Authors' response to reviewer #1

We thank reviewer # 1 for carefully reading our manuscript and the provision of many useful comments and detailed suggestions. We have considered all comments. They gave as useful hints where improvements of the paper were necessary to better understand our methodology and conclusions. Below all points raised by the reviewer are repeated; our comments are added in italics.

The changes (revised version vs. AMTD-paper) are highlighted as displayed by latexdiff ("diff.pdf", maybe renamed when uploaded as a supplement). For the sake of clarity only small changes are explicitly mentioned in our point by point replies, otherwise we refer to the corresponding parts of "diff.pdf" (in blue). Note, that some of our responses interact with comments of the other reviewers, so sometimes it is difficult to refer a change to one specific reviewer's comment.

Point by point replies

General comment

[...]

While these results are plausible and well explained, the paper currently suffers from a few serious deficiencies. First, the investigation comprises only two summer months, and day-night differences in the relationship MLH and pollutants are not at all considered. Second, the whole area of Berlin is represented by only one ceilometer. It is however well known that the mixing height will show some variation over such a large area both day and night, depending on the degree of urbanization and other surface related influences. A discussion on this issue is needed. The short investigation period and the use of only one ceilometer are currently briefly discussed in the conclusions, promising to tackle these issues in the future. However, these shortcomings have to be discussed more deeply, including references. An investigation of day-night differences has to be included in a revised version. Some of the figures need improvement (see below).

→ *The main concern of the reviewer is the limited length of the observation period, the number of ceilometers and the missing discussion of day-night differences (a similar comment was given by reviewer #2).*

The major focus of the BAERLIN2014 project was on ozone, secondary organic aerosol and the effect of urban vegetation. All of these effects are found at its maximum at summer especially at highest temperatures and oxidation strength. Because of the limited amount of resources the campaign must be concentrated on three months (June, July and August; this remark as a background information).

In principle we agree with the reviewer that more ceilometers would have been beneficial for the study. As already mentioned research projects are always limited with respect to money, personnel and hardware (temporal extent and spatial coverage of measurements, number of measured atmospheric variables, number of instruments, etc.). In case of BAERLIN2014 e.g. no external funding was available. As a consequence field campaigns always are limited in time: this was also true for the BERLIOZ campaign mentioned by reviewer #3.

Nevertheless we believe that BAERLIN2014 provided very valuable scientific results even if there was only one ceilometer available. We were able to demonstrate to what extent differences in MLH-retrievals play a role for calculating correlations between MLH and air quality parameters. By addressing standard retrievals (the proprietary software of the ceilometer manufacturer) and air quality measurements from an official monitoring network we think that the conclusions are relevant. These investigations could only be performed in the framework of a dedicated campaign because ceilometers do not yet belong to the standard equipment of urban air quality networks. To our knowledge only in Paris a network including three ceilometers for routine observations was recently established: the collaborative measurement platform "OCAPI". Results are not yet published. See extension of the introduction (page 3, lines 24 ff of diff.pdf):

Prospectively also the implementation of urban networks for air quality studies is likely at least for selected cities occasionally suffering from pollution events – recently three ceilometers were set up in larger Paris for this purpose (OCAPI: Observation de la Composition Atmosphérique Parisienne de l'IPSL).

Based on our research, open questions could be identified, one of them being the need for an in-depth investigation of the behavior of the mixing layer over a large municipality. So we hope that in future the wishes of the reviewer (and ours) to have more ceilometers and at least one full annual cycle of the MLH can be fulfilled, and that our paper will be *a motivation for setting up the corresponding infrastructure (see also our replies to the detailed comments of reviewer #1 below). In the conclusions (center of page 31 of diff.pdf) we have also stressed that numerical models (mesoscale, microscale) are required as well.*

As a consequence we have added several sentences, in particular we have clearly describe the motivation of our study to avoid misunderstandings (see introduction, page 3, line 6 ff of diff.pdf).

Following the suggestions of reviewers #1 and #2 we have extended Sect. 5.1 by discussing day-night differences and the influence of the wind field (reviewer #3). In addition a short comment on differences between working days and weekends has been added (see pages 24–26 of diff.pdf).

Reviewer #1 was primarily interested in day-night differences: The resulting correlation coefficients R of hourly values of MLH and PM_{10} *for all sites with* PM_{10} *measurements are shown in Figs. 1 and 2 (next page). Figure 1 covers the time period between 07:01 CET and 20:00 CET ("day time"), whereas Fig. 2 is for measurements before 07:00 CET and after 21:00 CET ("night time"). The four MLH retrievals are color-coded according to the legend. It can be seen that the absolute values of R are small as already stated in the original manuscript (for measurements of the whole day we have a range* $-0.3 < R < 0.1$*). For day time measurements (Fig. 1) we get* $-0.33 < R < 0.10$ *, for night time the correlation is slightly different (*$-0.27 < R < -0.09$*). The main difference is that during day time there are three out of 11 stations with positive correlations and 8 sites with* $\|R\| < 0.1$*, whereas during night time* $R < 0$ *for all sites and only one site with* $\|R\| < 0.1$*. These values are plausible as under ideal conditions an anti-correlation between MLH and* PM_{10} *is expected in view of the suppressed vertical mixing when the mixing layer is very shallow during night. Note, that during night the number of point source decreases, in particular at the outskirts, so that we find the lowest absolute values there. Nevertheless the small absolute values of R suggest that the MLH is not the only influencing factor. This is in accordance with several comments of all reviewers (and several statements in our manuscript) that there is no "simple" link between* PM_{10} *and MLH, on the contrary many processes are relevant.*

To avoid a substantial increase of the number of figures in the paper we have summarized the results of the above mentioned issues as a table

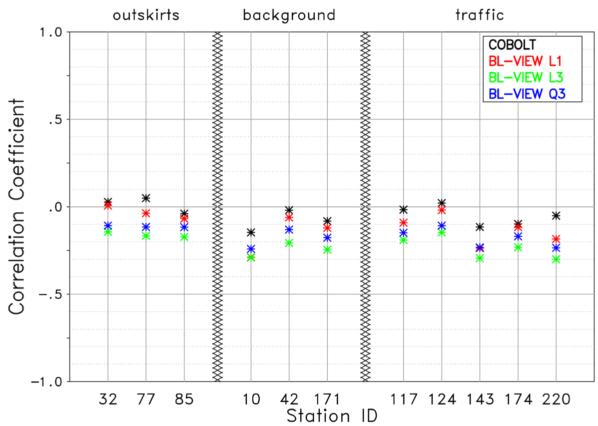

Figure 1: Correlation coefficient R of hourly values of MLH and PM_{10} shown for the 11 sites as indicated (station ID according to Table 1). The four MLH retrievals are color-coded according to the legend. Only measurements between 07:01 CET and 20:00 CET are considered.

and a new paragraph (see page 26 of diff.pdf and the new Table 3).

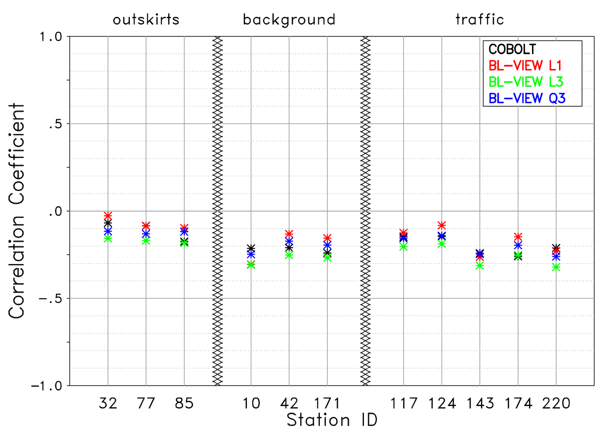

Figure 2: Same as Fig. 1, but only measurements before 07:00 CET and after 21:00 CET are considered. NOTE: THESE FIGURES ARE FOR DEMONSTRATION ONLY, NOT INCLUDED IN THE MANUSCRIPT.

Specific comments

- p. 3, line 1: The statement is too optimistic (frequently used approach); if true, provide more references. I think determining MLH from ceilometers in a reliable manner is quite a new subject.

  → *We have added more references: Haman et al. (2012), Caicedo et al. (2017) (see also response to reviewer # 3), de Bruine et al. (2017). A paper by Knepp et al. has been submitted to AMTD on 30. May 2017, and cannot be cited yet. More citations can be found in the already cited papers. Moreover we have added a short remark on "ceilometers" and "ALC" (automated lidars and ceilometers), see introduction, page 3 of diff.pdf.*

  In the last years several hundreds of ceilometers and ALC have been set up – not only by weather services but also by universities and research institutes. This triggered several activities to develop MLH-retrievals and improve their reliability – so it can be considered as a rather new subject. Most of the users rely on "atmospheric products" (primarily cloud bottom heights and mixing layer heights) that are automatically provided by the proprietary software, even if it is sort of a black box. Thus, from our point of view it makes sense to discuss associated problems, e.g., the risk to over-interpret the ceilometers' output (see our extended introduction, page 3 of diff.pdf). On the long term perspective it is likely that ceilometers will be the standard instrument to automatically monitor the aerosol distribution (this was the motivation of most weather services to establish these networks). We are sure that these instruments have a high potential if data are correctly exploited, as a consequence this subject will become even more relevant. Some future applications are briefly mentioned in the conclusions.

- p. 3, line 20: The short investigation period is mentioned here for the first time (see General comments).

  → *The duration of the campaign is mentioned in the abstract, and in the introduction after a general discussion of the topic (p. 3, line 20). We felt that this was a logical structure, but it can also be moved to the end of page 2 (see page 3, first lines of diff.pdf).*

- p. 4, line 7: The main shortcoming of sodar and RASS is that they usually cannot provide the whole diurnal cycle of MLH in Central Europe, especially in summer. A sodar alone can give a reliable estimate of MLH only with careful data analysis, see e.g. Bound.-Layer Meterorol. 124, 3-24 (2007).

  → *We agree and added a corresponding comment to the manuscript including the suggested reference (Piringer et al, 2007, page 4 line 27 of diff.pdf). In Seibert et al. (2000) and Emeis et al. (2012) – already cited – this topic is also discussed.*

- p. 4, lines 17-18: The advantage of spatial coverage of a network of ceilometers is not used in this study.

  → *We regret that our statement could be misunderstood. We mentioned "networks" here to highlight that there is an infrastructure of active remote sensing instruments that is very dense compared to research lidar networks (e.g. compared to the spatial and temporal coverage of EARLINET). Urban ceilometer networks are not known to us – at the EGU 2017 one of the authors (MW) learned that recently a few new instruments have been set up so that small scale investigations might be possible in future for selected cities, provided that it is known to potential users. Currently the implementation (mainly for scientific or educational purposes at universities) is not coordinated. The only exception seems to be Paris (OCAPI). We have added a corresponding note (see our reply to the general comment).*

- p. 8, line 5: please elaborate statement (one ceilometer is representative for a metropolitan area; see also General comments).

  → *We have written "is assumed to be representative", not "is representative". With only one ceilometer (see also our other replies) we cannot prove that such a strong statement is true for Berlin. So we refer to a previous case study in Munich (see the citation), and investigations of the diurnal cycles of the MLH in Munich, Freising and Augsburg, which are almost identical (Geiß, 2016). The distance between these stations however is somewhat larger than the size of Berlin (approximately 50 km). Similar findings were published by Lotteraner and Piringer (2016): they compared* *mean diurnal cycles for Vienna and Obersiebenbrunn, a village 26 km east of Vienna.*

Moreover we know of a short case study in the greater area of Paris including a lidar site in the city center (Jussieu), one site with a lidar and a ceilometer at SIRTA (outskirts) and one lidar site 105 km south of Paris, supplemented by mobile lidar measurements from a van. This study shows similar diurnal cycles for Jussieu and SIRTA, and slightly lower MLHs at the rural site (see citation on page 9, line 6 of diff.pdf). These results support our assumptions.

Another argument for our assumption is that the terrain around Berlin is quite flat. As the situation might be different in areas with pronounced orography (e.g. a city in a valley) we have explicitly mentioned this here, in the conclusions (page 31, line 9 of diff.pdf) and in the abstract (page 2 line 2 of diff.pdf, see also reply to reviewer #2).

As the most important argument we want to stress that our conclusions also hold if only air quality measurements very close to the ceilometer site are considered (see detailed comment to "p. 18" below).

- p. 10, last paragraph: A graphical sketch (Fig. 2 is not sufficient) on how the COBOLT algorithm works would facilitate understanding. How is "the parameter" defined?

  → *We have significantly extended the description of COBOLT including the most relevant equations. We believe that this extension is sufficiently clear to understand how COBOLT works, so that it is not necessary to add a flow chart to the manuscript as well. Examples of applications under different meteorological conditions and comparisons to independent data sets (e.g. radio sondes; see also comment of reviewer #3) would "overload" this paper and can only be presented in a separate publication. For modifications see pages 11-14 of diff.pdf.*

- p. 12, discussion of Fig. 2 (bottom): from visual inspection, L1 seems to work best in comparison with COBOLT. Do the quality flags really improve the comparison? This aspect is not discussed.

  → *Visual inspection cannot fully reveal all differences of the retrievals, especially as only one day is displayed in Fig. 2. This* *figure is only shown to illustrate the problem and the different solutions. For quantitative conclusions we have included Fig. 3: Here, the differences for the whole period are plotted, separately for the different retrievals. It can be seen that consideration of the quality flag in particular reduces the number of cases where the retrieved MLH is (much) larger than the COBOLT-retrieval (panels b and c).*

On the other hand, it is obvious that the number of successful retrievals is drastically reduced if the quality flags are used as described. With stricter requirements (e.g. quality flag must be 3, i.e. highest level) the number of MLH-retrievals drops from 8346 (if the quality flag is ignored) to 3331 (if only the highest quality is considered) to 2998 (if only the lowest candidate level is considered if it has the highest quality). This fact is described in Section 4.3 and is shown in detail (as a function of the time of the day) in Fig. 4c.

- p. 18, top: Only one ceilometer: this is indeed the main drawback of the investigation (see also General comment).

  → *As already mentioned, we would have been lucky if more ceilometers had been available. This was however not the case and cannot be changed afterwards. On the other hand five air quality stations are very close to the ceilometer site (within 6.4 km, see Table 1). On this spatial scale changes in the diurnal cycle of the MLH are very unlikely, especially as all are in the center of Berlin and no environments like forests or lakes are included. If we restrict ourselves to only these sites (#220, #143, #171, #174, and #124) our conclusions remain valid: it is demonstrated in Fig. 8 that correlations between MLH and PM_{10} are quite variable so that no generally applicable correlation coefficient can be found. We add a corresponding paragraph to the manuscript to emphasize this (pages 23 bottom to page 24 top of diff.pdf).*

  With respect to the correlation between MLH and NO_x the closest stations show positive R due to the strong contribution of traffic emissions. Ozone measurements were not available in the vicinity of the ceilometer site. We cannot exclude that at the outskirts of Berlin the MLH is different due to the different surface properties. The above mentioned investigations (Munich, Paris, Vienna; see reply to comment "p. 8 line 5") suggest that – if differences oc-

cur – they more likely concern the height of the mixing layer than the temporal development. So it can be assumed that the correlations are not very much affected. Nonetheless, for an ultimate clarification more ceilometers and further investigations would be highly desirable as mentioned in our conclusions.

- p. 24, lines 11 and 12: Robustness and representativeness are also not really investigated in this paper.

  → *Our statement was related to previous studies and their shortcomings. But we agree with the reviewer that we have not covered all open questions in our paper: we have only shown for a limited period at one place that the correlations are not representative. And we have investigated the role of the MLH retrieval. So we dropped the word "robustness" to avoid misunderstandings (see page 29 of diff.pdf, beginning of Sect. 6).*

Technical corrections:

- p. 1, line 5: "...has been investigated"

  → *Corrected*

- p. 1, line 7: July and August

  → *The campaign started in June, however, the ceilometer was installed only by the end of June. To be consistent with other papers we would like to leave this sentence unchanged.*

- p. 2, line 9: "...and when meteorological conditions..."

  → *Changed*

- p. 2, line 13: mass concentrations

  → *Changed*

- p. 2, line 27: either "for a chemical box model" or "for chemical box models"

  → *Changed (second option)*

- p. 3, line 3: In particular,

  → *Changed*

- p. 3, line 4: are established,

  → *Changed*

- p. 5, line 8: These findings

  → *Corrected*

- p. 19, Fig. 7: The lines for the outskirts stations are missing

  → *We don't understand this comment: dotted lines are included. Maybe it is a matter of the resolution; however we have checked this on a printed page and it was readable. In case there are problems this can be fixed during the type-setting process.*

- p. 21, line 13: probably ...larger at the outskirts sites.

  → *What is meant is the following: if the diurnal cycle of O_3 concentrations based either on averages or medians are compared, the maximum values during the afternoon are higher by approximately 5 μg m^{-3} in case of averages (page 26, line 29 of diff.pdf, Sect. 5.2). This is valid for all sites where ozone measurements were available. That ozone concentrations at the outskirts (dotted lines) are in general higher (i.e. the comment of the reviewer) is also true.*

- p. 22, Fig. 9, Fig. 10: The lines for the outskirts stations are missing

  → *See reply above on Fig.7.*

Additional references:

- de Bruine, M., Apituley, A., Donovan, D. P., Klein Baltink, H., and de Haij, M. J.: Pathfinder: applying graph theory to consistent tracking of daytime mixed layer height with backscatter lidar, Atmos. Meas. Tech., 10, 1893-1909, doi:10.5194/amt-10-1893-2017, 2017.

- Caicedo, V., Rappenglück, B., Lefer, B., Morris, G., Toledo, D., and Delgado, R.: Comparison of aerosol lidar retrieval methods for boundary layer height detection using ceilometer aerosol backscatter data, Atmos. Meas. Tech., 10, 1609-1622, doi:10.5194/amt-10-1609-2017, 2017.

- Haman, C. L., Lefer, B., and Morris, G. A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment, J. Atmos. Oceanic Technol., 29, 697710, 2012.

- Lotteraner, C. and Piringer, M.: Mixing-Height Time Series from Operational Ceilometer Aerosol-Layer Heights, Boundary-Layer Meteorol., DOI 10.1007/s10546-016-0169-2, 2016.

- Piringer, M., Joffre, S., Baklanov, A., Christen, A., Deserti, M., De Ridder, K., Emeis, S., Mestayer, P., Tombrou, M., Middleton, D., Baumann-Stanzer, K., Dandou, A., Karppinen, A., and Burzynski, J.: The surface energy balance and the mixing height in urban areasactivities and recommendations of COST-Action 715, Boundary-Layer Meteorol., 124, 3–24, doi:10.1007/s10546-007-9170-0, 2007.

Authors' response to reviewer #2

We thank reviewer # 2 for carefully reading our manuscript and the provision of many useful comments and detailed suggestions. We have considered all comments. They gave as useful hints where improvements of the paper were necessary to better understand our methodology and conclusions. Below all points raised by the reviewer are repeated; our comments are added in italics.

The changes (revised version vs. AMTD-paper) are highlighted as displayed by latexdiff ("diff.pdf", maybe renamed when uploaded as a supplement). For the sake of clarity only small changes are explicitly mentioned in our point by point replies, otherwise we refer to the corresponding parts of "diff.pdf" (in blue). Note, that some of our responses interact with comments of the other reviewers, so sometimes it is difficult to refer a change to one specific reviewer's comment.

Point by point replies

General comment

[...] Overall, the paper is well written and easy to follow, but however needs some more critical discussion on certain points.

In my point of view, using just one ceilometers/location might not be sufficient to answer the question given in the title. It is clear, that it is difficult to extent the study to other locations at that stage, but however, this aspect should be discussed more in detail. As highlighted by reviewer 1, I share the opinion that a day/night comparison might be interesting.

For meteorological conditions being a main driver of turbulent mixing it might be interesting to include some meteorological observations characterizing the measurement location and selected study period. With a observation height of about 5 m it might be interesting, which amount of the measured concentration is originated from the actual location and which amount is advected from neighboring areas or "removed" by vertical mixing. Here again a night/day difference would be interesting. How does this aspect influence on the analysis at one selected point?

→ *The first concern of reviewer #2 is the limited number of ceilometers and the missing discussion of day-night differences. This was also one* *of the major criticisms of the first reviewer (please see our reply to reviewer #1 as well).*

We agree with reviewer #2 that more ceilometers would have been beneficial for the study. It was however not possible to set up several ceilometers and/or to use mobile systems. Up to now ceilometers do not belong to the standard equipment of urban air quality networks, maybe this will change in future. So we had to rely on additional resources (in the framework of a campaign) with the inherent limitations (e.g. temporal availability).

Nevertheless we believe that BAERLIN2014 provided very valuable scientific results even if there was only one ceilometer available. We were able to demonstrate how differences in MLH-retrievals play a role for calculating correlations between MLH and air quality parameters. By addressing standard retrievals (the proprietary software of the ceilometer manufacturer) and air quality measurements from an official monitoring network we think that the conclusions are relevant. Based on our research, open questions could be identified one of which being the need for investigations of the variability of the mixing layer over a large municipality. So we hope that in future the wishes of the reviewer (and ours) to have more ceilometers and at least one full annual cycle of the MLH can be fulfilled, and that our paper will be a motivation for setting up the corresponding infrastructure (see also our replies to the detailed comments of reviewer #2 below).

As described also in the reply to reviewer #1 our conclusions on the large spatial variability of correlations between MLH and PM_{10} *are confirmed if we restrict our discussion to the stations nearby (less than 6.4 km from the ceilometer site). Over this small spatial domain the representativeness of a single MLH retrieval is very likely. Our discussion on the correlations between MLH and* NO_x*-concentrations also remain valid when focussing only on the vicinity of the ceilometer site.*

As a consequence we have added a new paragraph (see pages 23–24 of diff.pdf) and more references (see page 9 of diff.pdf). Following the suggestions of reviewers #1 and #2 we have extended Sect. 5.1 by discussing day-night differences and the influence of the wind field (see comment of reviewer #3) (see pages 24–26 of diff.pdf). In addition a short comment on differences between working days and weekends has been added.

A detailed discussion of the influence of the local sources to measure-

ments in 5 m height is beyond the scope of this paper (see the clarification of our objectives in the introduction: page 3, line 9 ff of diff.pdf): such small scale investigations require much better temporal and spatial resolution of the measurements (and associated models). For example, station #42 used for the BAERLIN2014 project as reference is being classified as urban background station. This determines major pollution sources such as major streets to be not within the direct surrounding area (>100 m) and includes usually residential areas. Therefore minor sources like smokers, restaurants, barbecue, and household sources determine the moderate emissions in the vicinity. A moped or car passing the station for a short period of time is not detectable in an averaging period of one hour. Note, that the altitude of the ceilometer (5 m above ground) is not relevant for the determination of the MLH (see also comment on "p.8 line 1" below).

Specific comments

- p.2 line 3-4: Is there evidence in your study? Otherwise put this sentence in the introduction or conclusion.

  → *From our point of view this message is important. That was the reason why it was included in the abstract. Note, that we have written "seems to be unrealistic ... a city like Berlin". It is not meant as a statement that is valid for all metropolitan areas worldwide at any time (for this we indeed do not have evidence); e.g., for cities surrounded by (high) mountain ridges or extreme pollution episodes the situation might be different. The sentence should be understood as a "warning" not to over-interpret correlations between MLH and concentrations of pollutants. To make this clearer we have modified the sentence in the following way: "seems to be unrealistic ... a city like Berlin (flat terrain)", and we have extended the introduction to better explain the scope of the paper (page 3 of diff.pdf).*

- p.2 line 19: measurements, data instead of techniques

  → *Improved*

- p.2, line 27: box models

→ *Corrected*

- p.3, line 16: COBOLT: add one sentence highlighting novelty, functionality

  → *We have added a much more detailed description of COBOLT according to the suggestion of reviewer #1, see pages 11-14 of diff.pdf.*

- p.3, line 19: aim to instead of may

  → *We don't want to change this. The reason is that our study aims to show the influence of the retrieval on the derived MLH and the heterogeneity of the concentrations and thus may help the user to draw conclusions. We don't aim to show a link between air quality and MLH because there are more variables than just the MLH that control pollutant concentrations (see several comments of all reviewers and the statements in our manuscript). However, we have substituted "assess" by "interpret" (page 4, top, of diff.pdf).*

- p.4, line 1: specify "active remote sensing networks" (e.g:...)

  → *We have added (e.g. the above mentioned ceilometer networks); page 4 line 19 of diff.pdf. This refers to the introduction where we have added (e.g. almost 100 instruments by the German Weather Service) (see page 3 of diff.pdf).*

- p.5, line 15: chemical processes? What about Ozone? Where does it come from – downward mixing, secondary formation?

  → *In this section of the paper (introduction) we give an overview over previous publications that are relevant for our study. In the Schäfer et al. (2012) paper ozone was not considered, thus, it is not mentioned here. However, later in our paper we discuss this issue (Sect. 5.2): downward mixing, destruction of ozone by sometimes high NO_x concentrations, production of ozone when NO_x levels are low because of the notable amount of green spaces (parks, forests and leisure areas), or ozone formation by photochemistry (page 27, lines 17 ff, of diff.pdf)*

- p.5 line 28-31: This is not part of your analysis and could be moved to the conclusion.

→ *Thanks for this remark: we agree and delete this part, as these ideas have already been included in the conclusions (so it was sort of a duplication).*

- p.6, line 9: specify "secondary material"

  → *We changed "secondary material" to "secondary aerosol compounds", see page 6, line 30 of diff.pdf*

- p.6, line 14: hourly measurement

  → *Corrected*

- p.8, line 1: how representative is the measurement location in 5 m height for near surface pm10 concentration? How does this impact the representativeness for the MLH measurements for this area?

  → *The ceilometer was installed 5 m above the ground. For the determination of the MLH a change of the altitude of the ceilometer in the range of a few meters is not relevant. Concentrations are measured at the BLUME stations approximately 3.5 m above ground. These values are expected to differ from measurements directly at the curbsite (see Bonn et al., (2016)). The latter might show much higher temporal fluctuations (e.g., passage of a car). Such microscale effects are not considered when correlations with the MLH are investigated. To resolve these problems certainly models at the building-resolving scale help. Moreover, during the transport from e.g. major traffic sites to the reference location strong vertical gradients will be smoothed (see also reply to "General Comment"). We have now briefly touched the topic of "scales" in the conclusions (page 31 of diff.pdf).*

- p.16 Figure 5: legend has to be added

  → *Done*

- p.17 line 2 ff: this chapter defines the scope of the study and in my opinion appears to late in the manuscript which results in a misbalance between introduction/methods and results. The first part until 5.1 is more an introduction to a new topic than a presentation of results. I might be helpful to include some of these aspects in the introduction (without changing the whole manuscript). Line 2: Ozone and NOx also measured at BLUME?

→ *Thanks for this suggestion. Indeed this paragraph does not fit here very well. We have completely rephrased and re-arranged this paragraph. We removed text that was not relevant for our study. We moved the modified text to section 3 (new caption: "The BLUME network and the BAERLIN2014 campaign"). Now, all information related to the underlying data sets are combined in one section. We have also explicitly mentioned the distance of the BLUME-stations to the ceilometer as this is an essential point in view of the correlations discussed later (end of Sect. 3, page 9 of diff.pdf).*

Ozone and NO_x are also measured by the BLUME network. This becomes clearer after moving text from Sect. 5 (introductory remarks) to Sect. 3, see above (from page 21 of the AMTD-version to page 7 of diff.pdf).

Moreover, we introduced a new paragraph to the introduction to make the scope of our study more clear (see also reviewer #1; page 3 of diff.pdf).

- p.18 line 8: it is unclear on which basis the median was calculated. 67 measurements each hour at every station?

→ *Yes, this is true for the concentration measurements at the BLUME-stations: the temporal resolution is one hour, and the whole measurement period of 67 days (i.e., when co-incident PM_{10} and MLH measurements were available) is considered. With respect to the MLH we rely on all available 10-minutes retrievals (up to six, depending on the MLH-retrieval) of the corresponding hour, for all 67 days. So, up to 402 MLH-values are considered for the MLH-median. An new paragraph has been added to the end of Sect. 4.4 (pages 19–20 of diff.pdf).*

- p.18, line 16-20: can you proof your assumptions by adding meteorological observation here? Is there a secondary circulation generated by the Urban Heat Island? Please specify the term "meteorological interpretation".

→ *We are aware that we missed to clearly outline the scope of the* *paper as necessary. We have corrected this now by adding a new paragraph to the introduction (page 3 lines 10 ff of diff.pdf).*

In this context the term "meteorological interpretation" should be understood as the interpretation of processes that control the development of the mixing layer and surface concentration – and their interaction. A thorough discussion of the meteorological reasons and atmospheric chemistry responsible for the observed distribution of pollutants was however not the goal of the study. Nevertheless we have included several comments to point at reasons for poor or unexpected correlations.

Finally we want to emphasize that we present diurnal cycles of MLH and concentrations averaged over 67 days. The analysis of the interactions between meteorological fields (e.g. wind), atmospheric chemistry and emissions should rather be carried out with a high temporal resolution. This analysis would certainly benefit from a "complete" set of observations. Such a data set is however unrealistic. Thus, tentative answers may be found by numerical models. But models do not necessarily display proof of understanding or concepts but provide a further tool and support understanding. Anyway, such investigations are far beyond the scope of this paper (see a short comment in the conclusions page 31 of diff.pdf).

- p.20, line 31f: see comment above

  → *See previous reply*

- p.21 general: here you mention briefly the problem of point measurements. This aspect could be further discussed. It is interesting if there is a mismatch between the timing of MLH and air quality observation. Does a low MLH mean a high concentration at the same time? What is the order of the processes? Where do meteorological conditions come into play?

  → *The answer to this question is closely related to the previous replies. It is not unlikely that a temporal delay between MLH and concentrations might occur, however, this delay is influenced by e.g. the wind field (upwind/downwind, low/high wind speed) or specific characteristics of the traffic (emissions). In case of secondary produced constituents it depends on the concentration* *of the precursors and the solar irradiance. These influences are certainly time dependent, so it is hardly possible to detect them when long temporal averages are considered as in our study. We have briefly discussed the influence of the wind speed at the end of Sect. 5.1 by adding several paragraphs and Table 3 (pages 24–26 of diff.pdf).*

- Chapter 6: It is not the extended mixing layer itself which is the initial precursor of dilution of pollutants near the ground. Several processes interact which each other which might as well lead to an extension of the mixing layer height.

  → *We agree with the reviewer. We have mentioned the complex distribution of pollutants several times in the paper. Our goal was to compare this complexity with different schemes to determine the MLH from ceilometer data.*

Authors' response to reviewer #3

We thank reviewer # 3 for carefully reading our manuscript and the provision of many comments and suggestions. They gave as useful hints where improvements of the paper were necessary to better understand our methodology and conclusions. However, some of the raised questions have already been discussed in the submitted paper at different places, and some are clearly beyond the scope of the paper and/or cannot be resolved with the available data sets. Nevertheless we have considered all comments of reviewer #3. Our replies are given in italics.

The changes (revised version vs. AMTD-paper) are highlighted as displayed by latexdiff ("diff.pdf", maybe renamed when uploaded as a supplement). For the sake of clarity only small changes are explicitly mentioned in our point by point replies, otherwise we refer to the corresponding parts of "diff.pdf" (in blue). Note, that some of our responses interact with comments of the other reviewers, so sometimes it is difficult to refer a change to one specific reviewer's comment.

Point by point replies

General comment

This is an interesting manuscript as it discusses the relationship of the mixing layer height (MLH) and near surface pollutant concentrations. The authors perform correlations of the MLH and PM10, NOx, and O3, and found varying results. The authors believe that the effects of the heterogeneity of the emission sources, chemical processing and mixing during transport exceed the differences due to different MLH retrievals. With regard to the use of the different MLH retrieval methods (Vaisala proprietary software, COBOLT), which are solely based on aerosol backscatter signal, I was wondering, if radiosondes have been used for a conclusive validation during the BAERLIN campaign.

→ *Intercomparisons between aerosol-based MLH retrievals (lidars, ceilometers) and retrievals based on temperature-, wind- or water vapor profiles (e.g. from radio sondes) have been carried out in several studies; some papers have been cited in the manuscript. COBOLT has been developed using ceilometer measurements in Munich and compared with radio sondes data of Oberschleißheim (distance 8 km only). So we don't feel that it is necessary to demonstrate this again in this* *paper. Moreover, it was not a goal of BAERLIN2014, and the closed radio sonde station is in Lindenberg, almost 60 km away from the ceilometer site!*

Also, I was wondering why other methods such as the Haar wavelet method or a cluster method have not been considered/discussed.

→ *The Haar wavelet method is one component of COBOLT when the Sobel operator is applied (see new citation of Comeron et al., 2013). This is mentioned in the revised version of the manuscript when we provide a much more detailed description of COBOLT (according to a suggestion of reviewer #1, pages 11–14 of diff.pdf). Moreover, we had already included three citations in the original manuscript (Cohn and Angevine, 2000, Brooks, 2003, Baars et al., 2008) that use this wavelet covariance transform. Caicedo et al. (2017) who applied the cluster method are cited as well (see response to reviewer #1).*

With regard to the relation of the MLH with air quality, it is well known that the local change of any given pollutant is not only controlled by the MLH, but by a combination of emission, chemical transformation, removal, advection, convection and turbulent mixing. Also, it is known that at the microscale level urban structures cause flow disturbations and thus deviations from the mean air quality of a larger, representative fetch in an urban area. An example is the well-known wind rotor system in street canyons. Thus the relationship of the MLH with surface concentration critically depends on the fetch area representative for a given measurement site. These well-known processes are not properly addressed in the paper.

→ *We agree with the reviewer that surface concentrations of pollutants do not only depend on the MLH and that our paper is not the first that points out these facts. Accordingly we have mentioned these processes e.g. in Sect. 5.1 (p 20, l27 ff. of the AMTD-version, p 20, l32 ff. including citations) and in the conclusions (p 25, l29 of the AMTD-version). As a consequence of the comments of reviewer #3 we have extended this discussion. Moreover, we have added a paragraph to the introduction where we describe the objectives of our study more clearly (see page 3 of diff.pdf). This was indeed not clear enough in the submitted manuscript: we want to focus on the ceilometer retrieval and the potential over-interpretation of correlations. These aspects have not yet been covered in the literature.*

The reviewer's statement on the influence of "microscale level urban structures" certainly points out a very important aspect, which in sum would however have resulted in exploding project costs. Some aspects e.g. of chemical transformation and deposition can be reasonably well while not perfectly described by a chemical boxmodel. However, the vertical mixing aspect in such a model, determined by the MLH cannot be reproduced acceptably well without observations. The information provided by BAERLIN2014 supplies the effect of the MLH and therefore vertical exchange to the change of pollutants, i.e. the fraction of change that can be explained by meteorology.

Due to the rather flat larger area of Berlin, it can be expected that transport processes may play a dominant role in the distribution of pollutants, both at the mesoscale and microscale level. I am surprised to see that the authors did not consider any of the findings associated with the BERLIOZ experiment in 1998 (mostly published in Journal of Atmospheric Chemistry 42, 2002, but also others), which focused on the upwind-downwind conditions found for the Berlin case, as well as the pollutant concentrations within the boundary layer and aloft in the same area and the impact of long-range transport.

→ *The BERLIOZ campaign (Berlin Ozone Experiment) was a huge campaign focussed on the impact of Berlin on the surroundings. It never investigated Berlin itself but a northwest-southeast transect through Berlin approximately 50 km on either side in Brandenburg, such as e.g. Pabsthum. In contrast the focus of BAERLIN2014 was the metropolitan area of Berlin and Potsdam and the influence of vegetation inside this area in detail. Thus, one could use references and results of BERLIOZ for broader discussions only.*

Reviewer #3 seems to focus rather on large scale effects than on small scale mixing. Berlin is not affected by the surrounding countryside, somewhat more the opposite. This actually caused the BERLIOZ project nearly to fail, because the anticipated effects were hardly found (e.g. huge ozone plumes downwind, large PM_{10} clouds etc.). As stated earlier the idea of the reviewer to conduct investigations for Berlin and Brandenburg including (very) detailed experiments and modelling approaches is far off realistic financial and personnel limits.

In conclusion, we don't feel it necessary to include any outcome from BERLIOZ as the scientific objectives of that experiment were quite different from our study. Neither the derivation of the MLH from *ceilometer measurements was part of BERLIOZ nor the distribution of pollutants inside the city.*

I would not expect an unambiguous relationship between the MLH and surface concentrations at any given location and under any given meteorological situation in the Berlin area. Rather, I would only expect a dominant role of the MLH on surface concentration, when advection is at a minimum, i.e. under stagnant wind conditions. In its current form this paper neglects the discussion of the MLH with regard to different wind regimes, both with regard to wind speed as wind direction. It should also be mentioned that not only pollutants can be transported, but also physical properties of the boundary layer including the MLH depending on the history of air masses. This extended in-depth analysis is a crucial requirement for a potential publication in AMT.

→ *We agree with the reviewer that advection plays a relevant role for the correlation between MLH and concentrations. This was briefly mentioned in the manuscript (see answer above). We have also elaborated this aspect in more detail in the revised version (pages 24–26 of diff.pdf) taking into account wind measurements of the German Weather Service at three sites in Berlin. We use these additional data set to select days when the wind was "predominantly stagnant". However, we want to emphasize that our mean diurnal cycles (MLH, concentrations) are averages over two months. So, the assessment of the contribution of a single process to the correlation between MLH and surface concentrations is hardly possible.*

More specific, mostly minor issues

- Page 2, L25-28: The paper by Czader et al. (2013) should be added as it is one of the earlier examples to use ceilometer derived MLHs for validation in conjunction with comprehensive air quality modeling.

  → *In Czader et al. (2013) we only find the reference to "remote sensing techniques" providing MLH at one site (Moody tower). Details were however found in Haman et al., 2012: here, CL31 measurements of almost two years have been evaluated for the diurnal cycle of the MLH in Houston, Texas. They use proprietary software of Vaisala. We have added both citations (pages 3 and 6).*

- Page 5, L18-19: I think both terms MLH and Hml mean the same. I suggest to use one term throughout the entire text.

  → *Our idea was to use MLH as "word" in the text, and H_{ml}, $H_{ml,v}$ etc. for mathematical expressions. We have checked this for consistency and changed it whenever necessary.*

- Page 5, L22: Please define what "width" would mean exactly: horizontal or vertical?

  → *Width is related to the MLH as derived from ceilometer measurements. As this could be misunderstood we changed the sentence to ...into intervals of 200 m. (see page 6 of diff.pdf).*

- Page 6, L9: Please define what is meant exactly by "secondary material"?

  → *We changed "material" to "secondary aerosol compounds", see page 6 of diff.pdf*

- Page 7, L1-3: "These data....in whole Germany". Is this statement important in understanding the contents of the paper? I suggest to remove it.

  → *We removed it as it is indeed not essential for the understanding. Anyway, for me it was an interesting information showing the extent of automatic air quality stations currently operational in Germany (see page 7 of diff.pdf).*

- Page 8, L4: What "information" is exactly meant?

  → *We have clarified this sentence: ...the option to combine in-situ measurements at the surface with data concerning the vertical direction (see page 9 of diff.pdf). The combination with aerosol optical depth would be another example. MLH is also useful to constrain model calculations as mentioned (see page 2 of diff.pdf).*

- Page 8, L6: Suggest to remove ", which is one hour different to UTC.", as UTC is not being used in the paper.

→ *This was included only as a explanation for readers who are more familiar with GMT. But we agree that it can be removed (see page 9 of diff.pdf).*

- Page 10, L16: Please explain what is actually meant by "cross-platform" here, and why it would be helpful?

  → *We wanted to emphasize that the code can be run on Windows and Linux platforms, so it is potentially useful for a large community. Moreover, Phyton is free of charge in contrast to e.g. MatLab. We have extended the whole section in accordance with the comments of reviewer #1; in this context we have also considered the comments of reviewer #3 (pages 11–14 of diff.pdf).*

- Page 15, L26: "Concentration measurements" of what?

  → *This could be "everything". In our study concentrations (PM_{10}, NO_x, O_3) are discussed but – if the corresponding data sets are available – the statement is also true for any other trace gas or e.g. $PM_{2.5}$. The sentence should only emphasize that problems may occur if data sets with low temporal resolution are considered during the rapid growth of the ML. To make this clearer we have substituted one word (page 18, line 21 on diff.pdf).*

- Page 15, L30: "The latter...(Pappalardo et al., 2014)". Please explain the schedule of EARLINET and explain whether the BAERLIN approach was important for the EARLINET approach or the other way round (which is more likely).

  → *The EARLINET schedule was defined in the year 2000. On the one hand it considers the diurnal cycle of the ML (measurements when the vertical extent is approximately constant for several hours) and on the other hand the performance of Raman lidars (they perform better during night). This was not influenced by our study, and our study is independent of the EARLINET schedule as we determine the full diurnal cycle. We only mentioned this because our (and similar) results confirmed that the selection made by EARLINET was reasonable (see COBOLT-retrieval shown in Fig. 5). For further illustration we have included Fig. 6. It shows the differences of the afternoon values of MLH when different MLH-retrievals are applied.*

To make this clearer we have modified the corresponding sentence as follows: The latter has been the reason for including a measurement around two hours after local noon in the regular EARLINET schedule (page 18 of diff.pdf).

- Page 15, L32-33: Please mention that these specific COBOLT results refer to the entire campaign period.

  → *We added ...for the whole period of 67 days (page 18 of diff.pdf).*

- Page 17, L11-12: What is exactly meant by "All measurements are performed under ambient conditions"? They way it is written it would mean that the air quality station was not air-conditioned.

  → *According to a comment of reviewer #2 we completely rephrased and re-arranged this paragraph. In this context we have also considered the comments of reviewer #3 and removed things that were not relevant in the context of our study (see page 7 of diff.pdf and answer to "Page 17, L15-16" below.*

- Page 17, L18: I think this "significant horizontal heterogeneity" refers to surface measurements here. Please clarify.

  → *Yes. Most of the measurements were made from bicycles. We have clarified this: Episodic mobile (bicycle) measurements from BAERLIN2014...(page 21 of diff.pdf)*

- Page 17, L15-16: What is exactly meant by "inorganic species": gas-phase, particle bound or both?

  → *Inorganic species refer to gaseous compounds like CO, NO and NO_2. The whole paragraph has however been rephrased and re-organized according to a suggestion of reviewer #2. Main parts have been moved to Sect. 3 (from page 21 of diff.pdf to page 7 of diff.pdf), and unnecessary information was deleted.*

- Page 17, L17: The reference "von Schneidemesser et al., 2017" is still in preparation and therefore not citable.

  → *We removed this citation and the text (lines before this citation) that was related to this paper which is currently still under preparation.*

- Page 17, L17-18: "Here we do not discuss these topics...". In this case please remove the preceding L13-17 as they are not within the scope of this paper.

  → *See our above response to the comment on "Page 17, L11-12"; the sentence was removed.*

- Page 18, L26-28: What is the justification for using these different correlations? The statistically most reliable quantity would be the median anyway, as it minimizes the impact of outliers. This is in particular true for such a quantity as PM10, which is mostly primarily emitted.

  → *We agree with the reviewer: that was the reason why we use the median in Figs. 7, 9 and 10 in the AMTD-manuscript. The same argumentation as the comment of the reviewer was given in the original manuscript (page 20, lines 6 ff). The different combinations of averages and medians as defined on page 18, lines 26 ff were only introduced to demonstrate the consequences on the correlations in the subsequent discussion. See also the new Fig. 7 (page 20 of diff.pdf).*

- Page 20, L23: "...with a lot of vegetation, a high density of buildings...". This sounds like a contradiction: where there is high density of buildings how can there be lot of vegetation at the same time?

  → *This description is made from the perspective of a German citizen. A "high density of buildings" does not mean that there is no space left for trees, bushes etc., often arranged as small "parks" of some tens of meters in length and width, or buildings organized as squares with trees inside a yard, to increase the quality of living. For example, southeast of the ceilometer is an area of approximately 100 × 70 m with "a lot of vegetation" whereas buildings dominate elsewhere. To avoid misunderstands we replace "high density of buildings" by "in a typical residential neighborhood in the inner part of a big German city; see page 24 of diff.pdf". A similar expression has also been used in Sect. 3.*

- Page 20, L17-18: The authors mention aerosol formation. Would PM10 data provide any indication for aerosol formation? If so, please explain.

→ *When we summarized our conclusions from Fig. 9, we mentioned different aspects that are responsible that no unique correlation coefficient (MLH vs. PM_{10}) has been found for entire Berlin. In this context only the absolute value of PM_{10} is relevant.*

- Page 20, L17: The authors mention that relative humidity may have an impact on PM10. Would PM10 concentration decrease or increase with relative humidity?

  → *The whole paragraph was significantly extended (see also replies to Reviewers #1 and #2) by including more investigations on correlations under special meteorological conditions (see pages 24–26 of diff.pdf). In this context the statement on the relative humidity became unnecessary (one can assume a small increase due to uptake of HNO_3).*

- Page 21, L1-2: What classes in addition would the authors recommend?

  → *We do not necessarily need more classes but the attribution might be reviewed. However we don't have any influence on this classification and the criteria for this classification. The same is true for the selection of the locations of the air quality stations. It is not unlikely that political reasons have a certain influence as well. We have added a short remark at the end of Sect. 5.1 (page 26 of diff.pdf).*

- Page 21, L2-4: This statement is obvious and has been considered in many urban air quality networks over many decades.

  → *This conclusion is indeed not unexpected. Nevertheless many publications do not clearly describe the conditions under which their correlation has been calculated or use only one site in a metropolitan area and leave it open how representative their conclusions are. So there remains room for misunderstandings, and we feel that it is justified/necessary to emphasize this statement (again). Accordingly we have expressed this objective in the updated introduction (see page 3 of diff.pdf).*

- Page 21, L31 - Page 23, L5: It is well-known that O3 can be mixed from the residual layer into the convective layer, also for the case of

Berlin (e.g. see BERLIOZ special issue in the Journal of Atmospheric Chemistry 42, 2002). The excellent correlation of the MLH with ozone in urban areas may not be surprising at all, as both processes are ultimately driven by incoming solar radiation provided there are sufficient precursors for O3 formation available. In other words the relation between the MLH and O3 is apparent, but not causally determined. This should be mentioned.

→ *We agree with the reviewer: we have used the same argumentation in that paragraph of the AMTD-version of the manuscript including the citation of a paper by Fallmann et al.; so it has already been mentioned. To make this clearer we slightly rephrased this paragraph (see page 27 bottom of diff.pdf).*

- Page 24, L7: "Whether ...studied". I suggest to remove this sentence. It is obvious that the potential impact of the MLH on ambient concentrations decreases with decreasing distance to the corresponding emission source.

  → *We have removed this sentence.*

- Page 24, L13: As I remember Xu et al. (2011) do not report MLH observations and thus no correlation with primary or secondary pollutants.

  → *Xu et al. (2011) discussed the influence of the MLH on surface concentrations of several trace gases in a general way. However, they did not use own measurements of the MLH. As a consequence we agree with the reviewer that this citation is not really relevant and dropped it.*

- Page 25, L10-12: I guess it is well-known that there is not one only parameter which controls surface concentrations.

  → *Again we agree with the reviewer, but our motivation was to investigate the role of the MLH-retrieval for correlation studies in view of its uncertainty and the inhomogeneity of urban air quality. This message was obviously not as clear as it should have been (see corresponding comments of all reviewers). As a consequence we have added a clear statement on our objectives to the introduction (see page 3 of diff.pdf).*

- Page 25, L27-28: In their paper the authors have tried to argue that MLH is not the only parameter which controls surface pollutant concentrations. Why then would it be of interest to perform a winter study in Berlin and why is it of importance that PM10 concentrations are 50% higher in winter compared to summer in Berlin. If there is no consistent correlation of MLH with PM10 in summertime why should it be different in wintertime?

  → *Reviewers #1 and #2 regret that only data from two months were available. This cannot be changed for obvious reasons. However, we believe (together with the reviewers) that a longer observation time would provide additional insight: in winter the MLH is expected to be shallower, the concentration of PM_{10} is larger, and the meteorological conditions (including atmospheric chemistry) are different. We do not expect that in winter the MLH is the only parameter that controls the concentration of pollutants, but it is not clear if the variability of R is more or less pronounced. It is scientific tradition to investigate any problem under different conditions if possible (see our comments on available resources). To point this out we have added an additional explanation: We do not expect that in winter the MLH is the only controlling parameter, but it is not clear if the correlation (and its variability) is more or less pronounced (see page 31, lines 17 ff of diff.pdf).*

- Page 26, L4-6: The authors state that MLH data is beneficial for box-model calculations and validation of chemistry transport models. While I would agree on the authors statement in this sentence I do not completely understand what the authors justification would be for this, since according to their paper the authors largely argue that there is no consistent correlation of the MLH with air pollutants. This should be clarified.

  → *In any case there are multiple counteracting processes merging in our findings as has been mentioned in the paper. As a consequence interpretation is much more complex than simply getting $R \approx \pm 1$ for all times, but this does not reduce the usefulness of a reliable determination of the MLH. This was our statement in the last paragraph of our conclusions. We have clarified this by extending the conclusions on validation and combination of models and measurements (see page 31, bottom, of diff.pdf). It would,* *e.g., be nice to tackle the question of the homogeneity of the MLH over a city like Berlin by models and compare the results with distributed ceilometer measurements (see our corresponding replies to similar questions of all reviewers), maybe possible in future. Moreover, model calculations can help to understand the interaction and the relevance of different meteorological and chemical processes; in this context it could be useful to have independent measurements to validate at least parts of the model output (again a question of resources to set up a adequate field campaign).*

Additional references (for more see also reply to reviewer #1):

- Czader, B. H., Li, X., and Rappenglueck, B.: CMAQ modeling and analysis of radicals, radical precursors and chemical transformations, J. Geophys. Res., 118, 11,376–11,387, doi: 10.1002/jgrd.50807, 2013.
- Haman, C. L., Lefer, B., and Morris, G. A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment, J. Atmos. Oceanic Technol., 29, 697710, 2012.

Mixing layer height as an indicator for urban air quality?

Alexander Geiß[1,6], Matthias Wiegner[1], Boris Bonn[2,7], Klaus Schäfer[3], Renate Forkel[3], Erika von Schneidemesser[2], Christoph Münkel[4], Ka Lok Chan[1], and Rainer Nothard[5]

[1]Ludwig-Maximilians-Universität (LMU Munich), Meteorological Institute, Theresienstraße 37, 80333 München, Germany
[2]Institute for Advanced Sustainability Studies (IASS), Berliner Straße 130, 14467 Potsdam, Germany
[3]Karlsruhe Institute of Technology (KIT) – Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany
[4]Vaisala GmbH, Weather Instruments, Notkestraße 11, 22607 Hamburg, Germany
[5]Senate Department for Urban Development and the Environment, Brückenstraße 6, 10179 Berlin, Germany
[6]now at: ESA ESTEC, Keplerlaan 1, 2201AZ Noordwijk, The Netherlands
[7]now at: Albert-Ludwig-University, Institute for Forest Sciences, Georges-Köhler-Allee 53, 79110 Freiburg, Germany

Correspondence to: Matthias Wiegner
(m.wiegner@lmu.de)

Abstract. The mixing layer height (MLH) is a measure for the vertical turbulent exchange within the boundary layer, which is one of the controlling factors for the dilution of pollutants emitted near the ground. Based on continuous MLH measurements with a Vaisala CL51 ceilometer and measurements from an air quality network, the relationship between MLH and near surface pollutant concentrations ~~have~~ has been investigated. In this context the uncertainty of the MLH retrievals and the representativeness of ground-based in-situ measurements are crucial. We have investigated this topic by using data from the BAERLIN2014 campaign in Berlin, Germany, conducted during June and August 2014. To derive the MLH three versions of the proprietary software BL-VIEW and a novel approach COBOLT were compared. It was found that the overall agreement is reasonable if mean diurnal cycles are considered. The main advantage of COBOLT is the continuous detection of the MLH with a temporal resolution of 10 minutes and a lower number of cases when the residual layer is misinterpreted as mixing layer. We have calculated correlations between MLH as derived from the different retrievals and concentrations of pollutants (PM_{10}, O_3 and NO_x) for different locations in the metropolitan area of Berlin. It was found that the correlations with PM_{10} are quite different for different sites without showing a clear pattern, whereas the correlation with NO_x seems to depend on the vicinity of emission sources in main roads. In case of ozone as a secondary pollutant a clear correlation was found. We conclude that the effects of the heterogeneity of the emission sources, chemical processing and mixing during transport exceed the

differences due to different MLH retrievals. Moreover, it seems to be unrealistic to find correlations between MLH and near surface pollutant concentrations representative for a city like Berlin (flat terrain), in particular when traffic emissions are dominant. Nevertheless it is worthwhile to use advanced MLH retrievals for ceilometer data, e.g. for the validation of chemical transport models.

1 Introduction

Air pollution is one of the major environmental issues in metropolitan areas because of its adverse effects on human health (e.g. Chen and Kann, 2008; Rückerl et al., 2011; Lelieveld et al., 2015). Strong emissions, e.g. from traffic, industry or heating, can drastically decrease air quality in particular when the emitted pollutants are captured below an inversion, and when meteorological conditions prevent an exchange of polluted and clean air. Without effective vertical mixing and advection pollutants can accumulate in the lowermost atmospheric layers and concentration thresholds as defined e.g. by the European Union Air Quality Standards (Directive 2008/50/EC) may be exceeded. For this reason several trace gases and particle mass ~~concentration~~ concentrations (diameter below 10 μm, PM_{10}) are continuously monitored by air pollution monitoring networks near the surface implemented by federal or state administrations. In case of an exceedance of legally binding thresholds measures to reduce pollution are mandatory. This could e.g. include restrictions for motorized individual traffic.

Surface concentrations of gaseous pollutants as nitrogen oxides (NO_x), ozone (O_3), sulfur dioxide (SO_2) or carbon monoxide (CO) as well as particulate matter are routinely measured by in situ monitoring stations. Gaps of in-situ measurement networks can be filled by data from remote sensing techniques (e.g. Gupta et al., 2006; Martin, 2008) or numerical models. To better understand – or supplement – direct observations, air quality may be linked to integral parameters such as the aerosol optical depth (e.g. Koelemeijer et al., 2006; Schäfer et al., 2008; Li et al., 2016) or to meteorological parameters such as the height of the mixing layer (henceforward referred to as MLH or H_{ml}). The MLH can be considered as a measure for the vertical mixing within the atmospheric boundary layer and determines the dilution of pollutants which are emitted near the ground. Therefore, the MLH is frequently examined in evaluation studies of regional chemistry transport models (LeMone et al., 2013; Scarino et al., 2014; Brunner et al., 2015; Kuik et al., 2016) or serves as an input parameter for chemical box ~~model~~ models (Knote et al., 2015). Due to the close relationship between turbulent vertical exchange and near surface air quality, several attempts have been made to establish correlations between MLH and near-surface pollutant concentrations (examples will be given in section 2). The underlying assumption is that high concentrations close to the surface may coincide with shallow mixing layers and vice versa. This assumption, which is used although vertical mixing is certainly not the only controlling process (e.g. Elminir, 2005; Tandon et al., 2010; Svensson et al., 2011), will be examined in this paper. Our study is based on two months of data in summer from the BAERLIN2014 campaign (Berlin air quality and ecosystem research: local and long-range impact of anthropogenic and natural hydrocarbons) in Berlin, Germany (Bonn et al., 2016) .

A frequently used approach to determine MLH is the implementation of so called ceilometers, automated and eye-safe single-wavelength backscatter lidars (Wiegner et al., 2014). For this purpose As there is no strict definition of the technical specifications of a "ceilometer" recently the term "ALC" (automated lidars and ceilometers) has been introduced and is often used synonymously. Though originally designed for only determining cloud base heights ceilometers are now used for a variety of more sophisticated activities such as the retrieval of the particle backscatter coefficient β_p and mixing layer height. Since ceilometers are commercially available including software providing "atmospheric products" (e.g. the MLH) we feel that it is necessary to scrutinize the application of such products. This is the main motivation and objective of our paper: to investigate the potential of proprietary software to derive the MLH, and the usefulness of correlations between such derived MLHs and surface concentrations of pollutants in an urban environment. The motivation of the latter is to increase the awareness that such correlations might be prone to over-interpretation. A thorough discussion of the meteorological reasons and atmospheric chemistry responsible for the observed distribution of pollutants is beyond the goal of the study.

For the determination of the MLH range corrected signals of ceilometers ALC can be analyzed (e.g. Morille et al., 2007). In particular since networks of such single-wavelength backscatter lidars, operating continuously, are established they gain increasing importance, often using proprietary software (e.g. Haman et al., 2012) . Recently ceilometer networks have been installed by several national weather services (e.g. almost 100 instruments by the German Weather Service), and it is expected that in near future dense networks providing data in real time will be available on a European scale. This might prospectively increase their relevance Prospectively also the implementation of urban networks for air quality studies is likely at least for selected cities occasionally suffering from pollution events – recently three ceilometers were set up in larger Paris for this purpose (OCAPI: Observation de la Composition Atmosphérique Parisienne de l'IPSL).

However, the retrieval of MLH is an issue even though state-of-the-art ceilometers provide a clear identification of aerosol layers; often several atmospheric layers are detected but it remains ambiguous which one is the mixing layer. This problem can be severe, especially in case of automatic retrievals optimized for specific atmospheric conditions. Retrievals might fail or lead to under- or overestimates if the aerosol concentration is extremely low or high, or if the range of incomplete overlap of the instrument is too large. Consequently any correlation between MLH and pollution – and thus the potential to use the MLH in discussions of air quality – might depend on the selected MLH retrieval technique. In this paper we want to investigate this topic by applying different MLH retrievals provided by the manufacturer of the ceilometer (in our case Vaisala) and a novel scheme COBOLT (Geiß, 2016). We have calculated correlations with concentrations of pollutants at different locations in the metropolitan area of Berlin to compare the effects due to the spatial inhomogeneity of pollutants and due to uncertainties of the MLH retrieval. The results may help to ~~assess~~ interpret possible links between air quality and MLH~~. Our study is based on two months of data in summer from the~~ , even though there was only one ceilometer available during BAERLIN2014~~campaign (Bonn et al., 2016; ?)~~ .

A selection of studies dealing with the link between MLH and pollutants is introduced in the next section. Then, we briefly describe the air quality network of Berlin and the measurement campaign. Section 4 provides a detailed description of different options to retrieve the MLH including a comparison. Correlations with concentrations of pollutants are discussed in view of their dependence on the selected MLH retrieval, and their location inside Berlin. A summary concludes the paper.

2 Relation between mixing layer height and surface concentrations

In this section a brief overview of studies dealing with the retrieval of the mixing layer height and its role with respect to air quality is given.

When discussing retrievals of the MLH it is important to note that it is defined in different ways, depending on the availability of specific measurement techniques and data sets. Most approaches are based on the analysis of either the temperature profile (e.g. Liu and Liang, 2010), the wind field (e.g. Schween et al., 2014) or concentration profiles of particles (e.g. Haeffelin et al., 2012). With the establishment of active remote sensing networks ~~the~~ (e.g. the above mentioned ceilometer networks) the latter approach is gaining importance, basically it is assumed that the concentration of particles considerably decreases at the transition from the mixing layer to the free troposphere. Thus, the analysis of particle backscatter is a promising approach to determine the MLH.

A thorough review of approaches to determine the MLH was given by Seibert et al. (2000). They emphasized the benefit of active remote sensing techniques as they allow measurements of the vertical ~~profiles~~ distribution of particles. Intercomparisons have shown (e.g. Emeis et al., 2004; Wiegner et al., 2006; Emeis et al., 2012) that sodar and RASS can ~~also~~ be used to monitor MLH, however, they usually cannot provide the full diurnal cycle of the MLH in Central Europe, especially in summer (Piringer et al., 2007) . Moreover, these techniques are less frequently applied mainly because of their more complicated implementation and higher expenses for investment and maintenance. The same is true for sophisticated multi-wavelength lidars (e.g. Baars et al., 2008), sodars (e.g. Beyrich, 1995), combinations of instruments (e.g. Cohn and Angevine, 2000), and combinations of models and measurements (e.g. Bachtiar et al., 2014). A large number of studies relying on lidar data has been published introducing different methodologies to determine MLH: among others Endlich et al.

(1979) and Flamant et al. (1997) used algorithms based on first derivatives of the backscatter signal, Menut et al. (1999) used second derivatives, Hooper and Eloranta (1986) the temporal variance, Cohn and Angevine (2000), Brooks (2003) and Baars et al. (2008) applied wavelet covariance transforms, de Bruine et al. (2017) used graph theory, Caicedo et al. (2017) cluster analysis, and statistical methods were used by (e.g. Eresmaa et al., 2006; Lange et al., 2014). With recent upgrades of their hardware and the hardware those methodologies can also be applied to ALC, and with the implementation of networks ceilometers they have become more attractive as they provide continuous monitoring and good spatial coverage.

The role of the mixing layer for pollution and its adverse effects on health have been highlighted since more than 50 years (Holzworth, 1964; Barlow, 2014). Consequently the link between air quality (in terms of particulate matter and concentrations) and MLH was investigated in many studies, primarily for urban areas. It should be emphasized that a comparison of different studies is inherently difficult, especially when only qualitative conclusions have been made. On the one hand different meteorological conditions and species are investigated, i.e. different gaseous pollutants and different sites in rural or urban environments. On the other hand there are conceptual differences, i.e. statistical analyses are based on hourly values, daily averages or diurnal cycles averaged over several weeks or even seasons, and there are different approaches to determine the MLH from measurements or numerical models. Moreover, there are differences with respect to the selection of a suitable MLH parameter used for correlation analysis: averages, medians or certain percentiles are used, maximum values, or MLHs are grouped in intervals.

Studies relying on numerical parameterizations were conducted e.g. by Tiwari et al. (2014) who used reanalysis data, and Du et al. (2013) who used routine meteorological observations to find an anti-correlation between $PM_{2.5}$ and MLH for Delhi, India, and Xi'an, China. Rost et al. (2009) found a strong anti-correlation between PM_{10} and MLH (derived from radio sonde data) for Stuttgart, Germany, with a coefficient of determination of $R^2 > 0.95$. The awareness of the potential of active remote sensing started at the end of the last century when the first generation of ceilometers was deployed. These systems suffered from low pulse energies so that their use was confined to winter measurements or clear atmospheric conditions when the measurement range of the instrument was sufficient to cover the complete vertical extent of the mixing layer. Schäfer et al. (2006) deployed CT25k- and LD40-ceilometers (Vaisala) but primarily relied on co-located sodar-data when they found a high anti-correlation between PM_{10} and MLH in Hannover and greater Munich, Germany, for winter conditions. They also found negative correlations for CO and NO_x with quite variable R^2 depending on the site and the horizontal wind. Differences between summer and winter measurements were also observed. These finding findings agree with results from a campaign in Budapest, Hungary, with a similar set of instruments (Alföldi et al., 2007). Examples with state-of-the-art ceilometers include Beijing, China (e.g. Sun et al., 2013; Tang et al., 2015), Essen, Germany (Wagner and Schäfer, 2015), and Paris, France (Pal and Haeffelin, 2015), or rural sites (Pal et al., 2014). Some of these studies also investigated correlations with gaseous pollutants~~.~~, e.g. Czader et al. (2013) for Houston, Texas. Significant negative correlations between surface NO concentration and MLH were reported for Beijing (Schäfer et al., 2012). However, surface NO_2 concentrations are only weakly affected by the MLH as they are mainly secondarily formed through atmospheric processes. For Paris Dieudonné et al. (2013) investigated the relationship between surface concentrations of NO_2, column amount of NO_2 and the MLH. Their results suggest that the discrepancies between NO_2 surface concentrations and column amount can be explained by the differences in the MLH. For seven cities in the North China Plain an anti-correlation between near-surface O_3 and ~~H_{mt}~~ MLH was found (Hu et al., 2014) however, this case study was confined to only one day. Wagner and Schäfer (2015) investigated conditions near a major traffic road in Essen, Germany, and found that correlations between several constituents and MLH are significantly negative, if the MLH from the ceilometer measurements is grouped into ~~classes~~ intervals of 200 m~~width~~.

Currently such investigations cannot ultimately demonstrate which correlations between surface concentrations and atmospheric stratification exist, how robust they are and how large their range of applicability is. One prerequisite for progress is a critical review of standard methods for the determination of the MLH. Then, the dependence of such correlations on season, meteorological conditions, or location can be investigated. ~~Moreover, mixing layer schemes in numerical weather forecast and chemistry transport models can be validated; a mandatory requirement as the MLH is not a prognostic variable, but estimated from diagnostic relations. If validation is successful, modelled MLHs can be beneficial, e.g. early warnings can be improved if episodes of heavy pollution are expected.~~

3 The ~~measurement~~ BLUME network and the BAERLIN2014 campaign

Berlin is the capital of Germany with about 3,500,000 citizens. The terrain is flat with altitude differences of not more than 25 m except some small hills of up to about 85 m. A considerable part (about 40%) of the area of Berlin is covered by forests, agricultural areas, lakes and rivers. Similar to many other metropolitan areas Berlin suffers from episodes of poor air quality, in particular when particulate matter (PM_{10}) and NO_2 concentrations exceed the EU limit values. Thus, measures have been implemented such as restrictions of the traffic in the city center. Air pollution in Berlin originates not only from anthropogenic emissions of urban sources, but also from long range transport of particulate matter from industrialized areas in Poland, biogenic emissions and formation of secondary ~~material~~aerosol compounds; their relative contributions are not yet agreed upon in detail (Bonn et al., 2016).

Routine measurements of the air quality of Berlin are conducted at 16 automated stations of the so called BLUME-network (SenStadtUm (2014), see Fig. 1) under the responsibility of the Senate of

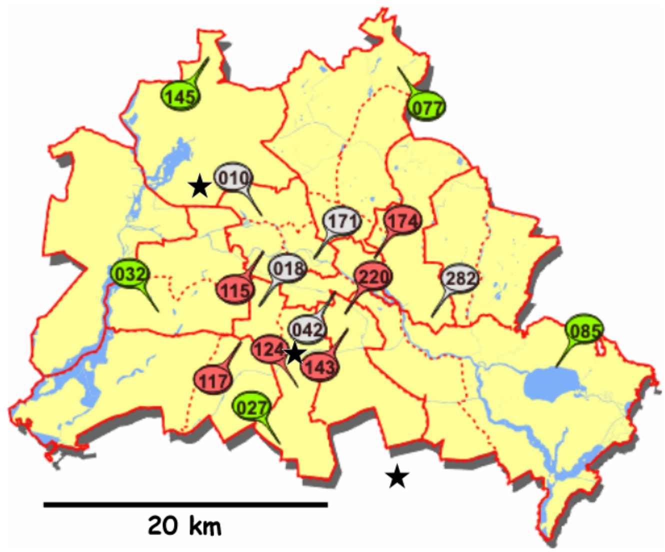

Figure 1. The 16 automated air quality station of the BLUME network. Ceilometer measurements reported in this paper were conducted at station #42, one of the urban background sites (in grey). Traffic stations are shown in red. The green flags are stations at the outskirts of Berlin and in forests (rural background). More details are given in Table 1. The three black stars indicate the wind measurements in Tegel, Tempelhof and Schönefeld (from northwest to southeast) mentioned in Sect. 5.1. Adaptation of a figure of the Senate of Berlin.

Berlin by European law. Their main purpose is the monitoring of surface concentrations of trace gases and particle mass concentration. For this study ~~data with a time resolution of 1 hour~~ hourly data are available. BLUME distinguishes three categories of stations: five of the stations are located at residential ~~areas~~ districts (labeled "urban background", grey flags), five at the outskirts of Berlin and forest areas ("rural background", green flags), and six at traffic hot spots (red flags). These data are reported to the Federal Environment Agency (UBA) of Germany and included into the European air quality database (AIR BASE). ~~More than 650 similar stations exist in whole Germany.~~

A summary of the automatic stations of the BLUME network and the monitored quantities used in this study are given in Table 1. Particulate matter is measured with the automatic PMI (particulate monitoring instrument), type FH 62 I-R (Thermo ESM Anderson), one of the standard systems for Germany's air quality network. It is based on attenuation of beta-radiation from a Krypton gas cell. It performs real time measurement of the suspended particulate matter on a filter. For the gaseous species discussed in this study Horiba's air pollution monitors (370-series) are deployed, i.e. an APNA-370 (for NO, NO_2 and NO_x, chemiluminescence method) and an APOA-370 (for ozone, absorption in the UV spectral range).

Table 1. Automatic stations of the BLUME network of Berlin: coordinates Names of the locations with the corresponding district are given in brackets. Coordinates are given as latitude (degree North) and longitude (degree East), d_{42} is the distance [in km] from station #42 (Nansenstraße). Listed are only measurements of pollutants discussed in this paper.

ID	location	coordinates	d_{42}	pollutants		
		outskirts (rural background)				
27	Schichauweg (Marienfelde)	52.3984°, 13.3681°	11.0		NO_x	O_3
32	Jagen (Grunewald)	52.4732°, 13.2251°	14.0	PM_{10}	NO_x	O_3
77	Wiltbergstr. (Buch)	52.6435°, 13.4895°	17.6	PM_{10}	NO_x	O_3
85	Müggelseedamm (Friedrichshagen)	52.4477°, 13.6471°	15.4	PM_{10}	NO_x	O_3
145	Jägerstieg 1 (Frohnau)	52.6533°, 13.2961°	20.3		NO_x	O_3
		urban background				
10	Amrumer Str. (Wedding)	52.5430°, 13.3491°	8.2	PM_{10}	NO_x	O_3
18	Belziger Str. (Schöneberg)	52.4858°, 13.3488°	5.6		NO_x	
42	Nansenstr. (Neukölln)	52.4894°, 13.4309°	0	PM_{10}	NO_x	O_3
171	Brückenstr. (Mitte)	52.5136°, 13.4188°	2.8	PM_{10}	NO_x	
282	Rheingoldstr. (Karlshorst)	52.4853°, 13.5295°	6.7		NO_x	
		traffic				
115	Hardenbergplatz (Charlottenburg)	52.5066°, 13.3330°	6.9		NO_x	
117	Schildhornstr. (Steglitz)	52.4636°, 13.3183°	8.2	PM_{10}	NO_x	
124	Mariendorfer Damm (Mariendorf)	52.4381°, 13.3877°	6.4	PM_{10}	NO_x	
143	Silbersteinstr. (Neukölln)	52.4675°, 13.4417°	2.5	PM_{10}	NO_x	
174	Frankfurter Allee (Friedrichshain)	52.5141°, 13.4699°	3.8	PM_{10}	NO_x	
220	Karl-Marx-Str. (Neukölln)	52.4817°, 13.4340°	0.9	PM_{10}	NO_x	

During summer 2014 a dedicated field campaign, BAERLIN2014 (Berlin Air quality and Ecosystem Research: Local and long-range Impact of anthropogenic and Natural hydrocarbons) was set up for three months (from 2. June until 29. August 2014) deploying several additional measurements from mobile and airborne platforms (Bonn et al., 2016; ?) . At focusing on ozone, secondary organic aerosols and the effect of urban vegetation (Bonn et al., 2016) . One Vaisala ceilometer was available at that time. It was installed at the BLUME station #42 (Nansenstraße, at the corner of Framstraße, 52.4894° N, 13.4309° E, see Fig. 1) a Vaisala ceilometer was installed on the roof of a children care take house (5 m above street level). This station is located in a residential neighborhood with trees and bushes. It is categorized as an "urban backgroundsite" site: in 2014 annual averages were 27 μg/m^3 and 21 μg/m^3 of PM_{10} and $PM_{2.5}$, respectively, 41 μg/m^3 of O_3, and 37 μg/m^3 of NO_x. The PM_{10}-threshold (daily average of

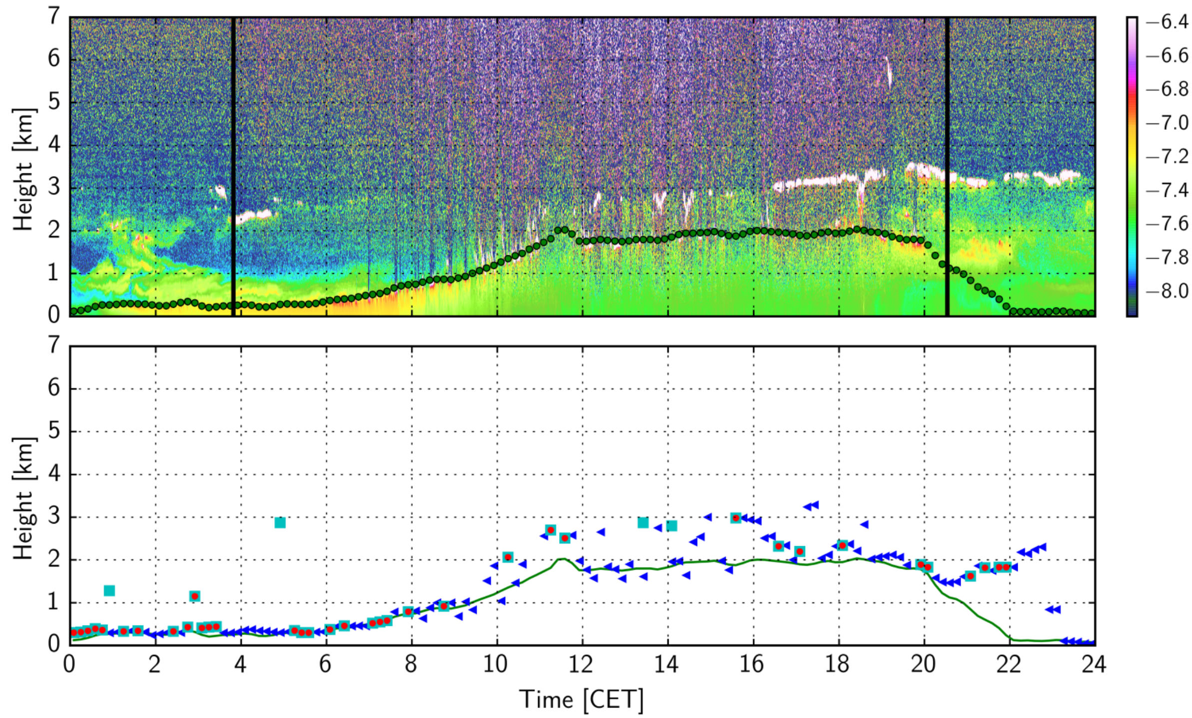

Figure 2. Top: Time-height cross section of range corrected ceilometer signals (Vaisala CL51) at the BLUME network site #42, Nansenstraße, on 1. July 2014 (in arbitrary units). The MLH as determined from COBOLT ($H_{ml,c}$) is marked by dark green dots. Bottom: Comparison of the MLH retrievals: $H_{ml,c}$ as above (green line), $H_{ml,v}$ with L1-criterion (blue triangles), the L3-criterion (red dots), and the Q3-criterion (cyan squares); for the definition see Table 2.

50 μg/m^3) was exceeded on 28 days which is below the limit of 35 days according to EU-regulations (http://ec.europa.eu/environment/air/quality/standards.htm).

The ~~main~~ objective of the ceilometer measurements ~~is~~ was the determination of the MLH and thus the option to combine ~~ground based and vertical information~~in-situ measurements at the surface with data concerning the vertical direction. Based on previous case studies for Munich (Wiegner et al., 2006) ~~it is assumed~~ and Paris (Pal et al., 2012) , long term observations for the region Munich/Augsburg/Freising (Geiß, 2016) and for Vienna (Lotteraner and Piringer, 2016) we assume that the so derived MLH is representative for ~~a metropolitan area.~~ Berlin. As can be seen from Table 1 all sites are within 20 km distance from the ceilometer with five stations being very close (less than 6.4 km).

Note, that all times are given in CET (central European time)~~, which is one hour different to UTC~~.

4 Mixing layer heights from ceilometer measurements

4.1 Ceilometer data

In the framework of BAERLIN2014 a Vaisala CL51 ceilometer (Münkel, 2007) was deployed. The instrument is fully automated and eye-safe. It provides backscatter signals at 910 nm. As this wavelength is influenced by water vapor absorption it is complicated to derive optical properties of particles in a quantitative way (Wiegner and Gasteiger, 2015), however the identification of aerosol layers is not affected as strong changes of the aerosol backscatter are not masked by the water vapor absorption. The height range of the measurements is more than 4 km, thus covering the typical range of MLH over a continental site like Berlin. Due to its optical design using the same lens for the emitter and the receiver optical paths, the minimum range is on the order of 50 m for the detection of aerosol layers and even lower for clouds. The spatial and temporal resolution is 10 m and 16 s, respectively. Ceilometer data (firmware version V1.032) are available for 67 days between 27. June and 2. September 2014 (except 15. July). The output signals are range corrected consistently for the whole measurement range, i.e., the "H2on"-parameter was set to 1 as discussed by Kotthaus et al. (2016). To improve the detection of aerosol layers close to the ground, ~~and~~ an additional overlap correction function, ~~based on the~~ similar to a concept outlined by Hervo et al. (2016), was applied.

4.2 Determination of the MLH

Virtually all retrievals of the MLH from ceilometer measurements are based on the shape of the range corrected signal (i.e. uncalibrated) or the vertical profile of the attenuated backscatter coefficient (i.e. calibrated, Wiegner and Geiß (2012)). Several methods to analyze the gradient of the profile or its temporal variability are available, different thresholds can be selected to distinguish between clouds and aerosol layers, and different temporal and vertical averaging can be applied to reduce the influence of noise.

The standard procedure for the MLH determination from Vaisala-ceilometers ($H_{ml,v}$) is the MATLAB-based software package BL-VIEW developed by the manufacturer. It provides up to three altitudes of aerosol layers (referred to as candidate levels in the following); they are counted upward, i.e., candidate level #1 is closest to the ground. They are determined from local minima of the gradient of the backscatter profile considering data of a 14 minutes-time period prior to the actual measurement; in case of low signal-to-noise ratios this time span is extended to 20 minutes. To improve the retrieval, signals are smoothed along the line of sight, thresholds are defined to identify cloud "contamination", and unrealistic outliers are deleted. In case of rain, no $H_{ml,v}$ is provided. Each candidate level is given with a quality flag based on the absolute value of the gradient and the "width" of the local minimum (Münkel et al., 2011). Quality flags are 1, 2 or 3, with 3 meaning the highest reliability. Candidate levels

Table 2. Overview over different approaches to determine MLH from BL-VIEW: the conditions with respect to the quality flag and the number of the candidate level

acronym	selected candidate level	quality flag
L1	#1	1, 2 or 3
L2	#1	2 or 3
L3	#1	3
Q3	lowermost of #1, #2 or #3	3

with quality flag 3 are not necessarily given for all times. This information is stored in an ASCII-file, and it is left to the user to find their own criteria to determine the MLH, i.e., different selection of the candidate levels is possible, and the quality flags might be considered or not. The advantage of the provision of three candidate levels is that different layers can be detected at the same time (e.g., stable layer, convective layer, residual layer), the disadvantage is that the attribution of the layers is more complicated (Schween et al., 2014). The details of BL-VIEW are not disclosed to the user.

In this paper we use different criteria. To facilitate further reading we introduce the acronym "L1" for the criterion "lowest candidate level if it has a quality flag of at least 1" (this is identical to the condition "lowest candidate level without considering the quality flag"'). "L2" and "L3" are defined accordingly. So in all cases the lowest candidate level (#1) is chosen if the quality flag fulfils the corresponding conditions, otherwise no MLH is retrieved. "Q3" stands for the criterion "lowermost candidate level with quality flag 3", meaning that any candidate level is chosen as long as it has the best quality flag. If more than one candidate level fulfils this quality criterion, the lowermost is selected. For reasons of clarity our nomenclature is summarized in Table 2. It is obvious that L1 is more often fulfilled than L3, and that any successful retrieval according to L3 is also a successful Q3-retrieval.

An alternative approach to determine the MLH ($H_{ml,c}$) has been developed by Geiß (2016), referred to as COBOLT ("continuous boundary layer tracing"). ~~With the code written in Python it~~ The code is written in the open source programming language Python and can be used ~~cross-platform~~on Windows and Linux platforms. The algorithm is based on a ~~parameter that~~ time and height dependent function $A(t,z)$ that has been defined according to Eq. 1:

$$A(t,z) = \frac{\epsilon_g\, M_g(t,z)}{99^{\text{th}}(\epsilon_g\, M_g(t,z))} + \frac{\epsilon_v\, M_v(t,z)}{99^{\text{th}}(\epsilon_v\, M_v(t,z))} \qquad (1)$$

It depends on the magnitude and orientation of gradients of the range corrected ceilometer signal (first term on the right hand side), and on the temporal variability of the aerosol layering (second term). Both terms are weighted according to ϵ_g and ϵ_v, respectively, and are normalized by the 99. percentile

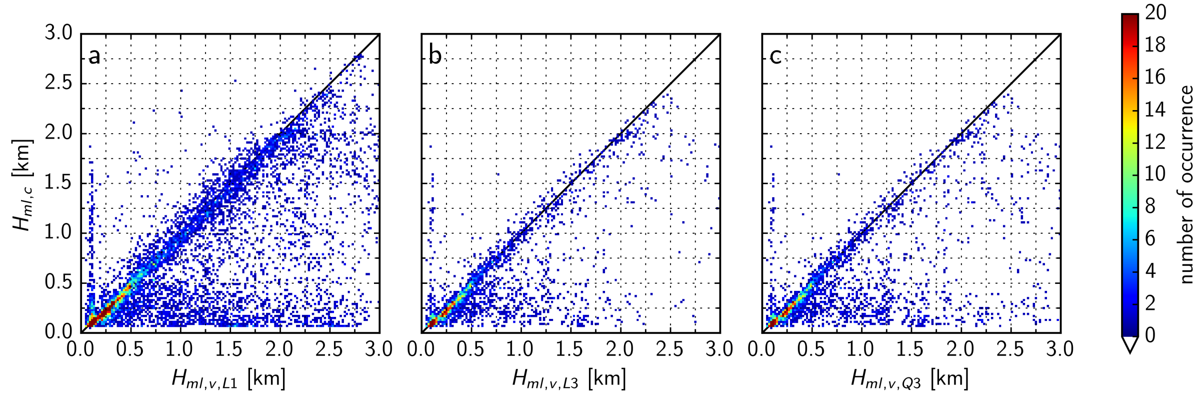

Figure 3. Comparison of MLH [in km] retrieved by COBOLT ($H_{ml,c}$) and different BL-VIEW approaches during the BAERLIN2014 campaign. (a): $H_{ml,v,L1}$ from L1-criterion, (b): ($H_{ml,v,L3}$ from L3-criterion, and (c): $H_{ml,v,Q3}$ from Q3-criterion. The number of occurrence is color coded.

of the function. By applying the Sobel operator (Duda and Hart, 1973) , in principle a two-dimensional gradient method, to X_Λ,

$$X_\Lambda(t,z,a,b) = \frac{1}{a}\int_{z_0}^{z_{max}} X(t,z)\,\Lambda\left(\frac{z-b}{a}\right)dz, \qquad (2)$$

with $X(t,z)$ as range corrected ceilometer signal and a low-pass filter $\Lambda\left(\frac{z-b}{a}\right)$ defined as

$$\Lambda\left(\frac{z-b}{a}\right) = \begin{cases} \frac{a}{2} - z + b & \text{if} \quad b - \frac{a}{2} \le z \le b \\ \frac{a}{2} + z - b & \text{if} \quad b \le z \le b + \frac{a}{2} \\ 0 & \text{elsewhere.} \end{cases} \qquad (3)$$

the function $M_g(t,z)$ and the direction of the gradients $\Theta(t,z)$ are obtained. The application of the Sobel operators to a low pass filtered ceilometer signal is equivalent to the wavelet covariance transform method using a Haar wavelet (Comeron et al., 2013) . Parameters a and on the magnitude and orientation of gradients of the b in Eq. 3 are the wavelet dilation and translation, respectively. The advantage of the Sobel operator is that both temporal and spatial changes can be evaluated simultaneously. The weighting function $\epsilon_g(t,z)$ is defined such that MLH that are unlikely in a meteorological sense are suppressed:

$$\epsilon_g(t,z) = \begin{cases} 0.1 & \text{if} \quad 0° \leq \Theta \leq 185° \\ 0.1 & \text{if} \quad 355° \leq \Theta \leq 360° \\ 1 & \text{elsewhere} \end{cases} \tag{4}$$

With this definition e.g. range corrected ceilometer ~~signal. This parameter has been~~ signals that increase with height ($\Theta \approx 90°$) – and most likely do not represent the top of the mixing layer – have a very low weight. In contrast, negative gradients caused by decreasing aerosol backscattering with height ($\Theta \approx 270°$) are emphasized. $M_v(t,z)$ is the temporal variance of $X_\Delta(t,z)$ and the weighting factor $\epsilon_v(t,z)$ is height-dependent in order to account for the decreasing signal to noise ratio with height. Specific gradient angles are excluded:

$$\epsilon_v(t,z) = \begin{cases} 0 & \text{if} \quad -5° \leq \Theta \leq 5° \\ 0 & \text{if} \quad 175° \leq \Theta \leq 185° \\ 1 - \frac{z}{3\,\text{km}} & \text{elsewhere, } z \text{ in km} \end{cases} \tag{5}$$

The function $A(t,z)$ was defined to especially determine the height of the convective boundary layer. ~~It has a maximum at the MLH~~ The empirical weights ϵ_g and ϵ_v had undergone extensive testing to find solutions that provide a reliable identification of the top of the mixing layer from the maximum of $A(t,z)$. For the determination of the diurnal cycle of the MLH, the maximum of $A(t,z)$ is traced in time. For the initialization of the time-height tracking procedure $H_{ml,c}$ at a starting time ~~and starting height are necessary. They are~~ t_0 is required. It is determined between 2.5 ~~h~~ hours and 3.5 ~~h~~ hours after sunrise, when the ~~MLH~~ convective mixing layer is assumed to be existent (Wildmann, 2015). Relying on the variance method which is especially sensitive to the beginning convection (Menut et al., 1999), the height of the maximum value of ~~the parameter~~ $A(t,z)$ is chosen as the initial ~~MLH~~$H_{ml,c}(t_0)$. Starting with ~~this initial MLH~~ $H_{ml,c}(t_0)$ a search window with a vertical extent depending on the solar zenith angle is moved backward in time to cover the period before sunrise, and forward until sunset. In case of rain ~~no MLH is determined~~$H_{ml,c}$ remains unchanged but is flagged; consequently, observations during (long lasting) precipitation events can be excluded by the user if desired. In the presence of convective clouds at the top of the boundary layer, the strongest decrease of the signal in the cloud is used to determine ~~the MLH~~$H_{ml,c}$, which is usually a few tens of meters above the cloud bottom. The analogue procedure as for the convective daytime MLH is applied after sunset for the nocturnal stable boundary layer. To account for the transition from decaying thermals in the well developed mixing layer to the establishment of a stable boundary layer a linear change of the MLH $H_{ml,c}$ between both layers is assumed to take place between 30 minutes before until 60 minutes after sunset (Grant, 1997; Grimsdell and Angevine, 2002).

In COBOLT an ensemble of 40 potential tracks $H_{ml,c}(t)$ is calculated with different initial conditions and search criteria. The track with the minimum length , e.g. different widths of the search window. The selection of the final result is performed by means of the function C_j for each ensemble member j (≤ 40) as defined in Eq. 6

$$C_j = \frac{\sum_{i=0}^{N-1} \sqrt{(t_{i+1} - t_i)^2 + (H_{ml,c}(t_{i+1}) - H_{ml,c}(t_i))^2}}{\sum_{i=0}^{N} A(t_i, H_{ml,c})} \tag{6}$$

with N being the number of time steps t_i within one day, i.e. N=144 for COBOLT's temporal resolution of 10 minutes. The track j with the minimum value of C_j is selected as the final result: the main idea behind this selection is that the MLH is assumed to develop smoothly in time, i.e., sudden "jumps" (that would increase the length of the track) do not occur in reality but are caused by wrong attribution of the mixing layer in case of multi-layered aerosol distributions. As a consequence, COBOLT retrievals do not have any temporal gaps, and unrealistic growth rates of the MLH $H_{ml,c}$ are suppressed. Otherwise, in particular in case of the detection of e.g. two layers it might happen, that the retrieved MLH $H_{ml,c}$ "switches" between the two those layers resulting in very strong and rapid changes. COBOLT output is typically provided for intervals of 10 minutes.

To make both approaches better comparable, time is assigned to the center of the interval of the BL-VIEW retrieval. Note, that a perfect temporal co-incidence is not possible because of the inherent properties of both algorithms, e.g. a the height-dependent temporal averaging in case of COBOLT.

4.3 Comparison of MLH retrievals

A typical example of CL51 measurements and the MLH retrieval is shown in Fig. 2. The attenuated backscatter signal (color-coded, in arbitrary units) up to 7 km above ground is shown in the upper panel for 1. July 2014. Sunrise at 03:46 CET and sunset at 20:32 CET are highlighted by the black lines. Visual inspection shows broken cloud fields from 09:00 CET to 16:00 CET at different altitudes, afterwards an almost continuous cloud deck at 3 km, and inhomogeneous aerosol layers up to 2.0 km before sunrise and up to 3.0 km after sunset. The MLH as identified by COBOLT ($H_{ml,c}$) is marked by dark green dots.

The results of all MLH retrievals are shown in a separate panel for reasons of clarity (Fig. 2 bottom): BL-VIEW ($H_{ml,v}$) with different selection criteria L1, L3 and Q3 are shown as blue triangles, red dots, and cyan squares, respectively, whereas $H_{ml,c}$ is shown as green line (same as in the upper panel). The temporal interval is 10 minutes. It can be seen that the overall agreement between COBOLT and BL-VIEW L3 is very good and coincides with what a human observer would have analyzed. Note that in general cloud bottoms were not misinterpreted as MLH by either approach. For L2 (not shown here) and L1 more cases of wrong assignments occur. Disagreements between COBOLT and L3 are rare, mainly between 20:00 CET and 22:00 CET when $H_{ml,v}$ is significantly higher – here the residual layer seems to be interpreted as the mixing layer by the Vaisala retrieval. Disagreements are more frequent when L1 or L2 is applied instead of L3, e.g. around noon, when BL-VIEW L1 selects the top of elevated aerosol layers and occasionally clouds as the MLH, or after sunset, when L1 selects the residual layer. It is obvious, that $H_{ml,v}$ is often not available during the daylight period, especially when L3 is considered. The main reason is the high temporal variability of the distribution of aerosols aerosol particles and clouds, e.g. under not well-mixed conditions with more than one aerosol layer that prohibit an unambiguous determination of $H_{ml,v}$. Consequently, candidate levels are rapidly changing, leading to lower quality flags (Münkel et al., 2011) and a failure of the MLH assessment. So it can be understood that the temporal coverage of $H_{ml,v}$ is quite low if L3 is applied. Figure 2 confirms that even the application of L1 (and L2, not shown here) does not fill all temporal gaps. As all MLHs from L3 are by definition also fulfilling the Q3-criterion, these results do not differ much. Only very few cases are added, e.g. before sunset, when the top of the residual layer was identified as the second or third candidate level and flagged with the highest quality.

These conclusions also hold for the whole period of BAERLIN2014. The intercomparison of the different MLH retrievals is summarized in Fig. 3. Figure 3a concerns BL-VIEW when the "weak" constraint L1 is applied: for each $\{H_{ml,v,L1}, H_{ml,c}\}$-pair the number of occurrence is color-coded. As expected from the example shown in Fig. 2, many cases with $H_{ml,v,L1} > H_{ml,c}$ exist. This is a consequence of multiple aerosol layers and the different behavior of the algorithms in the presence of a residual layer. The correlation coefficient according to Pearson is $R = 0.653$. The corresponding comparison if the stronger constraint L3 is applied is shown in Fig. 3b. Here, the correlation is obviously better with $R = 0.754$. Again, the number of cases with $H_{ml,v,L3} > H_{ml,c}$ is much larger than the opposite case, but less frequent than before (Fig. 3a). Similar results are found when Q3 is applied (Fig. 3c). It is the same distribution as the L3-case however with some additional cases when the lowest candidate level has a low quality flag, whereas one of the upper levels is considered as quite reliable. Consequently, the additional points concern primarily large $H_{ml,v}$ and the correlation is lower than before ($R = 0.650$). It is clear, that the application of more rigorous criteria leads to a drastic reduction of successful H_{mt} $H_{ml,v}$-retrievals: with the L1-criterion the total number is 8346, whereas it is only 2998 and 3331 for L3 and Q3, respectively. Note, that the largest possible number of MLH retrievals would be 9648 (67 times $\times$ 24 times $\times$ 6).

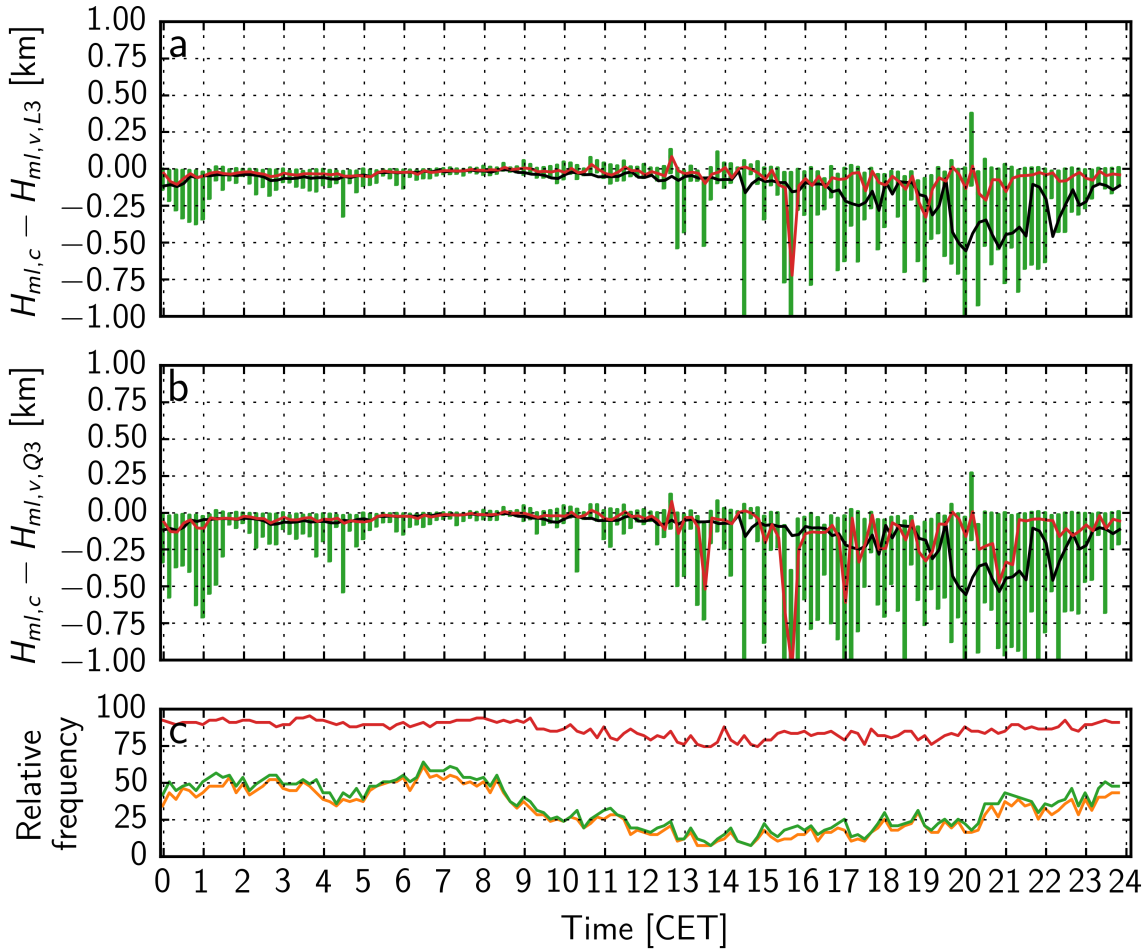

Figure 4. a: Difference ΔH of the retrieved MLH from COBOLT and BL-VIEW L3 during the BAERLIN2014 campaign: vertical lines indicate the interval from the 25th to the 75th percentile. The red line is the median of the distribution. For comparison the corresponding median from the L1-criterion is shown (black line). b: same as top panel but ΔH of the retrieved MLH from COBOLT and BL-VIEW Q3. c: relative frequency of successful $H_{ml,v}$-retrievals (L3 in orange, Q3 in green, L1 in red) in percent in relation to the COBOLT-retrieval.

To better understand the reasons for the discrepancies between the approaches the difference (ΔH)

$$\Delta H(t) = H_{ml,c}(t) - H_{ml,v}(t) \tag{7}$$

for each 10 minutes interval is calculated. Figure 4a concerns the L3-criterion. Green bars show the range between the 25th and 75th percentiles of ΔH at a given time. The red line illustrates the median value. For comparison the corresponding median of the L1-approach (black line) is also shown. It is obvious that the median is very small for both BL-VIEW approaches and stays between $+0.03$ and -0.11 km before noon. Between 16:00 CET and 23:00 CET ΔH is clearly shifted to negative values with a median reaching -0.33 km and -0.56 km for L3 and L1, respectively. This is a clear indication that with the establishment of the residual layer in the late afternoon and after sunset, the BL-VIEW algorithm tends to select the top of the residual layer as $H_{ml,v}$, especially if the user selects the L1-criterion. A similar effect is found in cases of complex aerosol particle distributions with several layers. L3 gives a much better agreement with COBOLT, however, as already ~~briefly~~ mentioned, the stricter L3-criterion leads to considerable temporal gaps in the $H_{ml,v}$-retrieval: in Fig. 4c it can be seen (orange line) that the relative number of 10 minutes intervals that allows to determine $H_{ml,v}$ is never larger than 61 %. Between 10:00 CET and 20:00 CET it is typically only in the 15–25 %-range because in the majority of cases the lowest candidate level does not have the highest quality flag (see Fig. 2). The low number of successful retrievals is also the reason for the rare cases (e.g., at 15:40 CET) when the absolute value of ΔH for L3 is larger than for L1. If the weaker L1-criterion is applied the availability of $H_{ml,v}$ is significantly increased (see the red line) and reaches a relative frequency of successful retrievals of more than 75 % throughout the day, however, at the expense of a in general good agreement between $H_{ml,v}$ and $H_{ml,c}$.

The corresponding comparison for the Q3-criterion is shown in Fig. 4b. The findings are similar ~~to~~ as before, however, the range of differences $\Delta H(t)$ is extended towards larger negative values (green lines) as expected. This concerns the whole diurnal cycle but the effect is strongest after sunset. The number of successful $H_{ml,v}$-retrievals is slightly larger than for the L3-criterion as can be seen in the lower panel (green line).

If we compare – as a consequence of these findings – only MLH retrievals before sunset the agreement between BL-VIEW and COBOLT is indeed improved. If the L1-criterion is applied to the complete diurnal cycle 23.4 % of the intercomparisons show large negative differences ($\Delta H < -0.5$ km). If only measurements before sunset are considered the number is reduced to 19.1 %. The corresponding numbers for the Q3-criterion are 20.2 % and 17.3 %, respectively. For the L3-criterion we find 12.3 % and 9.5 %. Retrievals when $H_{ml,c}$ is larger than $H_{ml,v}$ are quite rare. A difference ΔH of more than 0.5 km occurs in less than 1.5 % of the cases for all three BL-VIEW approaches.

Figure 5 shows the mean diurnal cycle of ~~H_{ml}~~ MLH from 67 days as retrieved by BL-VIEW L3 and COBOLT. The dark blue line corresponds to $H_{ml,c}$ whereas the orange line is for $H_{ml,v}$. The mean maximum vertical extent is approximately 1.5 km, similar to results from Lotteraner and Piringer (2016) found for Vienna. The light blue lines indicate the temporal variability as calculated from the standard deviation σ_c (COBOLT approach). It is in the order of 100 m before sunrise and up to 500-700 m in the afternoon. Though this finding is based on COBOLT that provides complete temporal coverage it remains open whether this is representative for summer months in Berlin. Similar values but less variability were found for Barcelona, Spain (Sicard et al., 2006). From summer observation over five years in Paris, France, Pal and Haeffelin (2015) found larger values (H_{ml}=1.95 $\pm$ 0.38 km), whereas maxima less than 0.8 km were observed during two years at Vancouver, BC, Canada (van der Kamp and McKendry, 2010) and Santiago, Chile (Munoz and Undurraga, 2010). The mean $H_{ml,c}$ at night is in the range of 0.2 km underlining the need of ceilometers with a very low overlap (or a reliable overlap correction function, see e.g. Hervo et al. (2016) for a CHM15k-ceilometer) for investigations of the mixing layer. The most prominent differences between BL-VIEW and COBOLT are the larger $H_{ml,v}$ during night, and the rapid changes of $H_{ml,v}$ around noon. The main reason for these "fluctuations" is the low number of retrievals when L3 is applied, e.g. for some of the 10 minutes intervals only in 5 out of 67 days $H_{ml,v}$ could be found. Thus, the significance is limited, nevertheless $H_{ml,v}$ is within the range of $H_{ml,c} \pm \sigma_c$.

The green line in Fig. 5 shows the first derivative of the COBOLT retrieval $H_{\mathrm{ml,c}}$. This quantity can be relevant in view of temporal averaging, e.g., when MLH is correlated with concentration measurements ~~with~~ having a lower temporal resolution. This topic is briefly discussed in the next section.

It is worthwhile to also determine a typical afternoon value of MLH. ~~On the one hand it describes the~~ Figure 5 confirms that this period provides the maximum volume for the mixing of emitted compounds, ~~on the other hand it is assumed to be~~ and that the MLH is representative for several hours~~(see Fig. 5). The latter is one of the reasons for the measurement schedule of EARLINET~~. The latter has been the reason for including a measurement around two hours after local noon in the regular EARLINET schedule (Pappalardo et al., 2014). Based on the mean diurnal cycle we define this value as the average over the 3-hour time period starting 30 minutes after noon. Figure 6 shows the results from COBOLT (blue dots) and L3 (orange dots) for the whole period of 67 days. Note, that BL-VIEW with the strict L3-criterion fails to determine $H_{ml,v}$ in 21 days (shaded areas) for the reasons mentioned above. If both values are available the general agreement is however good, only few cases exist when $H_{ml,v}$ is much larger than the respective COBOLT-result $H_{ml,c}$ (e.g., 27. June, 1. July, and 10. July).

We conclude that the main discrepancies between COBOLT and BL-VIEW origin from the presence of the residual layer and elevated aerosol layers during day time whereas broken cloud fields cause less

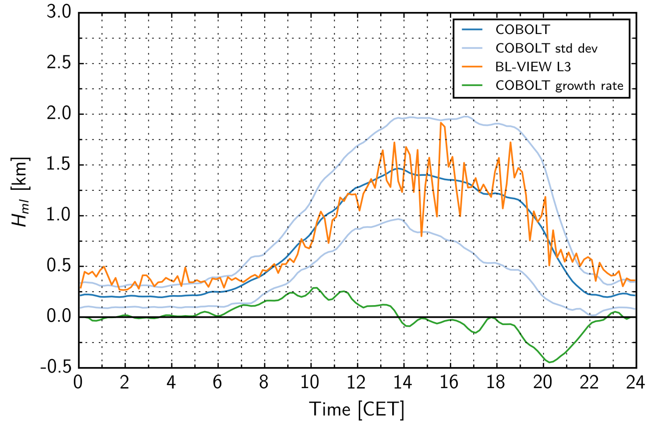

Figure 5. Mean diurnal cycle of $H_{ml,c}$ (dark blue) and $H_{ml,c} \pm \sigma_c$ as retrieved with COBOLT, and $H_{ml,v}$ from BL-VIEW L3 (orange) at the urban background site #42, in 10 minutes resolution, averaged over 67 days. The green line shows the growth rate of $H_{ml,c}$ (in km/h).

problems. The main drawback of the present version of BL-VIEW is the limited temporal coverage, when only retrievals with the highest quality flag are considered.

4.4 Temporal averaging of the mixing layer height

When evaluating ceilometer data a temporal resolution of ~~H_{ml}-retrievals~~ MLH-retrievals of the order of 10 minutes can be achieved. This is typically better than the resolution of air quality measurements of automated monitoring stations. To make ~~H_{ml}-retrievals~~ MLH-retrievals comparable with the in-situ measurements of the BLUME-network, 1-hour averages have to be calculated. In this context the growth rate of the ~~MLH~~ mixing layer (dH_{ml}/dt) is relevant; it is shown for the mean diurnal cycle derived from COBOLT as a green line in Fig. 5. It can be seen that the mean $H_{ml,c}$ rises with 150 – 200 m per hour between 08:00 and 12:00 CET with a maximum of 290 m. This is in good agreement with other continental cities (e.g. Baars et al., 2008; Pal and Haeffelin, 2015). The mean diurnal cycle of $H_{ml,c}$ shows its strongest decrease after sunset, reaching rates of -450 m per hour. For individual days these rates can be exceeded.

~~As a consequence if~~ However, in case of L3 or Q3 the MLH cannot be retrieved for each 10 minutes interval (see low values in Fig. 4c)~~but 1-hour averages shall be determined~~. As a consequence, hourly averages of the MLH can be biased on the order of $\pm$ 100 m due to the rapid growth of the mixing layer during strong convection events before noon. After sunset the uncertainty can be even larger ($\pm$ 200 m).

Medians of the MLH are derived from all available 10-minutes retrievals (up to six, depending on the retrieval) of the corresponding hour, for all 67 days. So, up to 402 values are considered. The resulting

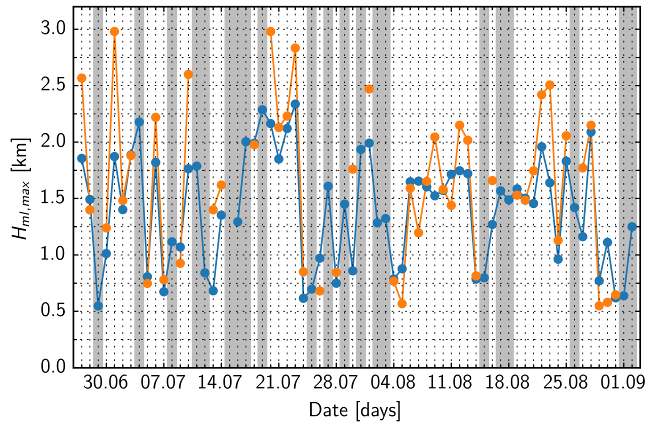

Figure 6. Daily afternoon value of H_{ml} (averaged over 3 hours starting 30 minutes after noon) as derived from COBOLT (blue dots) and BL-VIEW L3 (orange dots) between 27. June and 2. September 2014. The alignment of the labels of the x-axis (date) is defined by the position of the dots separating day and month. The shaded areas highlight days when $H_{ml,v}$ could not be retrieved applying the L3-criterion.

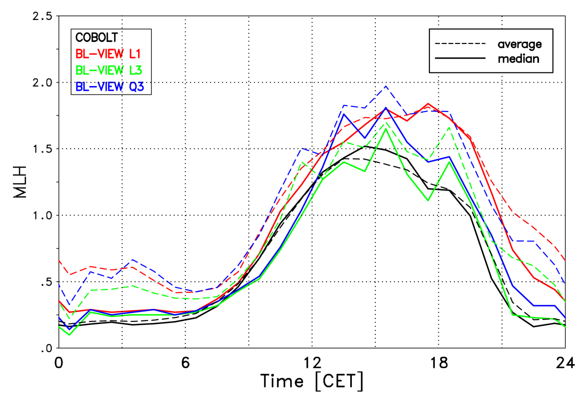

Figure 7. Diurnal cycle of hourly values of the MLH as determined from the different retrievals (see legend). Thick solid lines are for medians, thin broken lines for averages.

hourly values as they are used in the following discussion (Sect. 5) are shown in Fig. 7. In particular before sunrise averages are larger than the medians of MLHs. This is expected from Figs. 4a and 4b showing negative values of $\Delta H(t)$, i.e. there are cases of much larger MLH derived from the BL-VIEW retrievals.

5 Link to air quality

In the following we consider BLUME measurements of PM_{10} and concentrations of O_3 and NO_x. These measurements are available with a temporal resolution of 1 hour. For the MLH we use the arithmetic mean of up to six ~~H_{mt}-values~~ values from 10 ~~minute~~ minutes intervals.

~~With respect to particulate matter one of the standard systems for Germany's air quality network is the automatic PMI (particulate monitoring instrument) , type FH 62 I-R (Thermo ESM Anderson). It is based on attenuation of beta-radiation from a Krypton gas cell. It permits real time measurement of the particles on a filter and on-line measurement/display of the mass concentration of the suspended particulate matter. Inorganic species are continuously measured by Horiba's air pollution monitors (370-series). For the species discussed here the APNA-370 (for NO, NO_2 and NO_x, chemiluminescence method) and the APOA-370 (for ozone, absorption in the UV spectral range) are deployed. All measurements are performed under ambient conditions.~~

~~The mobile measurements from~~ Episodic mobile (bicycle) measurements during BAERLIN2014 ~~showed~~ have already shown that there is significant horizontal heterogeneity in ~~both~~ gas-phase pollutants ~~(O_3 and NO_x), as well as~~ and particle number concentrations (Bonn et al., 2016). ~~Additional measurements at the Nansenstraße site indicate that contributions from inorganic species contribute significantly to secondary aerosol and thereby PM concentrations, and that local emissions, both anthropogenic and biogenic, are relevant (?) . Here, we do not discuss these topics but rather focus on the importance of retrieving the MLH with an appropriate approach and discuss its influence~~ In the following we discuss the influence of different retrievals of the MLH on correlations with surface measurements of PM_{10} and concentrations of gases (O_3 and NO_x). Note, that in-situ measurements are available at different sites, whereas only one ceilometer was deployed, consequently an inherent assumption of the following discussion is, that the MLH is ~~similar over all parts of~~ the same over Berlin.

5.1 Correlation between MLH and PM_{10}

For the discussion of correlations between MLH and PM_{10} we can use measurements at the outskirts (#32, #77, #85)~~and~~ , at urban background stations (#10, #42, #171) and at five stations that are strongly influenced by traffic (#117, #124, #143, #174, #220, see Fig. 1 and Table 1). The diurnal cycles of PM_{10} (in μg/m^3) at ~~the~~ these eleven stations are shown in Fig. 8, calculated as medians of ~~up to 67 measurements for each hour of the day~~all measurements of the corresponding hour of each day of the measurement period (67 days). It can be seen that the concentrations at the traffic sites (solid lines) are in general slightly higher than at the urban background and ~~remote sites~~the outskirts. The amplitude of the mean diurnal cycle is quite small – between 4.4 μg/m^3 at #32 (red dotted line) and 10.6 μg/m^3 at #124 (green solid line) – whereas the day to day variations are comparably large at all sites and all times

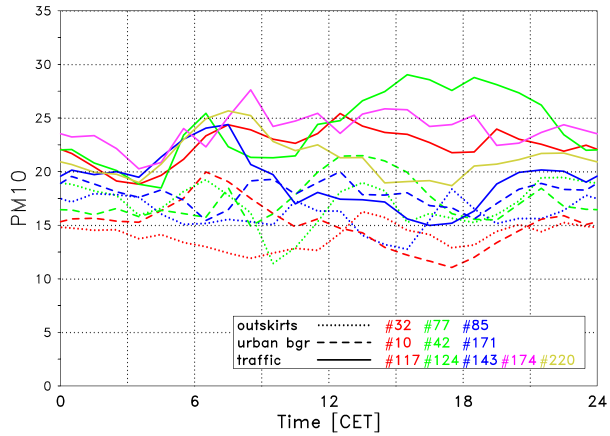

Figure 8. Diurnal cycle of PM_{10} in $\mu g/m^3$ at eleven BLUME stations, based on medians of measurements on 67 days. The temporal resolution is 1 hour. The locations at the outskirts of Berlin (dotted lines), urban background sites (dashed) and traffic sites (solid) are indicated in the legend; see also Table 1.

of a day. On the one hand the diurnal cycles have some common features; e.g., a distinct increase during the morning rush hours at all traffic sites and some of the urban background sites. This is plausible from vehicle emissions. At the urban background site #171 and sites at the outskirts, however, the strong increase occurs several hours later. The delay might be caused by the transport time from the main sources to the site. On the other hand ~~constantly~~ changing contributions of large scale transport from variable directions, local sources or particle removal by precipitation can lead to a quite different development in the course of a day including continuously increasing/decreasing PM_{10}, sporadic "peaks", or sudden drops at any time. The combination of these effects complicates a meteorological interpretation of mean diurnal cycles.

For the determination of the diurnal cycle of the MLH we have – as already mentioned in Sect. 4.2 – four different MLH retrievals available. For the correlations between the PM_{10}-measurements and the MLH retrievals, further options can be considered: either averages or medians of hourly values (67 or less) ~~.~~ as shown in Fig. 7 can be used. Figure 9 illustrates how these correlations depend on the site and the ~~properties of the~~ retrieval. Eleven blocks according to the eleven sites are separated and labelled following Table 1. For each site four different correlations are shown (from left to right): averages of MLH vs. averages of PM_{10} , medians of MLH vs. medians of PM_{10}, averages of MLH vs. median of PM_{10}, and median of MLH vs. averages of PM_{10}. The different colors indicate which ceilometer retrieval is used to determine the MLH: the COBOLT-approach is shown in black, and the BL-VIEW retrievals in red (L1-criterion), green (L3) and blue (Q3).

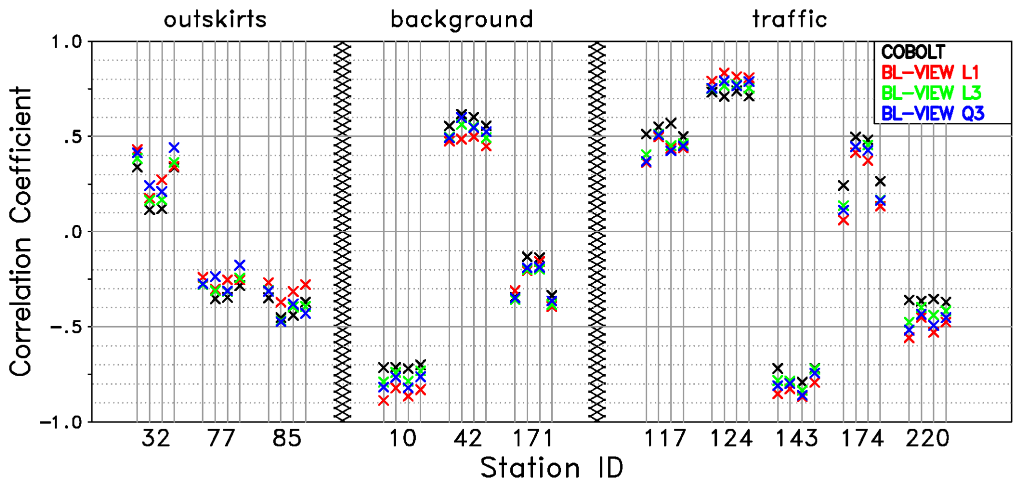

Figure 9. Correlation coefficient R for mean diurnal cycles of MLH and PM_{10} for eleven sites (from left to right): sites on the outskirts (#32, #77, and #85), urban background sites (#10, #42, and #171), and traffic sites (#117, #124, #143, #174, and #220). The results of the different retrievals are color coded as indicated in the legend. The four vertical lines of each block correspond to different options of correlation: MLH-average vs. PM_{10}-average (1), MLH-median vs. PM_{10}-median (2), MLH-average vs. PM_{10}-median (3), MLH-median vs. PM_{10}-average(4).

The wide range of correlation coefficients for the different locations is obvious: The strongest correlation between MLH and PM_{10} is found for the traffic site #124 ($R \approx 0.77$), the strongest anti-correlation for site #143 (traffic, $R \approx -0.79$) and site #10 (urban background, $R \approx -0.78$). So only for two sites a correlation is found as is expected if vertical dilution were the dominant process for the surface concentration of particulate matter. Compared to the large spatial heterogeneity the differences for different correlation options and MLH retrievals are with the exceptions of sites #32 (outskirts), #171 (urban background) and #174 (traffic) small: for a given MLH retrieval (i.e. same color) the range of R (maximum minus minimum) for different options is typically 0.08, for a given option (i.e. same vertical line) it is 0.11 on average. For the three sites mentioned the sensitivity to the correlation option is however 0.25–0.35. The reason is that correlations involving averages of PM_{10} (first and last vertical line of each block) clearly differ for those ~~involving~~ based on medians. The latter are less effected by short episodes of extreme concentrations which are not unusual for particulate matter.

As already mentioned these correlations are based on ceilometer measurements at one site, and that it was impossible in the framework of BAERLIN2014 to verify that the diurnal cycle of the MLH within the 20 km range of the air quality stations is identical. Large differences of the correlation coefficients are however also found if we restrict ourselves to the five BLUME-stations (#220, #143, #171, #174, and #124) that are closest to the ceilometer site ($0.9 \leq d_{42} \leq 6.4$ km, see Table 1). Over this small area in the center of the city changes of the diurnal cycle of the MLH are very unlikely. Nevertheless our previous conclusions are confirmed: as can be seen from Fig. 9 the correlations between MLH and PM_{10} are quite variable ranging from more than R=0.7 (#124, d_{42}=6.4 km) to less than $R = -0.7$ (#143, d_{42}=2.5 km).

At first glance it seems to be surprising that even within the same class catagory the correlations are quite different. The three stations at the margin of Berlin (outskirts) show however different characteristics with respect to their distance to major traffic sources. Station #32 (Grunewald) is only 0.8 km west of the AVUS-motorway, whereas stations #77 and #85 are more than 3.5 km from the next motorway. The latter station is close to a large lake. Thus there is in principle sufficient time for mixing during the transport from these sources towards the measurement site, of course depending on the wind direction that certainly changes during to observation period. The three urban background stations show even more pronounced differences. For station #10 the distance to the next main road is larger than for the other two sites, and due to the east-west orientation and the broad street ventilation is more effective than for the reference site #42 (Neukölln) with a lot of vegetation , a high density of buildingsin a typical residential neighborhood in the inner part of a big German city, and a comparably large distance to major roads. In contrast station #171 is close to a main road but it benefits from a good ventilation from the river Spree. For the traffic stations technical conditions, e.g., the number of lanes, the presence of traffic lights close to the monitoring station, and height and distance of the surrounding buildings becomes especially relevant because of the short distance between the emitters and the monitoring station. Consequently, the vertical dilution in the mixing layer is less relevant for PM_{10} concentrations, and correlations are rather governed by the diurnal cycle of the traffic which is not necessarily dominated only by the morning and evening rush hours, but could have a significant contribution from busses buses and trucks throughout the day. Thus, the quite different correlations shown in Fig. 9 are not unlikely and demonstrate the local differences of the relevance of mixing, transport, emission and aerosol formation (including the availability of precursor compounds).

We conclude that that the completely different correlations between mean diurnal cycles of MLH and PM_{10} at the different sites clearly demonstrates as shown in Fig. 9 clearly demonstrate that the surface concentration of particulate matter is not determined by the vertical stratification of the mixing layer alone, but also by local sources and sinks , the relative humidity and the wind field (see e.g. Tandon et al. (2010); Tai et al. (2010)). Moreover, the distance between the main sources and the measurement site is relevant. This is confirmed by absolute values of R below 0.15, if all

The lack of a unique correlation is confirmed if we consider sub sets of data with specific meteorological conditions. Two examples are briefly discussed: the consideration of the wind field and the differences of working days and weekends. If only days are considered when the average wind speed over Berlin was below a certain threshold a pronounced correlation is more likely because the

Table 3. Correlation coefficients R between medians of hourly MLH (derived from COBOLT) and PM_{10} for different sub sets of data: "all": diurnal cycles based on 67 days as shown in Fig. 9 (second vertical line of each block), "v40", "v30" and "v25": only consideration of days with average wind speed $\overline{v}$ below 4.0 m/s, 3.0 m/s, and 2.5 m/s, respectively, "m–f": Monday to Friday, "w-end": weekend only. The station IDs are according to Table 1.

station ID	all	v40	v30	v25	m–f	w-end
# 32	0.12	0.11	-0.05	-0.29	0.26	-0.46
# 77	-0.35	-0.41	-0.40	-0.44	-0.12	-0.80
# 85	-0.45	-0.50	-0.45	-0.42	-0.25	-0.74
# 10	-0.71	-0.77	-0.84	-0.83	-0.68	-0.76
# 42	0.62	0.54	0.20	-0.18	0.68	-0.14
#171	-0.13	-0.25	-0.62	-0.74	0.13	-0.64
#117	0.55	0.42	0.28	-0.10	0.52	-0.11
#124	0.71	0.74	0.67	0.52	0.76	0.19
#143	-0.80	-0.81	-0.80	-0.76	-0.63	-0.74
#174	0.50	0.35	-0.29	-0.47	0.58	-0.68
#220	-0.36	-0.42	-0.53	-0.55	-0.27	-0.53

vertical exchange can dominate advection. Hourly wind measurements in 10 m altitude were available at three stations in Berlin, i.e. Tegel (52.5644° N, 13.3088° E; d_{42} = 11.8 km), Tempelhof (52.4675° N, 13.4021° E, d_{42} = 3.1 km) and Schönefeld (52.3807° N, 13.5306° E, d_{42} = 13.9 km). They constitute a northwest to southeast transect through Berlin (see black stars in Fig. 1). For a simplified categorization of the wind field we use the daily averages of the wind speed $\overline{v}$. We found 52 (out of 67) days where $\overline{v}$ was below 4 m/s at all three stations, 28 days with $\overline{v}$ <3 m/s and only 16 days with $\overline{v} < 2.5$ m/s. In the latter case correlations between PM_{10} and $H_{ml,v,L3}$ or $H_{ml,v,Q3}$, respectively, suffer from the low number of successful retrievals (see Sect. 4.3). Therefore the correlation coefficients shown in Tab. 3 (columns "v40", "v30" and "v25", respectively) only refers to COBOLT-retrievals of the MLH. Though the correlation coefficients are in general more shifted to negative values compared to Fig. 9 (see also column labeled "all" in Table 3) and anti-correlations occur more frequently, the large spatial variability remains.

If we distinguish – as the second example – working days and weekends (columns "m–f" and "w-end" of Tab. 3, respectively), we find very pronounced differences with a tendency to stronger a anti-correlation for weekends. This is plausible as the diurnal cycle of the emissions is less pronounced. However, there were only ten weekends with ceilometer measurements during BAERLIN2014, so these findings should be treated as preliminary.

As an additional example one may focus on day/night differences of the correlation. For this purpose we use co-incident ~~1-hour measurements are correlated (~~hourly measurements (depending on the ceilometer retrieval up to 1608 ~~instead of 24 values, not shown here~~values) rather than the mean diurnal cycle as before to overcome the small number of samples. We define "day time" as the period between 07:01 CET and 20:00 CET, and "night time" as times before 07:00 CET and after 21:00 CET. Then, for day time measurements we get very low correlation coefficients $-0.33 < R < 0.10$, for night time the correlation is only slightly different ($-0.27 < R < -0.09$). The main result is that during night time $R < 0$ for all sites, and only one site with $\|R\| < 0.1$ was found. On the one hand these values are plausible as we expect an anti-correlation between MLH and PM_{10} in view of the suppressed vertical mixing in particular during night when the mixing layer is typically shallow (see Fig. 5). On the other hand the absolute values of R are too small for supporting a strict scientific interpretation.

We conclude that the heterogeneity of the city is obviously more relevant than the selection of the MLH retrieval and the correlation option. ~~It is demonstrated, that considering~~ The introduction of only three classes of monitoring stations (traffic, urban background, outskirts) cannot reflect the full complexity of pollution ~~exposure in a metropolitan area. As a consequence, the choice of a single representative measurement station for the entire city (and the determination of a single representative correlation coefficient) is impossible~~in the metropolitan area, and a re-assignment might be advisable when traffic flows have changed over years.

5.2 Correlation between MLH and gaseous pollutants

With respect to gaseous pollutants we restrict our discussion to O_3 and NO_x. Ozone measurements on a hourly basis are available at seven sites, NO_x at all 16 sites (see Table 1).

The mean diurnal cycle of O_3 is shown in Fig. 10 for the five stations located at the outskirts of Berlin (dotted lines) and two urban background sites (#10 and #42, solid lines); medians considering 67 days of data are plotted. It exhibits the typical pronounced diurnal cycle with a maximum of about 100 μg/m^3 between 14:00 and 16:00 CET. The minimum occurs shortly after sunrise which was between 04:00 and 05:00 CET during the BAERLIN2014-campaign. Note, that the diurnal cycles based on averages instead of medians are quite similar: ~~only the afternoon maximum is~~ during the afternoon (largest concentrations) averages are about 5 μg/m^3 larger than medians. There is a close agreement between all stations, not only for the mean diurnal cycle but also on a daily basis (not shown here) suggesting that the spatial dependence of ozone concentration is less pronounced. This can be expected as ozone is not emitted but formed in the atmosphere within several hours after release of precursors. Thus, transport and mixing are key driving forces.

The diurnal cycles for NO_x concentrations are shown in Fig. 11, again calculated as medians. The concentration at the stations at the outskirts of Berlin (dotted lines) are the lowest with a maximum

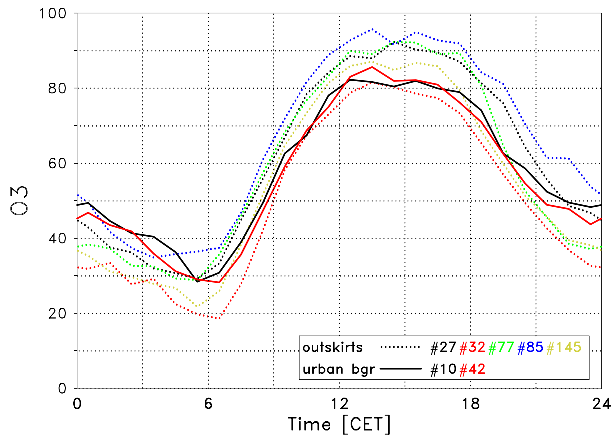

Figure 10. Diurnal cycle of O_3 concentration [in μg/m^3] at five stations at the outskirts (dotted lines) and two urban background stations (solid) as given in the legend (see also Table 1). Medians of the concentrations are plotted.

during the morning rush hours of not more than 25 μg/m^3. The urban background stations (solid lines) show larger concentrations with a morning maximum of up to about 40 μg/m^3. Significantly higher concentrations are observed at the traffic stations (dashed lines), again with a maximum during the morning rush hours. The absolute values and the development during the day are however much more diverse than at the less polluted locations. One reason can be that roadside NO_x-concentrations depend strongly on the distance from the source (e.g. Bonn et al., 2016; Richmond-Bryant et al., 2017), a similar situation as for PM_{10}. Due to the spatial variability of the mean diurnal cycles it is clear that for the traffic sites the correlations must vary as well. If averages instead of medians are considered NO_x-concentrations are somewhat larger (between 5 and 20 μg/m^3) and the morning maxima are slightly more pronounced.

Correlations between MLH and concentrations are shown in Fig. 12, separately for the three site-categories. For the outskirts of Berlin (leftmost block) and the urban background sites very strong positive correlations for ozone (circles) are derived. On average we find $R = 0.94$ for all sites and MLH retrievals. The differences between the sites are virtually negligible. One of the reasons for the very high correlations is that both MLH and O_3 concentration strongly increase after sunrise. The increase of the concentration is caused by the onset of ~~photochemistry and – later – possible mixing of~~ photochemical ozone production and by downward mixing of ozone from the residual layer ~~and the convective layer, while~~ in the morning hours when the mixing layer grows because of radiative heating of the ground and increasing convection. As shown by Fallmann et al. (2016) downward mixing of ozone from aloft can be a major source of near surface ozone for polluted urban sites with high NO_x levels. In "green" areas of

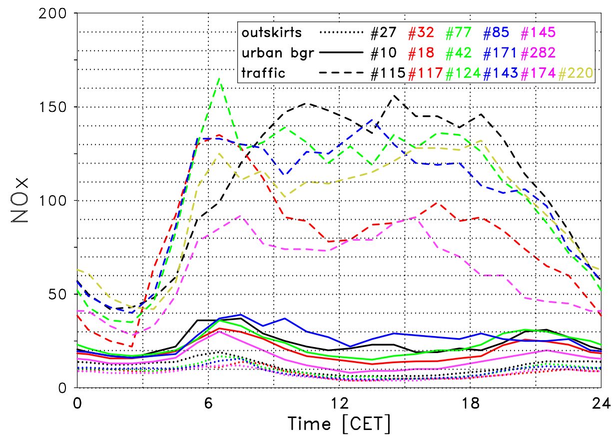

Figure 11. Diurnal cycle of NO_x concentration [in μg/m^3] at all BLUME stations: five at the outskirts of Berlin (dotted lines), five urban background stations (solid) and six traffic stations (dashed) as indicated in the legend (see also Table 1). Medians of concentrations are plotted.

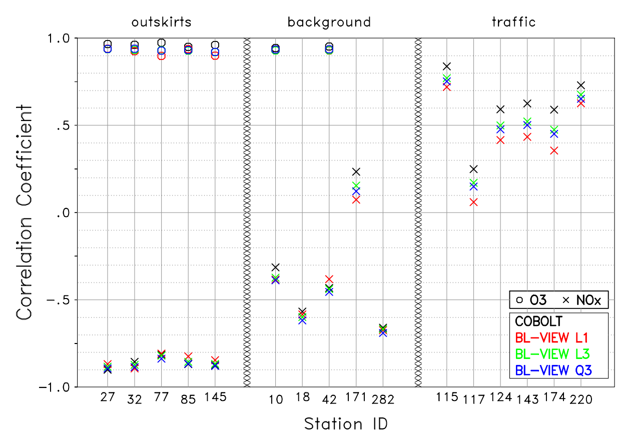

Figure 12. Correlation coefficient R of mean diurnal cycles of MLH and O_3 (circles) and NO_x-concentrations (crosses), respectively, shown for the 16 sites as indicated by the ID-number according to Table 1. The four MLH retrievals are color-coded according to the legend. Correlations based on MLH-averages and O_3- and NO_x-medians are plotted. Note, that at the traffic stations no O_3-measurements are available.

low NO_x concentration ozone production is also intensified. So, high correlations can be found, though there are manifold and partly different physical reasons are responsible.

The correlations between MLH and NO_x concentration at the sites on the outskirts are strongly negative, on average $R = -0.86$ is found. Again the spatial differences are almost negligible. An anti-correlation can be expected from mixing during the transport from the city center to the outskirts of Berlin. Negative correlations are also found for the urban background stations with the exception of station #171: due to their closer proximity to the main sources the absolute values on average are however slightly smaller ($R = -0.51$). Though labeled as "urban background" site #171 resembles much more the traffic sites. As already mentioned in Section 5.1, it is indeed close to a major road, but in contrast to the PM_{10}-concentration the presence of the nearby river does not counteract in a similar way the NO_x-distribution. For sites dominated by traffic a positive correlation is found, but with a wide spread of values from $R \approx 0.16$ for site #117 to $R \approx 0.77$ for site #115. Additionally, there is a pronounced dependence on the MLH retrieval, e.g. for site #174 $R = 0.36$ (L1 retrieval) and $R = 0.59$ (COBOLT).

These findings are confirmed by investigations on an hourly basis (up to 1608 cases, see Fig. 13). Correlation between MLH and O_3-concentration (open circles) are again high and virtually independent on the location. However, differences between the MLH retrievals are in the order of 0.2: for BL-VIEW Q3 on average $R = 0.52$ with a very small variation with the location (standard deviation of 0.02) is found, whereas the correlation is larger ($R = 0.71 \pm 0.01$) if COBOLT is applied. With respect to NO_x concentrations the correlation coefficients are approximately -0.36, -0.25 for outskirts and urban background stations, respectively, and much more dependent on the site. For the traffic sites there is however the correlation is weak with a large spread of $-0.1 \leq R \leq 0.3$. These results suggest that only in case of secondary compounds and primary pollutants in the absence of nearby traffic sources robust strong correlations between MLH and gaseous pollutants can be found. Whether this statement can be confirmed must be left to more extended studies.

6 Summary and conclusions

The height of the mixing layer (MLH) is expected to have an influence on air quality at the surface. It is assumed that extended mixing layers lead to dilution of pollutants and thus tend to decrease surface concentrations. Several publications have indeed reported such anti-correlations. However, neither the robustness and the representativeness of such correlations for metropolitan areas nor the role of choice of a MLH retrieval has the MLH retrieval have yet been investigated. Furthermore, ? reported different correlations for primary and secondary pollutants. This paper is devoted to these topics by examining the interrelationship relationship between MLH and near surface concentrations of particulate

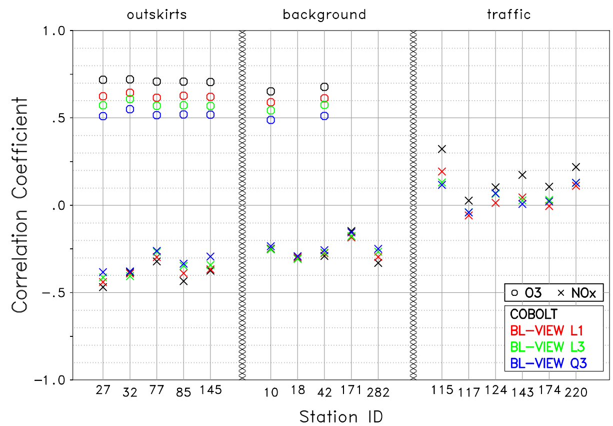

Figure 13. Same as Fig. 12 but correlation coefficient based on 1-hour measurements.

matter (PM_{10}), NO_x and O_3. It is based on two months of data from the field campaign BAERLIN2014 conducted in Berlin, Germany.

Frequently used ~~techniques~~ tools to determine the MLH are automated lidars and ceilometers ~~.~~ ~~Especially ceilometers~~ (ALC). Especially commercial systems with their unattended continuous operation are very promising since they are available as networks. Here, we compare four different approaches to determine the MLH, three of them based on proprietary software delivered by the manufacturer of the instrument (Vaisala), and the recently developed approach COBOLT (Geiß, 2016). ~~The properties of the retrievals are investigated using data from the field campaign BAERLIN2014 conducted during three summer months in Berlin.~~ It was found that a complete diurnal cycle with a high temporal resolution often cannot be derived from the proprietary software, and that there is a tendency to overestimate the MLH in the presence of the residual layer.

It is obvious that the differences of the retrieved MLH influence the correlation coefficients between MLH and pollutant concentrations. For mean diurnal cycles correlation coefficients differ by approximately 0.1 if different MLH retrievals are applied. These differences are smaller than the differences found when different locations in the city are compared – even if their distance is only a few ~~kilometres~~ kilometers from each other. In case of PM_{10} we found strong correlations as well as strong anti-correlations even if the sites are assigned to the same category (e.g. "urban background" or "traffic stations"). This clearly demonstrates that the MLH is not the only parameter controlling the surface concentration, and that local emissions and transport play a dominant role. This is in agreement to the pronounced heterogeneity over Berlin as reported by Bonn et al. (2016). In case of ozone as a secondary pollutant the correlations for different sites show only small differences. The strong correlation was found due to the similarity (although for different reasons) of the mean diurnal cycles of ozone and MLH with maximum values in the afternoon. An anti-correlation for near-surface concentrations of NO_x, as suggested by several previous studies, was only found in the absence of direct exposure to traffic sources.

We conclude that in case of ~~large cities~~ a large city as Berlin the MLH can be an indicator for urban air quality only in a very limited sense, and that any correlation between MLH and concentrations of pollutants should be treated with care: it is unlikely that they are representative for ~~an~~ the entire metropolitan area, in particular if the terrain is flat. At least for the observed summer period in Berlin this was not the case. Consequently, whenever links between MLH and near-surface concentrations are interpreted, it is mandatory to carefully describe the location, i.e., meteorological conditions, local sources etc., and the details of the MLH retrieval. Compared to the heterogeneity of the former we think that the selection of a certain MLH retrieval does not have the highest priority for correlation studies. It ~~is still open whether the situation is different for "more homogeneous" regions without pronounced changes in land use, and without significant local emissions. It would also~~ would be interesting to study winter time conditions when the PM_{10} concentrations in Berlin are about 50 % higher than in summer. We do not expect that in winter the MLH is the only controlling parameter, but it is not clear if the correlation (and its variability) is more or less pronounced. It remains open whether the situation is different for regions without pronounced changes in land use, without significant local emissions, or in areas with pronounced orography.

To better understand the complex interactions between the MLH, wind field, emissions, chemical processing etc. for air quality, there is a need ~~to have~~ for models down to a building-resolving scale as well as more extended data sets ~~in time and space~~ especially for heterogeneous areas. ~~In particular continuous~~ The specific setup of models and experiments must be defined according to the scale of interest. Continuous ceilometer measurements including at least one complete annual cycle ~~should be provided~~ can provide a significant contribution and help to investigate the generality of ~~our~~ the results, e.g. to check ~~whether they depend on the season or if there are significant~~ for seasonal changes or for differences between working days and weekends. It is obvious that ~~ceilometers~~ ALC with a very low overlap range are required for the observation of very shallow mixing layers typical during night time and in winter. Moreover, it would be nice to have co-incident ceilometer measurements at different sites or to have one or more mobile systems to check our hypothesis that the MLH does not change on a scale of a ~~few kilometres. In this context the differences of different MLH retrievals can be revisited.~~ large city.

It should be added that accurate retrievals of the MLH are beneficial for several applications: they can be used for box-model calculations ~~(e.g. Knote et al., 2015) , and are obligatory~~ and for the validation of meteorological models and the meteorological part of chemistry transport models. As the MLH is not a prognostic variable it is important to assess the accuracy of different parameterizations (e.g. Hu et al., 2010; Gan et al., 2011; Svensson et al., 2011; Banks et al., 2016). In this context a high accuracy of the MLH retrieval is crucial and a methodology that provides the full diurnal cycle without gaps and with high temporal resolution, and avoids wrong allocations of aerosol layers must be applied. Finally, we want to emphasize that state-of-the-art ceilometers ALC allow for the derivation of profiles of particle backscatter coefficients the particle backscatter coefficient β_p if the signals have been calibrated (e.g. O'Connor et al., 2004; Wiegner and Geiß, 2012). In case of ceilometers emitting in the spectral range near 910 nm, the signal must however be corrected for water vapor absorption (Wiegner and Gasteiger, 2015). These profiles Profiles of β_p can be used for the validation of chemistry transport models (e.g. Emeis et al., 2011) in a more direct way than the MLH as e.g. mixing ratios or mass concentrations of aerosols aerosol particles (or different aerosol components) are available as prognostic variables. Applying the adequate scattering theory, backscatter coefficients β_p can then be derived. On the long term perspective this is the preferable strategy for validation.

Acknowledgements. We like to thank Carsten Jahn (KIT/IMK-IFU) and Moritz Pickl during his internship at KIT/IMK-IFU for careful instrument operation and data analyses. We are grateful to Albrecht von Stülpnagel and Martin Schacht (both at Senate Department for Urban Development and the Environment, Berlin) for data-provision and technical assistance, and Holger Gerwig (German Environment Agency, Dessau) for helpful discussions.

The wind measurements discussed in Section 5.1 were obtained from the Climate Data Center of German Weather Service (DWD).

Alexander Geiß' contribution to this paper has received funding from the European Union Seventh Framework Programm (FP7/2007-2013) under grant agreement No. 262254. Ka Lok Chan was supported by the Marie Curie Initial Training Network of the European Seventh Framework Programme (Grant No. 607905).

References

Alföldy, B., Osán, J., Toth, Z., Török, S., Harbusch, A., Jahn, C., Emeis, S., and Schäfer, K.: Aerosol optical depth, aerosol composition and air pollution during summer and winter conditions in Budapest, Sci. Total Environ. 383, 141–163, 2007.

Baars, H., Ansmann, A., Engelmann, R., and Althausen, D.: Continuous monitoring of the boundary-layer top with lidar, Atmos. Chem. Phys., 8, 7281–7296, doi:10.5194/acp-8-7281-2008, 2008.

Bachtiar, V. S., Davies, F., and F. M. Danson, F. M.: A combined model for improving estimation of atmospheric boundary layer height, Atmos. Environ. 98, 461–473. doi: 10.1016/j.atmosenv.2014.09.028, 2014.

Banks, R. F., Tiana-Alsina, J., Baldasano, J. M., Rocadenbosch, F., Papayannis, A., Solomos, S., and Tzanis, C. G.: Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign, Atmos. Res. 176–177, 185–201, doi:10.1016/j.atmosres.2016.02.024, 2016.

Barlow, J. F.: Progress in observing and modelling the urban boundary layer, Urban Climate 10, 216–240, doi:10.1016/j.uclim.2014.03.011, 2014.

Beyrich, F.: Mixing-height estimation in the convective boundary layer using sodar data, Boundary-Layer Meteorology 74.1, 1–18, doi: 10.1007/ BF00715708, 1995.

Bonn, B., von Schneidemesser, E., Andrich, D., Quedenau, J., Gerwig, H., Lüdecke, A., Kura, J., Pietsch, A., Ehlers, C., Klemp, D., Kofahl, C., Nothard, R., Kerschbaumer, A., Junkermann, W., Grote, R., Pohl, T., Weber, K., Lode, B., Schönberger, P., Churkina, G., Butler, T. M., and Lawrence, M. G.: BAERLIN2014 – the influence of land surface types on and the horizontal heterogeneity of air pollutant levels in Berlin, Atmos. Chem. Phys., 16, 7785-7811, doi:10.5194/acp-16-7785-2016, 2016.

Brooks, I. M.: Finding Boundary Layer Top: Application of a Wavelet Covariance Transform to Lidar Backscatter Profiles, J. Atmos. Oceanic Technol. 20 (8), 1092–1105. doi:10.1175/1520-0426(2003)020<1092:FBLTAO>2.0.CO;2, 2003.

Brunner, D., Eder, B., Jorba, O., Savage, N., Makar, P., Giordano, L., Badia, A., Balzarini, A., Baro, R., Chemel, C., Forkel, R., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Knote, C., Kuenen, J.J.P., Makar, P.A., Manders-Groot, A., Davis, L., Perez, J.L., Pirovano, G., San José, R., Savage, N., Schröder, W., Sokhi, R. S., Syrakov, D., Torian, A., Werhahn, K., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S.: Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2, Atmos. Environ. 115, 470–498, doi:10.1016/j.atmosenv.2014.12.032, 2015.

Caicedo, V., Rappenglück, B., Lefer, B., Morris, G., Toledo, D., and Delgado, R.: Comparison of aerosol lidar retrieval methods for boundary layer height detection using ceilometer aerosol backscatter data, Atmos. Meas. Tech., 10, 1609-1622, doi:10.5194/amt-10-1609-2017, 2017.

Chen, B., and Kann, H.: Air pollution and population health: a global challenge. Environ. Health Prev. Med. 13 (28), 94–101, doi:10.1007/s12199-007-0018-5, 2008.

Cohn, S. A. and Angevine, W. M.: Boundary Layer Height and Entrainment Zone Thickness Measured by Lidars and Wind-Profiling Radars, J. Appl. Meteor. 39 (8), 1233–1247, 2000.

Comerón, A., Sicard, M., and Rocadenbosch, F.: Wavelet Correlation Transform Method and Gradient Method to Determine Aerosol Layering from Lidar Returns: Some Comments, J. Atmos. Oceanic Technol., 30 (6), 1189–1193, doi: 10.1175/JTECH-D-12-00233.1, 2013.

Czader, B. H., Li, X., and Rappenglueck, B.: CMAQ modeling and analysis of radicals, radical precursors and chemical transformations, J. Geophys. Res., 118, 11,376–11,387, doi: 10.1002/jgrd.50807, 2013.

de Bruine, M., Apituley, A., Donovan, D. P., Klein Baltink, H., and de Haij, M. J.: Pathfinder: applying graph theory to consistent tracking of daytime mixed layer height with backscatter lidar, Atmos. Meas. Tech., 10, 1893–1909, doi:10.5194/amt-10-1893-2017, 2017.

Dieudonné, E., Ravetta, F., Pelon, J., Goutail, F., and Pommereau, J.-P.: Linking NO_2 surface concentration and integrated content in the urban developed atmospheric boundary layer. Geophys. Res. Lett., 40 (6), 1247–1251, doi:10.1002/grl.50242, 2013.

Du, C., Liu, S., Yu, X., Li, X., Chen, C., Peng, Y., Dong, Y., Dong, Z., and Wang, F.: Urban Boundary Layer Height Characteristics and Relationship with Particulate Matter Mass Concentrations in Xi'an, Central China, Aerosol Air Qual. Res. 13, 1598–1607, doi: 10.4209/aaqr.2012.10.0274, 2013.

Duda, R., and Hart, P.: Pattern Classification and Scene Analysis, John Wiley & Sons Inc., 1973.

Elminir, H. K.: Dependence of urban air pollutants on meteorology, Sci. Total Environ. 350, 225–237, doi:10.1016/j.scitotenv.2005.01.043, 2005.

Emeis, S., Münkel, C., Vogt, S., Müller, W., and Schäfer, K.: Atmospheric boundary-layer structure from simultaneous SODAR, RASS, and ceilometer measurements, Atmos. Environ. 38, 273–286, 2004.

Emeis, S., Forkel, R., Junkermann, W., Schäfer, K., Flentje, H., Gilge, S., Fricke, W., Wiegner, M., Freudenthaler, V., Groß, S., Ries, L., Meinhardt, F., Birmili, W., Münkel, C., Obleitner, F., and Suppan, P.: Measurement and simulation of the 16/17 April 2010 Eyjafjallajökull volcanic ash layer dispersion in the northern Alpine region, Atmos. Chem. Phys., 11, 2689–2701, doi:10.5194/acp-11-2689-2011, 2011.

Emeis, S., Schäfer, K., Münkel, C., Friedl, R., and Suppan, P.: Evaluation of the interpretation of ceilometer data with RASS and radiosonde data, Bound.-Lay. Meteorol. 143, 25–35, 2012.

Endlich, R. M., Ludwig, F. L., and Uthe, E. E.: An automatic method for determining the mixing depth from lidar observations, Atmos. Environ. 13 (7), 1051–1056. doi:10.1016/0004-6981(79)90015-5, 1979.

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

Fallmann, J., Forkel, R., and Emeis, S.: Secondary effects of urban heat island mitigation measures on air quality, Atmos. Environ.,125 A, 199–211, doi:10.1016/j.atmosenv.2015.10.094, 2016.

Flamant, C., Pelon, J., Flamant, P. H., and Durand, P.: Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer, Boundary-Layer Meteoro. 83 (2), 247–284, doi:10.1023/A:1000258318944, 1997.

Gan, C.-M., Wu, Y., Madhavan, B. L., Gross, B., and Moshary, F.: Application of active optical sensors to probe the vertical structure of the urban boundary layer and assess anomalies in air quality model PM2.5 forecasts, Atmos. Environ. 45 (37), 6613–6621, doi:10.1016/j.atmosenv.2011.09.013, 2011.

Geiß, A.: Automated calibration of ceilometer data and its applicability for quantitative aerosol monitoring, PhD thesis, Ludwig-Maximilians-Universität, München, 185 pp., 2016.

Grant, A. L. M.: An observational study of the evening transition boundary-layer, Quart. J. Royal Met. Soc. 123, 657–677, 1997.

Grimsdell, A. W. and Angevine, W. M.: Observations of the Afternoon Transition of the Convective Boundary Layer, J. Appl. Meteor. 41, 3–11, 2002.

Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Leed, Y., and Kumare, N.: Satellite remote sensing of particulate matter and air quality assessment over global cities, Atmos. Environ. 40, 5880–5892, doi:10.1016/j.atmosenv.2006.03.016, 2006.

Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G. P., Lolli, S., O'Dowd, C. D., Sauvage, L., Xueref-Rémy, I., Wastine, B., and Feist, D. G.: Evaluation of mixing-height retrievals from automated profiling lidars and ceilometers in view of future integrated networks in Europe, Boundary Layer Meteorol., 143, 49–75, doi: 10.1007/s10546-011-9643-z, 2012.

Haman, C. L., Lefer, B., and Morris, G. A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment, J. Atmos. Oceanic Technol., 29, 697–710, 2012.

Hervo, M., Poltera, Y., and Haefele, A.: An empirical method to correct for temperature-dependent variations in the overlap function of CHM15k ceilometers, Atmos. Meas. Tech., 9, 2947-2959, doi:10.5194/amt-9-2947-2016, 2016.

Holzworth, G. C.: Estimates of mean maximum mixing depths in the contiguous United States. Mon. Wea. Rev. 92, 235–242, doi:10.1175/1520-0493(1964)092<0235:EOMMMD>2.3.CO;2, 1964.

Hooper, W. P. and Eloranta E. W.: Lidar Measurements of Wind in the Planetary Boundary Layer: The Method, Accuracy and Results from Joint Measurements with Radiosonde and Kytoon, J. Climate Appl. Meteor. 25 (7), 990–1001, doi:10.1175/1520-0450(1986)025<0990:LMOWIT>2.0.CO;2, 1986.

Hu, X.-M. and Nielsen-Gammon, J. W.: Evaluation of three planetary boundary layer schemes in the WRF model, J. Appl. Meteorol. Clim. 49, 1831–1844, doi:10.1175/2010JAMC2432.1, 2010.

Hu, X.-M., Ma, Z., Lin, W., Zhang, H., Hu, J., Wang, Y., Xu, X., Fuentes, J. D., and Xue, M.: Impact of the Loess Plateau on the atmospheric boundary layer structure and air quality in the North China Plain: A case study, Sci. Total Environ., 499, 228–237, doi:10.1016/j.scitotenv.2014.08.053, 2014.

Knote, C., Tuccella, P., Curci, G., Emmons, L., Orlando, J. J., Madronich, S., Baró, R., Jiménez-Guerrero, P., Luecken, D., Hogrefe, C., Forkel, R., Werhahn, J., Hirtl, M., Pérez, J. L., San José, R., Giordano, L., Brunner, D., Yahya, K., and Zhang, Y.: Influence of the choice of gas-phase mechanism on predictions of key gaseous pollutants during the AQMEII phase-2 intercomparison, Atmos. Environ., 115, 553–568, doi:10.1016/j.atmosenv.2014.11.066, 2015.

Koelemeijer, R. B. A., Homan, C. D., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmos. Env., 40 (27), 5304–5315, doi: 10.1016/j.atmosenv.2006.04.044, 2006.

Kotthaus, S., O'Connor, E., Münkel, C., Charlton-Perez, C., Haeffelin, M., Gabey, A. M., and Grimmond, C. S. B.: Recommendations for processing atmospheric attenuated backscatter profiles from Vaisala CL31 ceilometers, Atmos. Meas. Tech., 9, 3769-3791, doi:10.5194/amt-9-3769-2016, 2016.

Kuik, F., Lauer, A., Beukes, J. P., Van Zyl, P. G., Josipovic, M. Vakkari, V., Laakso, L.,and Feig, G. T.: The anthropogenic contribution to atmospheric black carbon concentrations in southern Africa: a WRF-Chem modeling study, Atmos. Chem. Phys., 15, 8809–8830, doi:10.5194/acp-15-8809-2015, 2016.

Lange, D., Tiana-Alsina, J., Saeed, U., Tomas, S., and Rocadenbosch, F.: Atmospheric Boundary Layer Height Monitoring Using a Kalman Filter and Backscatter Lidar Returns, IEEE Transactions on Geoscience and Remote Sensing, 52 (8), 4717–4728, doi:10.1109/TGRS.2013.2284110, 2014.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature, 525, 367–371, doi: 10.1038/nature15371, 2015.

LeMone, M. A., Tewari, M., Chen, F., and Dudhia, J.: Objectively Determined Fair-Weather CBL Depths in the ARW-WRF Model and Their Comparison to CASES-97 Observations. Mon. Weather Rev., 141, 30–54, doi:10.1175/MWR-D-12-00106.1, 2013.

Li, S., Joseph, E., and Min, Q.: Remote sensing of ground-level PM2.5 combining AOD and backscattering profile, Remote Sensing of Environment, 183, 120–128, http://dx.doi.org/10.1016/j.rse.2016.05.025, 2016.

Liu S. and Liang, X.-Z.: Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height, J. Clim., 23, 5790–5809, doi: 10.1175/2010JCLI3552.1, 2010.

Lotteraner, C. and Piringer, M.: Mixing-Height Time Series from Operational Ceilometer Aerosol-Layer Heights, Boundary-Layer Meteorol., doi: 10.1007/s10546-016-0169-2, 2016.

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

Menut, L., Flamant, C., Pelon, J., and Flamant, P. H.: Urban Boundary-Layer Height Determination from Lidar Measurements Over the Paris Area, Appl. Opt. 38 (6), 945–954. doi: 10.1364/AO.38.000945, 1999.

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

Münkel, C.: Mixing height determination with lidar ceilometers – results from Helsinki Testbed, Meteorol. Z. 16, 451–459, 2007.

Münkel, C., Schäfer, K., and Emeis, S.: Adding confidence levels and error bars to mixing layer heights detected by ceilometer, Proc. SPIE 8177, 817708–1 – 817708–9, 2011.

Muñoz, R. C. and Undurraga, A. A.: Daytime mixed layer over the Santiago Basin: Description of two years of observations with a lidar ceilometer, J. Appl. Meteorol. Clim., 49, 1728–1741, 2010.

O'Connor, E. J., Illingworth, A. J., and Hogan, R. J.: A technique for autocalibration of cloud lidar, J. Atmos. Ocean. Tech., 21, 777–786, 2004.

Pal, S., Xueref-Remy, I., Ammoura, L., Chazette, P., Gibert, F., Royer P., Dieudonné, E., Dupont, J.-C., Haeffelin, M., Lac, C., Lopez, M., Morille, Y., and Ravetta, F.: Spatio-temporal variability of the atmospheric boundary layer depth over the Paris agglomeration: An assessment of the impact of the urban heat island intensity, Atmos. Environ. 63, 261–275, doi: 10.1016/j. atmosenv.2012.09.046, 2012.

Pal, S., Lee, T. R., Phelps, S., and De Wekker, S. F. J.: Impact of atmospheric boundary layer depth variability and wind reversal on the diurnal variability of aerosol concentration at a valley site, Sci. Total Environ., 496, 424–434, doi: 10.1016/j.scitotenv.2014.07.067, 2014.

Pal, S., and Haeffelin, M.: Forcing mechanisms governing diurnal, seasonal, and interannual variability in the boundary layer depths: Five years of continuous lidar observations over a suburban site near Paris, J. Geophys. Res. Atmos., 120, 11,936–11,956, doi:10.1002/ 2015JD023268, 2015.

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

Piringer, M., Joffre, S., Baklanov, A., Christen, A., Deserti, M., De Ridder, K., Emeis, S., Mestayer, P., Tombrou, M., Middleton, D., Baumann-Stanzer, K., Dandou, A., Karppinen, A., and Burzynski, J.: The surface energy balance and the mixing height in urban areas – activities and recommendations of COST-Action 715, Boundary-Layer Meteorol., 124, 3–24, doi:10.1007/s10546-007-9170-0, 2007.

Richmond-Bryant, J., Owen, R. C., Graham, S., Snyder, M., McDow, S., Oakes, M., and Kimbrough, S.: Estimation of on-road NO_2 concentrations, NO_2/NO_x ratios, and related roadway gradients from near-road monitoring data. Air Qual. Atmos. Health, doi:10.1007/s11869-016-0455-7, 2017.

Rost, J., Holst, T., Sähn, E., Klingner, M., Anke, K., Ahrens, D., and Mayer, H.: Variability of PM_{10} concentrations dependent on meteorological conditions. Int. J. Environ. Pollut. 36, 3–18, 2009.

Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., and Peters, A.: Health Effects of Particulate Air Pollution – A Review of Epidemiological Evidence, Inhal. Toxicol. 23, 555–592, 2011.

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

Schäfer, K., Emeis, S., Hoffmann, H., and Jahn, C.: Influence of mixing layer height upon air pollution in urban and sub-urban areas, Meteorol. Z. 15, 647–658, 2006.

Schäfer, K., Harbusch, A., Emeis, S., Koepke, P. and Wiegner, M.: Correlation of aerosol mass near the ground with aerosol optical depth during two seasons in Munich, Atmos. Environ. 42, 4036–4046, 2008.

Schäfer, K., Wagner, P., Emeis, S., Jahn, C., Münkel, C. and Suppan, P.: Mixing layer height and air pollution levels in urban area, in Remote Sensing of Clouds and the Atmosphere XVII; and Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing VIII, edited by E. I. Kassianov, A. Comeron, R. H. Picard, K. Schäfer, U. N. Singh, G. Pappalardo, Proc. of SPIE, 8534, doi:10.1117/12.974328, 2012.

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

Seibert, P., Beyrich, F., Gryning, S. E., Joffre, S., Rasmussen, A., and Tercier, P.: Review and intercomparison of operational methods for the determination of the mixing height, Atmos. Environ. 34, 1001–1027, 2000.

SenStadtUm, Senate Department for Urban Development and the Environment, Württembergische Straße 6, 10707 Berlin: http://www.stadtentwicklung.berlin.de/umwelt/luftqualitaet/de/messnetz/index.shtml, 2014.

Sicard, M., Pérez, C., Rocadenbosch, F., Baldasano, J. M., and Garcia-Vizcaino, D.: Mixed-Layer depth determination in the Barcelona coastal area from regular lidar measurements: Methods, results and limitations, Bound.-Lay. Meteorol., 119, 135–157, 2006.

Sun, Y., Song, T., Tang, G., and Wang, Y.: The vertical distribution of $PM_{2.5}$ and boundary-layer structure during summer haze in Beijing. Atmospheric Environment 74, 413–421, doi:10.1016/j.atmosenv.2013.03.011, 2013.

Svensson, G., Holtslag, A. A., Kumar, V., Mauritsen, T., Steeneveld, G. J., Angevine, W. M., Bazile, E., Beljaars, A., de Bruijn, E. I., Cheng, A., Conangla, L., Cuxart, J., Ek, M., Falk, M. J., Freedman, F., Kitagawa, H., Larson, V. E., Lock, A., Mailhot, J., Masson, V., Park, S., Pleim, J., Soderberg, S., Weng, W., and Zampieri, M.: Evaluation of the diurnal cycle in the atmospheric boundary layer over land as represented by a variety of single-column models: The second GABLS experiment. Boundary-Layer Meteorology, 140, 177-206, doi:10.1007/s10546-011-9611-7, 2011.

Tai, A. P. K., Mickley, L. J., and Jacob, D. J.: Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change, Atmos. Environ. 44, 3976–3984, doi:10.1016/j.atmosenv.2010.06.060, 2010.

Tandon, A., Yadav, S., and Attri, A. K.: Coupling between meteorological factors and ambient aerosol load, Atmos. Environ. 44, 1237–1243, 2010.

Tang, G., Zhu, X., Hu, B., Xin, J., Wang, L., Münkel, C., Mao, G., and Wang, Y.: Impact of emission controls on air quality in Beijing during APEC 2014: lidar ceilometer observations, Atmos. Chem. Phys., 15, 12667–12680, doi:10.5194/acp-15-12667-2015, 2015.

Tiwari, S., Bisht, D. S., Srivastava, A. K., Pipal, A. S., Taneja, A., Srivastava, M. K., and Attri, S. D.: Variability in atmospheric particulates and meteorological effects on their mass concentrations over Delhi, India, Atmos. Res., 145–146, 45–56, doi:10.1016/j.atmosres.2014.03.027, 2014.

van der Kamp, D. and McKendry, I.: Diurnal and seasonal trends in convective mixed-layer heights estimated from two years of continuous ceilometer observations in Vancouver, BC, Bound.-Lay. Meteorol., 137, 459–475, 2010.

~~von Schneidemesser, E., Bonn, B., Butler, T. M., Ehlers, C., Gerwig, H., Hakola , H., Helln, H., Kerschbaumer, A.,Klemp, D., Kofahl, C., Kura, J., Ldecke, A., Nothard, R., Pietsch, A., Schfer, K., Schauer, J. J., Villalobos, A.M., and Lawrence, M. G.: BAERLIN2014 – stationary measurements and source apportionment at an urban background station in Berlin, Germany, in preparation, 2017.~~

Wagner, P. and Schäfer, K.: Influence of mixing layer height on air pollutant concentrations in an urban street canyon. Urban Climate, doi: 10.1016/j.uclim.2015.11.001, 2015.

Wiegner, M., Emeis, S., Freudenthaler, V., Heese, B., Junkermann, W., Münkel, C., Schäfer, K., Seefeldner, M., and Vogt, S.: Mixing Layer Height over Munich, Germany: Variability and comparisons of different methodologies, J. Geophys. Res. Atmos. 111, D13201, doi:10.1029/2005JD006593, 2006.

Wiegner, M. and Geiß, A.: Aerosol profiling with the Jenoptik ceilometer CHM15kx, Atmos. Meas. Tech., 5, 1953–1964, doi:10.5194/amt-5-1953-2012, 2012.

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

Wiegner, M. and Gasteiger, J.: Correction of water vapor absorption for aerosol remote sensing with ceilometers, Atmos. Meas. Tech., 8, 3971–3984, doi:10.5194/amt-8-3971-2015, 2015.

Wildmann, N., Rau, G., and Bange, J.: Observations of the Early Morning Boundary-Layer Transition with Small Remotely-Piloted Aircraft, Boundary-Layer Met. 157, 345–373, doi:10.1007/s10546-015-0059-z, 2015.

~~Xu, W. Y., Zhao, C. S., Ran, L., Deng, Z. Z., Liu, P. F., Ma, N., Lin, W. L., Xu, X. B., Yan, P., He, X., Yu, J., Liang, W. D., and Chen, L. L.: Characteristics of pollutants and their correlation to meteorological conditions at a suburban site in the North China Plain. Atmos. Chem. Phys., 11, 4353–4369, doi:10.5194/acp-11-4353-2011, 2011.~~