We would like to thank the Anonymous Referee #3 for devoting time in reading and commenting on our manuscript. Following his/her suggestions, the content of the manuscript has been significantly revised. More specifically, comparisons were directed between ground-based and space-borne lidar measurements as well as numerical estimations from WRF. The comparison between lidar and radiosondes is not included in the new version of the manuscript, since the nature of the two methods is different. Furthermore, we decided to include one atmospheric model in the comparison. Therefore, we selected the WRF, which has higher horizontal resolution than ECMWF. The comparison between WRF and radiosonde profiles and the application of the temperature, potential temperature and relative humidity criteria are excluded, since the comparison with lidar is indirect. As suggested by the two reviewers Section 5, (Comparison to another location) was left out. A new Section (4.2) has been added, in which the sensitivity analysis regarding the WCT threshold is discussed. Emphasis is given on the reasons that can explain the discrepancies with ancillary sources. In the same manner, the Conclusions and the Abstract of the manuscript have been rewritten. Excessive details regarding the Figures and Tables have now been removed.

1) **P3, L85: Please check the latitude and longitude of the two stations.**
   The coordinates of the lidar site and the radiosonde launch site have been checked. The radiosonde site is located NE of the lidar station, not NW.

   The corresponding text in the manuscript has been now corrected.

2) **P3, L93-108: This paragraph presents very detailed information about the meteorological conditions of the observation site, but I think the messages that authors have presented are not well selected and structured…authors could shorten this paragraph and present the information that serves the topics and results that will be present in the following sections.**

   The paragraph has been reconstructed and shortened as the reviewer suggests. The temperature and precipitation anomalies are used later in the new Section 4.4.2, where the seasonal cycle of PBLH is investigated in relation to the meteorological conditions. The new paragraph is as follows:

   Temperature and precipitation patterns can potentially reflect the state of sensible and latent heat fluxes within the PBL as well as the exchange of moisture and momentum with the Earth's surface. Thus, climatologies of meteorological parameters can be considered a valuable tool for assessing the representativeness of PBLH seasonal cycle with respect to long-term measurements. Such a comparison is performed in Section 4.4 based on the 30-year anomalies of maximum temperature and accumulated precipitation (Figure 1).

3) **P3-P4: I think the lidar system deserves more 'words', even it has been presented in other papers. Main messages should be addressed to the readers.**
   More details regarding the technical specifications of the portable Raman lidar system FMI-Polly^{XT} have been added in the manuscript (Section 3.1.1), such as the emitted wavelengths, the beam divergence and the telescope type and field of view:

The measurements were conducted with a six-channel Raman lidar called FMI-Polly^{XT} (Finnish Meteorological Institute - Portable Lidar sYstem eXTedend). The lidar system was entirely remotely controlled via an internet connection, with all the measurements, data transfer and built-in device regulation being performed automatically. The instrument was equipped with an uninterruptible power supply (UPS) and an air conditioning system (A/C) to allow for safe and smooth continuous measurements. A rain sensor was also connected to the roof cover in order to assure a proper shutdown of the instrument during rain.

FMI-Polly^{XT} used a Continuum Inline III type laser. The pulse rate of the laser was 20 Hz and it delivered energies of 180, 110 and 60 mJ simultaneously (with external second and third harmonic generators) at three different wavelengths, i.e. 1064, 532, 355 nm, respectively. A beam expander was used so as to enlarge the beam from approximately 6 mm to 45 mm. The remaining beam divergence after expansion was less than 0.2 mrad. The backscattered light was collected by a Newtonian telescope, which had a main mirror with a diameter of 30 cm and a field of view of 1 mrad. The output of the instrument included vertical profiles of the particle backscatter coefficient at three wavelengths, i.e. 355, 532 and 1064 nm (retrieved with the Klett method; Klett (1981) and Klett (1985)), extinction coefficient at 355 and 532 nm (retrieved with the Raman method (Ansmann et al., 1990; Ansmann et al., 1992) by using the Raman shifted lines of N_2 at 387 and 607 nm) and linear particle depolarization ratio at 355 nm. The system vertical resolution was 30 m and the vertical range covered the whole troposphere under cloudless conditions.

4) **P4, L145: "A certain range of signal is cut to avoid strong gradients in the lower range", so how do you choose the range to be cut? Is it automatic?**
In cases, where strong signal gradients appeared in the first hundred meters from ground, we made use of the option to cut the lower parts of the signal. This procedure was not automated, but it was performed manually. More specifically, we started cutting the first height bin (30 m) above ground and we tested whether the WCT algorithm managed to omit these strong gradients. If the latter was not successful, then we repeated the above mentioned procedure for the first two height bins (60 m). The algorithm offers the capability to perform the cutting off procedure up to 29 height bins (870 m). In case the algorithm did not manage to detect a significant gradient, then it was not possible to detect the PBL height.

5) **P4, L152: Comment the abbreviation 'L-R'; and the paper "Wandinger and Ansmann, 2002" is not included in the bibliography.**
The abbreviation L-R was commented as Laser-Receiver. The publication of Wandinger and Ansmann, 2002 has now been added in the bibliography of the manuscript.

6) **The Subsection '3.1.3 Data coverage' is very 'dry' and not so interesting. It does provide some information, but it is too detailed with unnecessary technical issues, and some numbers are not well defined so I got confused and sometimes had to redo these simple calculations. I think this paragraph could be shortened.**

The Data Coverage Section has now been shortened. An emphasis is given on the factors that either prohibited the operation of the lidar system or hampered the detection of the PBL top. The new Section is as follows:

During the one-year long measurement campaign FMI-Polly^{XT} was measuring on 139 days. Due to technical problems with the laser, the data coverage from September to January was sparse. Furthermore, precipitation prohibited lidar measurements, since the lidar system had to shut down. Hence, sufficient data availability was achieved during 72 days. Multiple aerosol layers appeared mainly between March and May, whereas low clouds were present mostly in the monsoon period and both complicated PBL top detection. Additionally, some technical issues arose due to photomultiplier supersaturation and signal problems. A lack of a significant decrease in the backscatter profile was observed in only a few cases. The latter was a first indication that the modified WCT method can detect the PBL top efficiently, as long as the signal decrease threshold is tuned properly. The data coverage is presented on a monthly basis in Figure 2. The highest PBLH detection frequency was achieved in February, which can be attributed to favorable meteorological conditions, since low clouds appeared sparsely without any rainfall events.

7) **P8, L203: And in Figure 3, Please comment the PBLH close to 0 m, why do the models produce such low values? There are two T profiles in the figure at right, is it a mistake? What are the white squares in Figure 3 upper panel?**
ECMWF produces low PBLH values. The surface layer scheme, which is utilized in the ECMWF model for describing the turbulent transfer of heat, momentum and moisture between the surface and the lower parts of the atmosphere, allows a consistent treatment of different roughness lengths for momentum, heat and moisture. However, it has been found that the revised stability functions reduce diffusion in stable situations producing a shallower stable boundary layer (ECMWF, 2010b, p. 37).
Regarding numerical weather prediction (NWP) models, the understanding of turbulence in nocturnal stable boundary layer (SBL) and its parameterization is rather slow and not well established (Mahrt et al., 1999; Beare et al., 2006; Hong 2010). As a result, there is a tendency of the PBLH to remain at the lowest model level mainly due to the deficiency in SBL mixing and partly due to the poor vertical resolution in NWP models. In particular, the PBLH usually becomes the height of the lowest model level right after sunset. In the present study, this is partly addressed by the revised SBL scheme (Hong 2010) that computes the exchange coefficients with a parabolic function with height as in the mixed layer, in which the top of the SBL is determined by the bulk Richardson number (Ri), following the study of Vickers and Mahrt (2004). This leads to a gradual-and not abrupt-collapse of the mixed layer after the sunset due to the residual superadiabatic layer near the surface even in the presence of negative surface buoyancy flux. However, the fact that neither anthropogenic heat sources nor heat storage in buildings were included in the simulations could also explain the model underestimation during the night.
Furthermore, it should be noted that often the measurements depict different layers from the simulated ones, as in the case of the residual aerosol layer. The comparison could be improved if overall consistency in PBLH retrieval approaches between the model and lidar observations was obtained.The two T profiles were plotted by accident. These profiles, as stated above, are not included in the new manuscript.White squares in the contour plots of the lidar range-corrected signal indicate 15 min values of cloud base height.

8) **P9, L311: I miss a clear and quantitative difference between the PBL and RL. In Figure 3, there is a noticeable decrease of lidar signal at 1000-1200 m, do you consider it as the residual layer? Please also check the color scale of Figure 3. I saw a clear**

discontinuity at 06:00 UTC, 12:00 UTC and 17:30 UTC, are the lidar signals ploted with the same color scale? Moreover, the width of the black zone near the surface in Figure 3 is also changing, what does it mean? The authors should be more careful in preparing scientific figures.

During night-time, the configuration of FMI-Polly^{XT} permitted the determination of the Residual Layer height (RLH). The study of Wang et al. (2016) which was performed at a station of similar latitude, Wuhan, China, revealed that the RLH lies mostly in the range 0.5–1.3 km, following a seasonal variation. Hence, for most of our night-time cases we considered that the lidar system detected the top of the residual layer, which contained the aerosol of the previously mixed layer. In particular, if a layer top more than 500m was detected between sunset and sunrise, it was associated with the RLH.

This definition is now clarified in Section 3.1.1 of the manuscript.

The lidar data was available in 6-hour datasets. For this reason, the algorithm of the WCT method was applied separately to each 6-hour dataset. Furthermore, the color scale of the range-corrected signal contour plots is normalized with respect to the maximum signal recorded in each 6-hour dataset. The 6-hour quicklooks of the lidar-range corrected signal are made available by TROPOS (Leibniz Institute for Tropospheric Research) and can be accessed at http://polly.rsd.tropos.de/?p=lidarzeit&Ort=21.
The width of the black zone in lower part of Figure 3 (top panel) is representing the number of cut-off heights that are used. More specifically, during 00:00-12:00 UTC no cut-off heights were used, 12:00-18:00 UTC 3 cut-off heights (90 m) were used, while during 18:00-00:00 1 height bin was cut off.

These aspects have now been clarified in the manuscript (Label of Figure 3).

9) **P9, L315: PBLH=435 m is not found in Figure 3, check the text, table 2 and Figure 3.**

We thank the reviewer for noticing this mistake. In Figure 3 (lower panel), the PBLH from radiosonde measurements is given correctly according to the θ_crit and RH_crit method at 219 m Above Ground Level (AGL). However, in the text (line 315) and in Table 2 we gave the height in meters above mean sea level (ASL) (219 m AGL+ 216 m elevation = 435 m ASL). In the manuscript we kept the PBLH in meters AGL, since the PBLH derived from all of the methods are discussed in meters AGL.As stated above, the WRF-radiosonde profile comparison is not included in the new version of the manuscript.

10) **P9, L333, define PBL cycle, when does this cycle start and end?**
The PBLH cycle is defined as follows:
The PBLH growth period begins when the PBLH started to increase (typically 2-4 h after sunrise) and is complete when 90% of the daily maximum height is reached (typically between 08:00 and 10:30 UTC). More specifically, in the case of 2 March 2009, the PBLH growth period was completed at 7:30 UTC, which was one hour earlier compared to the completion of the PBLH cycle on the previous day. Furthermore, the PBLH started to grow approximately at 4:00 UTC on both days. In the manuscript, the PBL Cycle is defined in Section 3.1.2.

11) **P10, L354: which data did you use to derive this 553m/h? WRF?**

The PBLH growth rate between 3:00 and 5:00 UTC was determined by the cloud base height, which was assumed to be indicative of the PBL top and was derived using the WCT method as described in Section 3.1.2.

12) **Figure 4, and Figure 5: Again, check the color scale please. The discontinuity is quite obvious.**

The lidar data was available in 6-hour datasets. For this reason, the algorithm of the WCT method was applied separately to each 6-hour dataset. Furthermore, the color scale of the range-corrected signal contour plots is normalized with respect to the maximum signal recorded in each 6-hour dataset. The 6-hour quicklooks of the lidar-range corrected signal are made available by TROPOS (Leibniz Institute for Tropospheric Research) and can be accessed at http://polly.rsd.tropos.de/?p=lidarzeit&Ort=21. These aspects have now been clarified in the manuscript (Label of Figure 3).

The colorbars have been removed from Figure 3, 4 and 5, since the colorscale of each 6-hour contour plot is different.

13) **P10, Subsection 4.2.2: The data points are so few, and I do not think it is enough for a statistical study. "Based on the analyzed cases, it was found that the overpass distance (here 20 and 101 km) from the lidar station and time difference between the measurements did not affect the PBL heights. " This conclusion does not convince me, because the dataset is so small and cannot represent the spatial and temporal variability.**

The reviewer is right. Therefore we added the following comment after the statement ''based on…PBL heights'':

… *However, the small number of measurements does not allow us to generalize these findings. Hence, longer measurement periods or more extended comparison to ground stations are needed in order to draw more robust conclusions.*

14) **P11: PBLH diurnal Cycle might be more specific than PBL diurnal Cycle, because the authors investigated only the PBL height, not other parameters in the PBL**

In this work, the only parameter that we analyzed was the PBL height (PBLH). For this reason, as the reviewer suggests, we replaced the term PBL with PBLH in the manuscript.

15) **P12, L426: why is the comparison made only between lidar detection and ECMWF, how about WRF?**

The comparison of the PBLH diurnal cycle was performed between lidar and ECMWF Reanalysis because their data availability was sufficient throughout the measurement campaign. On the other hand, WRF simulations were dedicated to specific case studies that are analyzed in Section 4.1 so as to justify the PBLH derived by the WCT method under different aerosol load and meteorological conditions. As mentioned above, in the new manuscript we do not include ECMWF results due to its low horizontal resolution and the need to perform comparison with one atmospheric model in the paper.

16) **P14, Section 5: not relevant and too short to be a section.**

This section has been removed. The comparison with the PBLH characteristics over Elandsfontein site is performed in parallel with the corresponding results (PBLH diurnal and seasonal cycle) from Gual Pahari.

17) **P15 conclusion: this section is long and not conclusive, and it is repeating what have been said previously.**
We thank the reviewer for this comment. After revision of the manuscript content, the Conclusions Section has been rewritten. More emphasis is given on the factors that can explain the patterns of diurnal and seasonal PBLH cycle and contribute to the formulation of PBLH. Moreover, the sources of discrepancies between lidar and numerical estimations are discussed and suggestions for future studies are made. In the same sense, the Abstract has also been revised.