Reply to the comments of the reviewer.

We would like to thank the reviewers for finding the time to read our manuscript and for giving valuable comments to improve the manuscript. We have made modifications to the manuscript accordingly and replied to the specific comments below. The reviewer comments are **in bold** and our replies *in italic.*

Kind Regards,

Kimmo Korhonen, on behalf of all co-authors

Comments by A.D. Papayannis, *and our replies:*

1) The overlap height of the lidar system used should be provided, so as to know if it is within the PBL height or not.

The overlap information is added in the Section 3.2.

2) Some previous work on PBL studies and methods should be cited, e.g.

The citations which you suggested (Menut et al., 1999; Matthias et al., 2004; Tsaknakis et al., 2011 and Haeffelin et al., 2012) have been added into text, we thank you for bringing more information.

3) The 3.7.2 section on the PBL height retrieval is not convincing, mostly the part describing how the PBL height is calculated, as they write: "The inversion was determined subjectively using measured vertical profiles of T and RH". This kind of phrase should be replaced by stronger and more documented arguments, as we know that T and RH are not sufficient (some times) to determine the correct PBL height. Why the authors do not mention the Richardson number in connection to previous studies on the PBL determination?

We changed the method of PBL height determination for radio sounding and the SAWS model following the method introduced by Troen and Mahrt (1986) and the corrected text in Section 3.1 now reads: "For PBL height determination we used the bulk Richardson number (BRN) method (Troen and Mahrt, 1986) as follows. Potential temperature profile was calculated for each sounding according to Stull (2000). Then, using the potential temperature and wind profile, the vertical BRN profile was determined using the formula introduced by Menut et al. (1999). The critical value RiCr was set to 0.25, which is also being used in the ECMWF model. We used linear interpolation with 10-metre interval in order to have the profiles on a standard grid. The PBL height was determined to be the lowest altitude where BRN reaches the critical value."

Comments by Anonymous Referee #1, *and our replies:*

This manuscript presents analysis of the atmospheric boundary layer height over the South African site Elandsfontein using lidar and radiosonde measurements and comparisons to model data. In particular the lidar data are valuable as extensive measurements are rare in this part of the earth. However, the recommendation is major revisions of the manuscript before publication in ACP. General comments: The intention of the paper is not clear. Is the main component of the manuscript the analysis of the PBL height or the comparisons between lidar, model and radiosonde data? The authors should focus on one topic and work it out in more accurateness. The manuscript is not well structured. A better arrangement of the different topics would lead to a better comprehensibility. This point might be correlated to the missing focus of the paper. The interpretation of the data is often subjective (e.g. time shifting for comparisons between ECMWF and lidar data; statements like 'detected reliably'). Conclusions are drawn out of the manuscript without justification or links to former studies (e.g. Summary and Conclusion). Some specific comments:

Our main aim was to compare different methods for PBL height retrieval and to study how the analysis of PBL depends on the chosen method. The lidar measurement offers better temporal and vertical resolution

for measurement so first we inter-compared between traditional radiosonde and lidar (convection strength vs. vertical aerosol mixing, respectively) in PBL height determination. We wanted to study the performance of the three models via comparing their output PBL heights to observations acquired from both measurement methods. The manuscript has also been re-structured to improve the readability as follows: The methods section was re-structured and made more compact. The sub-section order in results section was re-arranged and some sub-sections were combined to decrease the amount of sub-section headings. We have also removed presentation of subjective methods, time shifting experiment and related statements, from the manuscript.

We have added references to previous studies, e.g. Menut et al. (1999); Matthias et al. (2004); Tsaknakis et al. (2011) and Haeffelin et al. (2012), to the manuscript.

Page 17410, line3-4: 'only a few….' This sentence is repeated several times in the manuscript. In my opinion the value of the lidar measurements would not be less if the authors would omit this sentence; especially as the statement is not completely stable considering the number of boundary layer studies over the northern hemisphere. Introduction: The introduction is in large parts not introductive for a PBL study (e.g. Page 17411, line 25-Page 17412, line 18 dealing with general aerosol lidar studies). References of former studies are missing.

The sentence on page 17410, line 3-4 has been omitted as suggested. References to former studies on both PBL height and lidar technology have been added into the introduction section.

Page 17410, line 18-19: Which are the many methods for measuring the vertically resolved atmospheric properties of the PBL?

For example radiosonde, lidar, sodar, mast measurement and aircraft measurement according to Seibert et al. (2000), which is being referred as "a comprehensive study on the comparison of different operative measurement methods for PBL top height". The methods are now presented in the manuscript.

Page 17412, line 4: Is the reference to Ansmann et al., 2009 the right one at this time? Perhaps it would be better to give the reference in connection to SAMUM and give a second reference for the vertically profiling (with lidar).

A reference to Ansmann et al. (2011), a manuscript about SAMUM campaigns, has been added.

Page 17412, line 8: Give a reference and explanation for SAFARI 2000.

We modified the text as follows: "dense biomass smoke event during South African Regional Science Initiative (SAFARI) 2000 (Swap et al., 2003)".

Page 17412, line 13-15: What is the intension of presenting this information? Why do the authors present the analysis of lidar ratios from an instrument (CALIPSO) not able to measure the lidar ratio independently? What is the value of this information while talking of the limitations of the elastic backscatter measurement technique in the following sentence?

The reviewer is right that this information is not relevant and, therefore, we removed this part from the manuscript.

Page 17412, line 28-29: repetition

The repeated sentence has been omitted from the text.

Page 17413, line 23: '... and the lowest and the lowest…'

Corrected

Page 17413, line 25: Which are the four major synoptic circulation types? Why are they important? They are not used anymore in the manuscript.

As the reviewer points out, this information wasn't used in the analysis and therefore the paragraph about circulation types has been removed from the manuscript.

Page 17414, line 18: Give a reference for the long-term climate statistics.

A reference to the web page of World Meteorological Organization has been added to the text.

Page 17415, line 25: '… typically…; are there others? – Be consistence with radiosonde or radio sounding

The routinely used radiosonde measurements give only the data mentioned in the text. Radiosondes are also being used for measuring vertical ozone profiles (ozonesondes) and radiation in the upper atmosphere. We chose using "radiosonde" as the name of measurement method and it has been used throughout in the text.

Section 3.2: Whow is the PBL height derived? Does the wind direction matter for the comparisons of lidar and radiosonde measurements?

Sections 3.1-3.6 now explain how the PBL top height was derived in detail for each method and model. We performed comparison for wind speed and direction between the radiosonde launch site and Elandsfontein station. We found that neither wind speed, nor direction is correlated with PBL top height differences between radiosonde and lidar measurements.

Section 3.3: Give more information and references for the ECMWF model

We have added a reference to the full documentation of the model, published by ECMWF. Also a brief general description of the model has been added to the text.

Section 3.4: Give references for the SAWS model

We added reference to Landman et al. (2012) where a description of the model can be found.

Section 3.5: What does LAPS and GASP mean?

We clarified in the text that LAPS stands for Limited Area Prediction System and GASP for Global AnalysiS and Prediction.

Section 3.6: Do the authors use level 1 or level 2 data?

We state in Section 3.7.4 that "In this study, we used the Level 2 aerosol layer product."

Page 17418, line 23-24: This sentence is not clear.

An example has been added, and the corrected sentence now reads "The comparable hourly PBL top height was calculated from lidar data using the average of the three closest data points of the hour considered (e.g. for 12:00 the PBL height would be the average of the three points between 11:45 and 12:15)."

Section 3.7.2: Not understandable. What does major period mean; more than 2 per day? What is for the other periods? What is the intension of this section?

The commented section has been rewritten due to change of method in PBL height determination. The word "major" has been changed to "long". The intention is to explain the method we used in PBL height determination and periods when radiosonde data was unavailable during 2010.

Section 3.7.3: Give references!

We have provided references to describe the measurement instruments and models in Sections 3.1-3.6 and in this section discuss how the PBL has been derived from their data. We have added the reference Troen and Mahrt (1986) to describe how the bulk Richardson number was used in PBL height determination.

Section 4.1: Why can the CALIOP data be used for interpreting the POLLYxt data? In the introduction the authors stated that CALIOP data are not so valuable for boundary layer studies.

The reviewer is correct. The statement in the introduction was misleading and therefore we removed it from the manuscript.

Section 4.2: What about daily/monthly/seasonal comparisons? What about the variability? What is the value of an annual comparison above 120 km without considering the external conditions (e.g. wind

direction)? Why shifting the lidar? What is the value of shifting the lidar data if it does not give a realistic view? What about the deviation due to temporal resolution?

We performed comparison for wind speed and direction between the radiosonde launch site and Elandsfontein station. We found that neither wind speed, nor direction has a correlation to PBL top height differences between radiosonde and lidar measurements. This is an indication that the 120 km distance does not have major impact on the analysis. We have also removed the time shifting experiment from the manuscript.

Section 4.3: The conclusions presented in this section are not clear and comprehensible. E.g. results from TAPM and SAWS model data are in the same order and no systematic deviation can be seen for the ECMWF data.

The section has been rewritten as a consequence of changing the PBL top height determination method for the SAWS model.

Section 4.4: Mainly repetitions! Line 4-6: Can this conclusion really be drawn from this data set?

Solar radiation and surface temperature are crucial factors in daily development of the PBL. We compared the measured temperature to long-term climate statistics between 1961 and 1990 and the measured temperature cycle agreed well with the 30 year average temperature, as discussed in Section 2. In addition to that, the annual PBL cycle agreed well with the measured global radiation cycle. Based on these two observations we concluded that the observed PBL characteristics of 2010 could be used as a generalization to some extent.

Line 8-10: This conclusion is not clear! What about January and December?

The measurements were started on January 27th and therefore the number of measurement days was low. The corrected text now reads "On average, the PBL top was highest in spring (September and October with heights of 2170 and 2260 m with standard deviations of 790 m and 940 m, respectively), while it was the lowest in winter (May-August with heights of 1450-1790 m) and in January (1210 m), which may be due to low number of measurement days (5). The standard deviations of monthly averages were from 17 % (January) up to 42 % (October), which indicates high variability between the PBL daily top maxima." Additionally, in the original manuscript there was a typing error, i.e. "January" had been accidentally written as "December". This was also corrected.

Line 20-23: What is the height of the stacks? Give references for this statement!

A reference to height of the stacks has been added and the statement has been corrected as "The industrial emissions are most probably released and lifted to the RL because of tall stacks. As an example, all but two power plants owned by the national power company have stacks rising up to 200 metres or more (Bethlehem and Goldblatt, 1997). Therefore the immediate effect of industrial emissions on air quality is smaller during the night, when the SBL is low."

Section 4.6: line 14: What does 'detected reliably' mean? This is a subjective conclusion.

A description of reliable detection has been added, and the corrected sentence now reads "Figure 7a shows the monthly mean difference between each comparison for days when PBL maximum height was considered to be detected reliably with the lidar, i.e. when measurement time was distributed over separate hours before and after the solar noon to detect PBL evolution and no wet removal was observed."

Line 14-15: What is meant by 'variability'? What is the 'strong anomaly in October'? From Figure 8 only the ECMWF shows a large difference.

The sentence mentioning variability has been rewritten as follows: "As is evident in Figure 7a, the tendencies of over- or underestimation in PBL top height altered between the change between the dry and the rainy season for the models. In comparison to lidar measurements, the ECMWF and the SAWS models tend to show larger overestimation during the rainy season, while TAPM gives larger values than the lidar.". The word "anomaly" has been removed from the text.

Also the following subsections have been corrected to follow this statement.

Line 17: Where does the information about the cloudiness come from? Give references! Line 17: Seasonal or monthly differences?

The cloudiness has been observed with the Polly^{XT}, as it has been mentioned in Section 4.3 of the manuscript. The word "seasonal" has been changed to "monthly".

Subsection 4.6.1: Mainly repetitions! What is meant by 'uncertainties'? Uncertainties or errors are not discussed before. It is crucial to discuss errors and uncertainties of the measurements and the used method. This is missing in this manuscript.

The sentence has been rewritten now reads "Maintenance was carried out during two months in the rainy season (January and December) resulting in a smaller dataset for that particular season, which could have affected the observations of monthly and seasonal PBL characteristics." We have added discussion of errors in using the BRN method for radiosonde and the SAWS model because we found cases where the method failed to determine the PBL top height. The discussion in Section 4.1 was accordingly.

Standard deviations have been calculated, added and discussed for all relevant numerical PBL top heights and growth rates throughout the results section.

Subsection 4.6.6: Is not the good correlation between POLLYxt and the CALIOP lidar contradictory to the statement earlier in this manuscript, that the CALIOP lidar is not suitable for PBL studies? Which CALIOP data have been used?

The reviewer is absolutely right. The statement in the introduction was misleading and therefore we removed it from the manuscript. Retrieval of PBL from CALIOP Level 1 data has large uncertainties and, therefore, we used Level 2 aerosol layer data in the analysis as described in Section 3.7.4.

Section 5: Line 3-4: Repetition!

The sentence "Only a few studies have been performed in Europe and Asia, most of them with less data coverage." has been removed from the text.

Section 6: Page 17432, line 2-3: This was not discussed in the paper!

Section 2 discusses the main sources of emissions: "The main anthropogenic emission sources in this area include high-capacity power production with coal-fired power plants (Lourens et al., 2011), yielding nearly half of all electricity produced on the African continent. In addition, there are many other industrial sources of nitrogen and sulphuric oxides, such as petrochemical industry and mining activities. The area surrounding the measurement site is globally regarded as one of the top five hotspots of nitrogen oxide emissions (Lourens et al., 2011; 2012). Other anthropogenic emissions in this area include household combustion (for space heating and cooking) and controlled, as well as uncontrolled burning of vegetation. Wildfires and controlled burning of vegetation are significant sources of particulate emissions, especially during May-September. During the measurement campaign the fire frequency in the surroundings of Elandsfontein was highest in September (the end of the dry season), and the lowest in March (the end of the wet season) (http://earthdata.nasa.gov/data/near-real-time-data/firms)."

Section 3.2 discusses, among the data coverage and lidar performance, also how rain, cloudiness and detected complex aerosol plumes depend on season: "Failures in the PBL height determination were attributed to low clouds (including fog) and thick aerosol plumes, which occasionally caused the detection of complex aerosol layer structures. The latter was caused by strong aerosol sources (e.g. originating from power plants, wildfires) in the proximity of the lidar site.".

Page 17433, line 19-23: This was not discussed in the paper!

The text has been modified to be consistent, i.e. this statement is being discussed in the results, while the summary and conclusions already include the discussed statements.

Page 17434, line 1-5: How can this conclusion be derived from this manuscript?

The PBL top height defines the volume into which particulate pollutants are distributed, as it has been discussed in the beginning of the introduction section. Also the definition we have used states that the aerosols inside the PBL are vertically mixed. Therefore unrealistic estimation of PBL top height leads into unrealistic estimation of particulate emission concentration inside it.

Comments by Anonymous Referee #2, *and our replies:*

The subject of this paper deals with the observation/diagnosis of atmospheric boundary layer top heights in South Africa using lidar and radiosonde data and the comparison with the outcome of three atmospheric models. The lidar observations are very relevant, in particular for this part of the world. However, the comparison with radiosonde data and the models calculations is troublesome and inconsistent. The authors use inconsistent and subjective methods for the derivation of the pbl depth on data and from the models. The lidar method provides basically an aerosol boundary layer depth which is known to be a good estimate for the pbl depth only during daytime in convective conditions (see also discussion by Harvey et al in QJMS). It is not correct to use this as a basis for comparison with night-time turbulent mixing depth since during nights over land the aerosol layer coincides more with the depth of the outer region (or the top of the previous daytime boundary layer depth). Also the methods utilised for estimation of the pbl on basis of radiosonde data and using the three models is not consistent. My suggestion is to use one general method for all the model and radiosonde cases using mean profiles (and perhaps surface fluxes), such as in Seidel et al on basis of the ECMWF method which itself is based on the original Troen and Mahrt method (1986) and used by many others afterwards (and was recommended by Seibert et al). Also the claim that lidar is suitable for continuous measurements for pbl depth cannot be true given the comments above, and this claim needs to be relaxed in the paper. The time shifting in section 4.2 between ECMWF and the lidar "to improve the correlation" (as the authors write), is another example of a subjective method not acceptable for a scientific paper. Finally, the authors also need to give a better account on papers which appeared in the literature on related subjects (like Summa et al and Cimini et al in recent volumes of Atmos. Meas. Tech Discuss, 2013) and should avoid the use of so many acronyms in their paper.

We changed the method of PBL height determination for radio sounding and the SAWS model following the method introduced by Troen and Mahrt (1986) and the corrected text now reads: "For PBL height determination we used the bulk Richardson number (BRN) method (Troen and Mahrt, 1986) as follows. Potential temperature profile was calculated for each sounding according to Stull (2000). Then, using the potential temperature and wind profile, the vertical BRN profile was determined using the formula introduced by Menut et al. (1999). The critical value RiCr was set to 0.25, which is also being used in the ECMWF model. We used linear interpolation with 10-metre interval in order to have the profiles on a standard grid. The PBL height was determined to be the lowest altitude where BRN reaches the critical value."

The statement that lidar is suitable for continuous measurement has been relaxed, the new statement is "is suitable for continuous measurements of daytime PBL and night-time RL with high temporal data coverage.*", and a reference to Cimini et al. (2013) has been added to the text. We have also added multiple references related to previous studies, e.g. Menut et al. (1999); Matthias et al. (2004); Tsaknakis et al. (2011) and Haeffelin et al.(2012), to the manuscript.*

The comparison between models that predict the SBL height during night-time and the lidar is carried out for only daytime values as we have stated in Sections 3.3 and 3.5.

We also dropped out the experimented time shifting from the text.

The use of acronyms for convective and stable boundary layer (CBL and SBL, respectively) and residual layer (RL) is necessary for keeping Table 1 compact. TAPM and ECMWF, despite being acronyms, are generally recognized names for the corresponding models. We have removed acronym "SA" and refer to the country

as “South Africa” throughout the text, as well as “PM” for particulate matter has also been removed. Also the acronyms “UM”, “GR” and “RH” have been removed due to infrequent usage in the manuscript.