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

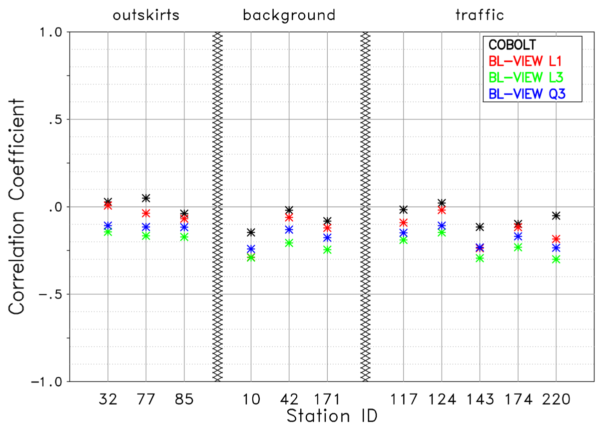

Figure 1: Correlation coefficient R of hourly values of MLH and PM_{10} shown for the 11 sites as indicated (station ID according to Table 1). The four MLH retrievals are color-coded according to the legend. Only measurements between 07:01 CET and 20:00 CET are considered.

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

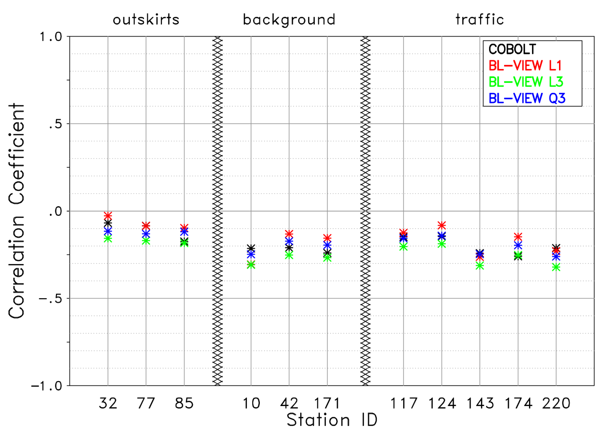

Figure 2: Same as Fig. 1, but only measurements before 07:00 CET and after 21:00 CET are considered. NOTE: THESE FIGURES ARE FOR DEMONSTRATION ONLY, NOT INCLUDED IN THE MANUSCRIPT.

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.