as seen for 35.5% of cases referred to as "algorithm errors" or the BLH could be identified as the height of the lofted aerosol layer. The later BLH identification case would not be called an algorithm error but rather an error due to the assumptions used.

We have edited this sentence to clarify the distinction in p.9, l.23-33.

The cluster method works best to retrieve the top of the residual layer or of the fully developed mixed layer, assuming no clouds, high signal-to-noise ratio and no lofted aerosol layers above the BL. The noise aspect in particular, affecting the temporal variance, seems to have been an important issue here, and you

could for e.g. say that this issue should be kept in mind and tackled when using the clusteror other variance-based algorithms on the CL31 (even if you say that later in the conclusions, p.10, l.22). A similar statement has been added to p.12, l.26-27.

p.9, l.4 Add: "capturing the BLH. Also, it was less affected by noise than the gradient method or the cluster method." Remark: These conclusions are not surprising, since the wavelet method with 300m dilation means taking the difference of two 150m-averaged backscatter signals. With this, you increase the total amount of signal considered and decrease the variability. The drawback is that with a big dilation of 300m, you risk to detect the top of a lofted aerosol layer instead of the top of the BL, which happens 12.5% of the cases, as you mentioned. This has been edited as follows "This method was also less affected by noise than the gradient method or the cluster method" p.10, l.22. The selection of dilations is certainly a compromise and one that was shown to work for the majority of cases.

p.9, l.6 soften the appreciation: "The identification of the BLH is more difficult in the presence of clouds". Indeed, you say yourself in p.9, l.13-15 that in case of cloud the gradient and wavelet methods often compare well with the thermodynamic BLH. This has been edited as suggested in p.11,l.2.

p.9, l.6-7 What about precipitation or fog events? Are they included into this category of BLH retrieval with cloud signals? Please specify. Fog and precipitation events are not included in this category. We cannot include these conditions since the comparable radiosonde launches were not performed during fog or precipitation events. This has been specified in p.11, l.5.

p.9, l.26 "The presence of clouds limits the detection of the BLH for all methods due to either the extinction of aerosol backscatter signals above the cloud, or the presence of low clouds mistakenly identified as the BLH." In case of cumulus or stratocumulus, even if you have extinction above the cloud, the error you

make with your ceilometer BLH detection compared to the thermodynamic BLH is acceptable and not a major limitation. The presence of low clouds mistakenly identified as the BLH is similar to the cases where you have a lofted aerosol layer above the BL, in the sense that the algorithms get "distracted" by those (strong) gradients and or variances and fail to pick the correct gradient at the top of the BLH. This is an attribution problem (linked to the processing of the measurement), but not a problem of the ceilometer measurements. We have replaced this statement with "The presence of clouds creates difficulties in the detection of the BLH for all methods due to the extinction of aerosol backscatter signals above the cloud, the presence of low clouds mistakenly identified as the BLH, or the detection of high cloud signals above the skew-T log-P derived BLH" p.11, l.30. Indeed the cloud problem is similar to lofted aerosol layers where the algorithms are confused by these strong signals. However, the difficulty lies within the difference of low-level clouds, cumulus or stratocumulus clouds were a BL algorithm cannot differentiate between types of clouds, therefore is not able to accept (cumulus/stratocumulus cloud) or reject (low level clouds) BLHs calculated with cloud signals. Additionally, according to literature radiosonde BLH detection methods disagree or sometimes fail in some cloud conditions. For example, Hennemuth and Lammert (2006) show detection methods failing in non-fully convective conditions where the cloud top does not equal the inversion corresponding to the BLH. This case is seen in our data set as well (Figure 13) where radiosonde BLHs show a lower BLH than the aerosol backscatter methods which find the BLH either at the base or top of a high-level cloud. These situations make the BLH detection particularly difficult in the presence of cloud signals.

p.9, l.27-28 soften the appreciation: "Hence the removal of profiles with cloud signals is preferred for the automatic retrieval of the BLH". This has been edited as suggested in p.11, l.32-33.

The removal of profiles with cloud signals is per se not needed for the automatic retrieval of the BLH. Clouds introduce simply additional uncertainty. The

Printer-friendly version

Discussion paper

algorithms you use in this paper do not really tackle the attribution problem: retrieved BLHs are here simply defined as the strongest gradient in the profile, or the first gradient from the ground, or the first separation point between two clusters, etc. There are more robust algorithms that have been developed in the last few years and advances have been made in this attribution problem issue, see my comments below for the "conclusions" section. For this reason we feel cloud signals should be removed when using these specific algorithms in order to get the most reliable BLH detection. Backscatter data alone cannot tell us how to calculate the BLH when clouds are present. Figure 13 shows the effect of these cloud signals on the overall correlation. As can be seen here the correlation decreases for all methods, especially the cluster analysis method, which constantly finds the BLH at the cloud base level. This uncertainty also arrives quite often - as mentioned above - when BLHs are lower than that of the cloud layer in not fully developed convective conditions.

p.9, l.29 be more specific: "since the moving time averaging performed before the application" This has been edited as suggested in p.11, l.33 - p.12, l.1.

p.10, l.2-3 be more specific: "were compared to over 40 cloud-free daytime radiosonde profiles" In order to specify the number of radiosondes used we have added the following sentence "This comparison used 47 radiosondes for the aerosol gradient method, 45 for the cluster analysis method, and 48 for the Haar wavelet method due to limitations implicit to each algorithm (see Results Section 4)" p.12, l.6-8.

p.10, l.9 replace: "corresponding to the top of the BL by calculating the wavelet transform" This has been edited as suggested in p.12, l.13.

p.10, l.20 "Profiles were also mistakenly divided into relative variance concentrations rather than a peak in variance," I do not quite understand the term "relative variance concentrations". Replace with: "Profiles were also mistakenly divided due to the increasing noise with height rather than a peak in variance," This has been edited as suggested in p.12, l.24-25.

p.10, l.21-22 "Previous knowledge of the BL will aid in identifying these algorithm errors, but is not necessary to obtain a BLH with this method." Not very clear. Replace with: "This method being automatic, previous knowledge of the BL aids in identifying such algorithm errors, but is otherwise not necessary." This has been edited as follows "With this automated cluster analysis method, a previous knowledge of the BL aids in identifying such algorithm errors, but is otherwise not necessary" p.12, l.25-26.

p.10, l.24 "without previous knowledge of the BL required to derive at a BLH." Not very clear. Replace with: "without previous knowledge of the BL required, as this method is also automatic." This has been edited as suggested in p.13, l.2.

p.11, l.1 replace: "signal and a cluster for cloud-free signal", i.e. signal without ending s, because it is along one backscatter profile. This has been edited as suggested in p.13, l.12-13.

p.11, l.5 be more specific: "more sensitive due to the moving time averaging applied" This has been edited as suggested in p.13, l.16.

p.11, l.3-5 "Limited detection of the BLH in aerosol profiles with cloud signals is seen for all methods due to either the extinction of aerosol backscatter signals above the cloud, or the presence of low clouds mistakenly identified as the BLH" You did not really show this "limited detection of the BLH in aerosol profiles with cloud signals". (85-48)/85 = 43.5% of the measurements were not analyzed because of cloud presence, which is not negligible. I would be very interested to see a scatter plot exactly like the one in Figure 5, but for the 37 days with clouds. You said yourself in Section 4.4 that the methods tested here, especially the gradient and the wavelet methods, correlated quite well with the skew-T log-P derived BLH, so why actually exclude them from the analysis? This statement corresponds to the correlation using cloud-free BLHs (Figure 3). We have added Figure 13 to show the effect of clouds in the overall correlation if these were not removed from the analysis. As can be seen, correlation decreases for all methods when comparing to skew-T log-P derived BLHs. The gradient methods performs reasonably well in fully convective cloud-topped BL were the difference between the BLH by radiosondes is close to the aerosol derived gradients. However a fully convective cloud-topped BL is not always present and radiosondes will show strong inversions at a lower height than the cloud-top as the gradient methods find. In fully cloud topped BLs, the gradient methods find the BLH at similar heights though usually lightly higher than radiosonde BLHs. The cluster method however finds the cloud base as the BLH significantly underestimating the BLH. Correlation coefficients when using only cloud signals is poor for all methods. We calculated an r2 of 0.36, 0.14, and 0.33 for the aerosol gradient, cluster analysis, and wavelet methods respectively (not shown).

p.11, l.8-10 "The errors found with this method were due to the basic assumptions used to derive BLH from aerosol backscatter concentrations rather than errors with the algorithm itself." Again, as in p.8, l.21-22, I am not satisfied with this formulation. Replace with: "The errors found with this method were due to lofted aerosol layers, low-level clouds and differences in determining BLHs using aerosols and thermodynamically using radiosondes." This has been edited as suggested in p.13, l.22-24.

p.11, l.10-13 "For this reason, the use of this automated method is recommended for BLH observations with careful determination of the dilation values needed for individual instrument and locations. It is suggested to employ the wavelet method in future studies, in particular for long-term boundary layer evolution studies and spatial analysis of the BL using multiple LIDAR aerosol backscatter measurements." More results are needed in order to accept these two phrases, as I will explain below. Please reformulate: "In order to use this method on other instruments and locations, dilation values should be determined carefully and individually. Out of the three methods tested in this study, it is suggested to employ the wavelet method in future studies, in particular for long-term bound-

Printer-friendly version

Discussion paper

ary layer evolution studies and spatial analysis of the BL using multiple LIDAR aerosol backscatter measurements.“ This has been edited as suggested in p.13, l.24-27.

What do you mean by “evolution” in “boundary layer evolution studies”? Seasonal diurnal cycle? Annual cycle of the e.g. 13:00 CST value? Please specify. This had been indicated in p.13, l.26.

Here you did not show any “climatological results” (i.e. do we recover a diurnal and annual cycle with this automatic wavelet method and do they compare qualitatively well with the cycles depicted in the Haman et al. (2012) study?) This would be the expected second part of the validation of the algorithm (wavelet method) you preconize in your conclusions, now that you have good comparison results with the radiosonde measurements, and especially if you suggest this algorithm as suitable for PBL studies on multiple LIDARs. We now show the algorithms for a full day (October 24th, 2013) as a case study to demonstrate the continuous detection efficiency of these algorithms. This day was chosen as it is one of the few days with two radiosonde launches, which qualify for the cloud-free signal analysis at the same time. The result of all algorithms is shown in the added Figure 8 and discussion in Results Section 4 – 4.3.

I am not sure (at least I need to see it) that you are able to see the full diurnal cycle of the ML with the wavelet method as implemented here. For example, during the morning hours, by taking the strongest local minimum, you risk to still detect the top of the RL instead of the top of the developing CBL. But for the annual cycle of the CBL at say 13:00 CST (where you could roughly assume that the CBL is fully developed), I could imagine it works reasonably well. This is correct. The algorithms will most likely detect the top of the RL and for this reason we have added a simple height detection limit of 500m during nighttime hours continuing to two hours after sunrise for which resulting BLHs are shown in Figure 8 and detailed in p.8, l.5-9.

Printer-friendly version

Discussion paper

There are also other automatic methods than the wavelet method or the cluster method, or in a more general sense gradient or variance methods. I think in particular of the "fitting an idealized backscatter profile to observed profiles" method of Steyn et al. (1999), which is a publication you cited. We did not review the idealized backscatter method since we felt it was less reliable and flexible than the method we chose. Studies by Eresmaa et al. (2006), Munkel et al. (2007) and Munoz and Undurraga (2010), show the idealized method to work well under unstable conditions but less so during stable conditions where this method will often mistakenly choose the RL as the BLH. Munoz and Undurraga showed that the use of the idealized method had trouble with a non-idealized BL such as that in the study of the complex Santiago de Chile basin. We felt the idealized method to be less malleable to a stratified BL often found in nocturnal conditions.

In addition, you might be confronted to physical inconsistencies if you only take the largest peaks in gradient or variance (especially during the development of the CBL in the morning, you risk to jump back and forth between the still existing RL and the CBL). More robust algorithms aiming to tackle these issues exist (among others, gradient-variance coupling with temporal height tracking (Martucci et al. (2010), coupling with a BL-model (di Giuseppe et al. (2012), Biavati (dissert., 2014)), coupling with surface measurements (Pal et al. (2013)), coupling with graph theory (de Bruine et al. (AMTD, 2016))). Such a more robust method has not been investigated or commented here, and your suggestion to use the Haar-wavelet method (out of all existing methods) is thus not justified enough. Note that there is a review study of (Haeffelin et al., 2012), where several state-of-the-art algorithms at that time on three different ceilometers, including the CL31, are compared. Here you did not do a review of all state-of-the-art algorithms, but compared three rather simple methods (one manual, two automatic) with radiosonde measurements, and showed that the wavelet method with a large dilation is a robust method on the CL31 to derive the height of strong gradients in the aerosol backscatter profiles, and that using the temporal variance should be

done with precaution on this ceilometer. This is valuable information for future work. This is correct, we did not review all algorithms available for aerosol backscatter BL detection, but reviewed what we felt are the most used retrieval methods. The purpose of this study was to arrive at the most reliable and (hopefully) automated algorithm to use when dealing with extensive data sets as is the need when using long-term data sets such are those of the CL31. Although methods exist which attempt to tackle the known problems of residual layers, multiple aerosol layers, and clouds they still very much rely on the manual inspection of all data or extensive supplemental measurements. For example, such as the temporal-height-tracking method as that proposed by Martucci et al. (2010) requires manual identification of the BLH from the two local minima detected. The method proposed by de Haij (2006) requires almost a threshold value to prevent the detection of incorrect BLHs yet can fail in various conditions as this threshold is calculated from a pronounced convective BL. The method using surface measurements (Pal et al., 2013) show great correlation with radiosondes when using micrometeorological measurements to aid in the detection of the BLH using the three gradients calculated from the STRAT2D method. However most CL31 measurements sites are not coupled with extensive micrometeorological measurements. The use of a BL-model by di Giuseppe et al. (2012) shows great improvement, especially during nighttime hours when the RL signals are often incorrectly identified. Yet this can also be corrected as seen in Figure 8 by limiting the detection height limits of both wavelet and cluster methods. In contrast, the coupling with graph theory (de Bruine et al. 2016) method is shown to decrease errors coming from other aerosol layers and decreases the jumps in BLHs implicit for example, to STRAT2D calculation of multiple gradients. This method showed very promising results and which could potentially be applied to a long-term automatic detection of the BLH. Overall, the existing methods are aiding in the correct identification of the BLH. However they do not completely tackle the problem, again due to some limitations inherent to BLH detection based on aerosol backscatter methods. We feel that, although the methods reviewed in this study are simple, equally simple fixes, such as the height detection limit during nighttime hours and the removal of cloud signals tending to elude the algorithms, give reliable detection of the BLH, in a suitable simple automated manner. This is in particular helpful in air quality studies, which mainly occur in less cloudy conditions.

You also showed that there are cases where the aerosol layer based and thermodynamic based PBLHs do not match perfectly, which would be a further example giving input to the discussion on how well both methods compare. To my knowledge, it is also the first time the Cluster method, which is based on a good idea but seems quite sensitive to noise, is evaluated with the CL31 ceilometer, and you showed also that it was a potentially interesting technique to detect cloud layers. This was very interesting to us as well. For future studies we would like to look into this further. The detection of clouds layers with this method coupled with the detection of cloud-free BLH might give us a more robust data set for evaluating BL evolution such as seasonal and diurnal BL evolution and how this compares to the presence of cloud layers.

Figure 2:Remove the word "Density" in the z-label "Aerosol Backscatter Density" of the figure, because "backscatter density" does not make sense and it will be more consistent with the "Aerosol Backscatter" that you write in all other plots. Change the legend: "Aerosol backscatter time series for September 26, 2013." This has been edited as suggested to Figure 3. The image was also changed to keep consistency with Figure 8 of the day of October 24, 2013.

Please also note the supplement to this comment:
http://www.atmos-meas-tech-discuss.net/amt-2016-340/amt-2016-340-AC1-supplement.pdf

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-340, 2016.