comparison between the three sensors, presented in section 5.1, should be rather considered as an evaluation of the performances of the two ceilometers respect to the lidar. Lidar capability to retrieve boundary layer height has been largely investigated. ASL3000 higher performances, mentioned several times in the manuscript, are obviously related to the basic differences existing between an elastic backscatter lidar system and a ceilometer, that are related to the basic concept of the two sensors, designed for different purposes. The differences mainly concerns the laser source, sampling rate of the atmosphere and optical receiver (e.g. Dupont et al., 1994 - Bound. Layer Meteor; Matthias et al., 2002 – Atmos. Research).

We thank the Reviewer for her/his comments. In our manuscript no assumptions are made regarding the a priori correctness of the Lidar with respect to the ceilometer in detecting the structure of the boundary-layer (BL). We appreciate that the state-of-art of the boundary layer studies and retrieval techniques to obtain its height largely proves that the lidar is capable to detect the vertical structure of the BL (we refer to the cited works in the manuscript), however we feel that a neutral analysis of the lidar-ceilometer skills to retrieve the BL structure and its comparison to the in-situ observations can not be simplified by the assumption of lidar's higher performances. We believe that it should not be taken into account the assumption of lidar superiority when compared to ceilometers when dealing with the cloud-free lower troposphere with small signal attenuation and at heights in the overlap region. Moreover, as the Reviewer highlighted in his/her review, the signal-to-noise ratio (SNR) plays an important role in the detection of the BL structure, however the SNR can not be assumed to be always higher for a lidar since the ratio depends equally on the averaged pulsed power and on the background noise impinging onto the receiver. With the advent of the new-generation ceilometers like the CHM15K with solid-state laser source and enhanced power, the gap between lidars and ceilometers is significantly reduced.

My opinion is that the authors should address several aspects before providing a quantitative assessment of ceilometers vs. lidar performance in retrieving boundary layer (BL) height. This manuscript sounds more like a technical report of an intercomparison campaign with not strong advances in the scientific knowledge.

We beg to disagree with the Reviewer, our manuscript presents a scientific analysis on SML and DRCL retrieval using different remote sensors and in-situ observations, with Mace Head being of strategic importance in Europe. We believe that the 20-day events analyzed during the ICOS campaign will be of substantial interest to the scientific community especially for what concerns the assessment of different retrieval techniques to determine the BL height Moreover, this study provides the opportunity to perform a three-way comparison amongst a lidar, a previous-generation and a new-generation ceilometer where the last two are certainly the most utilized devices by national Met Offices in Europe as primary instruments to detect the BL height

Also the attempt to assess THT algorithm performances on the selected cases should be performed on a larger dataset that refers to a larger set of atmospheric scenarios. However, if one of the aim of manuscript is to assess THT performances, this should be explicitly mentioned in the paper. Moreover, a brief description of the error on the BL height retrieved using the THT algorithm should be reported. Moreover, their conclusions come out from a measurement dataset that is not sufficient for a significant statistical analysis.

The time frame specified by the duration of the field campaign along with the number of cases collected for the comparison with the in-situ observations is certainly not as large as to cover the full range of meteorological and environmental conditions or to provide a “definite word” on lidar-ceilometer-radiosounding intercomparison. However, the studied dataset is certainly large enough to perform the instrument intercomparison with clear indications of the different sensors capabilities in retrieving the BL structure. In order to achieve this result the authors of this manuscript selected cases that would show clear details of intercomparisons in terms of statistical result as it is summarized in tables 1 and 3, i.e. mean values and mean distances between the retrievals as well as the related statistical uncertainty in assessing the instruments and instruments-radiosoundings correlation..
It is not an explicit objective of this study to assess the performances of the THT algorithm, even though we understand that this may lead to confusion since all retrieval are obtained by applying the THT to the different lidar and ceilometer profiles. We will then add clear statement in the abstract and through the manuscript about the role of THT as neutral methodology to assess the lidar-ceilometer-sounding comparison.

A description of the error on the BL height retrieved using the THT algorithm will be provided in the revised manuscript as stated below:
“The error-bars in fig 3 and 4 are the standard deviation representing the variability of contiguous BLH detections over a fixed interval of time (Martucci et al., 2010).”

I also encourage the authors to reconsider the section 5.2 where a comparison with in-situ measurements is described. The comparison is based on a restricted dataset that is not sufficient to perform a quantitative assessment of the reliability of both the definition of BL, based on the aerosol distribution, and the instrumental performances in detecting the BL.

We refer here to our previous answer concerning the aim of our study and the dimension of the dataset.

The statistics reported in this section aims at assessing the difference between the two methods for the retrieval of the boundary layer height. This difference are within a few hundreds of meters. This differences are consistent with the results presented by Seidel et al. 2010 (JGR), where several methods reported in literature for retrieving the boundary layer height from radiosounding measurements are compared over 10 years of data. Considering that a comparison between different instruments and different methods for the retrieval of the BL height is proposed in the manuscript, I suggest the authors to investigate larger dataset of co-located measurements before assessing the adequacy or inadequacy of the BL retrieval obtained using aerosol lidar measurements.

We will add reference to Seidel et al, 2010. Please, do refer to our previous answer concerning the aim of our study and dimension of the dataset

Section 5.4 is very short and it is not necessary to let it separated from 5.3. I suggest to put together 5.3 and 5.4 sections.

The two sections will be merged into single Section 5.3.

2. The authors should be aware that they are dealing with different sensors, with different features, as different overlap functions, temporal sampling, laser sources, and soon. This means that an assessment of the performances of a technique or of an algorithm in retrieving the PBL should be assessed also in terms of different aerosol optical thickness scenarios before providing final conclusions that could be incomplete and wrong.

Across the studied dataset the atmospheric column above the sensors have varied accordingly to the different air masses advected over Mace Head bringing higher or lower aerosol load depending on continental polluted or clean marine conditions. Thus, the changing air masses from clean to polluted determined small and high aerosol optical thickness, respectively.

page 569, line 10: the ASL300 is an elastic backscatter lidar; this means that it is not able to provide profiles of the extinction coefficient without making a few strong assumptions.Please specify what assumptions are made in the retrieval of the extinction and backscatter coefficient.

A more concise version of the following extended explanation will be added to the text: The ALS300 retrieves aerosols optical properties using Klett backward inversion (Klett, 1981). This means that the backscattering-to-extinction ratio (BER) should be assumed in the inversion procedure. The BER depends on the microphysical, chemical and morphological properties of the particles and ranges from 0.05 to 0.01 sr-1. During the instrument set-up, it is possible to choose the suitable BER from software database depending on surrounding environment. Moreover, the retrieval needs as calibration a zone where the backscattering coefficient is known. Being in the UV, when it is possible, the reference zone is chosen in an aerosol-cloud-free region, where only Rayleigh scattering is not negligible. In this region, the value of the backscattering coefficient is then calculated theoretically using a radiosounding for air density when available or the values of standard atmosphere (USSA1976)

page 569, line 25: in order to avoid confusion, please specify that the CHM15k laser source is diode-pumped Nd-YAG laser (1064 nm) yielding about 8 μJ per pulse at 5-7 KHz repetition rate. Moreover, additional information about differences in the divergence of the systems could be useful. As a whole, please stay consistent in the description of the three systems giving the same information for all of them.

We will add the additional details and keep to consistent description of the devices where possible based on the manufacturing companies delivered specifications.

page 570, line 3: in the eq (1), please mention that the additional term B is the background noise.

Please see lines 18-19 page 570: “The last term B is the sum of the electronic and optical background noise.”

Moreover, the authors mention the optical efficiency of the system using the term O(h). What about the quantum efficiency?

Please see lines 16 page 570: “O(h) is the overlap function”. The overall instrumental optical efficiency is the term K. The quantum efficiency is related to the efficiency of photomultipliers and it is not specified here.

page 570, line 22: please include a reference for supporting this approximation. For the utilized wavelengths (355 nm, 910nm and 1064 nm) the relation $\alpha_{aer}, \beta_{aer} >> \alpha_{mol}, \beta_{mol}$ can be applied

We will add a reference.

page 572, line 10-13: the following sentence is quite unclear “The intercomparison’s output in terms of SML and DRCL detections retrieved by the THT algorithm can vary significantly from case to case depending on meteorological conditions and on the different instrumental skills “. Please explain better.

Will change to
Results returned for the SML and DRCL detections when the THT algorithm is applied to each of the instruments backscatter profiles can show significant differences from one case to another mainly due to the instruments detection abilities and the different meteorological conditions

page 572, line 18; please specify that X vs Y indicates you are referring to a couple of sensors.

We will make correction as below:
For each X vs. Y (where X and Y refers to the instruments sensors)

9. page 574, line 2-4: please include a reference for supporting the BL retrieval method from the inversions in the radiosounding temperature profile. For example, I suggest Seidel et al., 2010. Martucci et al 2007

We will include reference as detailed below:
Maxima in the vertical gradient of the temperature profile of the radiosoundings were used to determine SML and DRCL layers which then could be compared to the sensors’ retrievals (Seidel et al., 2010;. Martucci et al 2007)

10. page 574, line 22-25: this sentence again shows that your study should be considered as an assessment of ceilometer vs lidar performance in retrieving BL height and not as an intercomparison between three sensors. This should be explicitly mentioned throughout the paper.

We would like to highlight again that this study is an intercomparison amongst three sensors and this must be intended as the answer we provided above, i.e. “this study provides the opportunity to perform a three-way comparison amongst a lidar, a previous-generation and a new-generation ceilometer”. In the revised manuscript we will add this sentence in order to avoid any misleading interpretation of the nature of the intercomparison.

11. page 576 line 3-4: the authors could also include information relative to wind speed or distance of the radiosonde from the launch station. This represents a useful information to address the possible co-location mismatch between radiosonde and lidar/ceilometer data and to better evaluate the time average solution used for comparing the BL retrieval algorithms using radiosonde and lidar/ceilometer observations.

We will include discussion and a table with these details.

12. page 577, line 9-13: this sentence is quite unclear "Also, rapidly-ïnˇC´ uctuating SML and DRCL upper boundaries are 10 hard to detect using lidar and ceilometer especially when the DRCL is shallow and the instrument vertical resolution can not resolve properly the weak gradients". Please explain better.

We will change this to

SML and DRCL can be difficult to detect when the distance between the SML and DRCL is very short or the instrument vertical resolution can not resolve the weak gradients at the layers' edges; this coupled with cases where the SML and DRCL heights are moving vertically at high temporal frequency adds to the complexity of detecting the two layers.

13. page 578, line 26-28: Please include a reference for supporting the aerosol-based boundary layer definition.

We will add the reference.

14. page 579, line 3-5: the authors should be very careful in providing any remark relative to the efficiency of the aerosol-based boundary layer definition using the results obtained from only one case study. This is definitely not sufficient. Moreover, this remark is relative to the detection of an inversion in the temperature profile at 315 m of altitude. Therefore, see next comment.

15. page 579, line 18: the authors are comparing an inversion detected in the radiosonde at 315 m with the lidar and ceilometer observations and they found no corresponding gradients in the lidar/ceilometer backscatter profiles. First of all, it is necessary to clarify that just one case does not make sense in a statistical analysis.

We have analyzed two case studies in Section 5.3 and 5.4 in which the detection of an aerosol layer using the gradient-based THT algorithm fails when compared to the in-situ detection; the aim of Sections 5.3 and 5.4 is not to provide statistical analysis but rather to present a case study and to hint at future studies in which to perform assessments of the different definition of BL based on different tracers.

.

However, this comparison is dubious. Your results could strongly depend on the lidar/ceilometer systems you are considering in this study. The disagreement could be related for example to the presence of a small amount of aerosol and to the low sensitivity of the lidar/ceilometer system to the small aerosol loading.

Whilst low sensitivity of the Lidar / celiometer can in general cause issues, the temperature inversion detected by the radiosounding for this specific case is at a level (315 m a.g.l.) where we expect the instruments to receive backscatter signal with very little attenuation. Moreover, both the ALS300 and the CL31 are in the region where the overlap is ~1 and can then provide reliable profiles. On the same profile at a higher height where the SNR is expected to be lower a second inversion is clearly detected by the instruments, its our opinion that the instruments have the ability to detect such an inversion and this case can not be explained by the sensitivity of the instruments. Alternatively, the authors have interpreted the failed detection as a temperature inversion forming in the already developed surface layer so that aerosols are homogeneously distributed below and above the inversion and no gradient is observed in the lidar data (but this is rather speculative so we decided not to include it in the paper).

Do you have any data relative to the sensitivity of these systems respect to different aerosol optical thickness? Moreover, the ALS 300 has a nominal full overlap height of 200 m, that is very close to the inversion height observed in the radiosonde temperature profile (315 m).
Has the nominal overlap of the ALS300 system been never assessed before by the authors?

In this case we show the overlap becoming full at 200 meters. Looking at figure 8 (top left hand panel) the overlap of the ALS300 clearly rises and becomes full above 200 meters. This proves the correctness of 200 meters as the full overlap level. With this in mind this case does not appear to be influenced by the incomplete overlap.

Finally, according to the data reported in the plots, the case study the authors are considering
for supposing the existence of possible limitations in using aerosol as a tracer
for the retrieval of BL height is also the only case where the BL is located below 450 - 500 m of altitude. This seems to agree with the above mentioned sensitivity or overlap issues.

Please see previous answer

All technical corrections as detailed by the referee will be corrected in the revised paper.