Point-by-point response to the referees ' comments

We thank the two referees for their pertinent comments which certainly helped to improve the manuscript. The referees have brought up some important points and we appreciate the opportunity to address them based on a point-by-point response to their comments, which follows below (in black the original comments from the referees and in red our responses).

Referee # 1

First results are presented for a case study, which look very promising. Overall, it is a very impressive effort. The paper is well written and clearly structured. The English is very good. In my opinion, this paper deserves publication in ACP. I only have minor modifications to suggest.

We are very pleased for the enthusiastic words expressed by the referee. We also appreciate the possibility for further improve the manuscript based on his/her precious suggestions.

Minor comments. The potential for BASIL to measure profiles of moments of turbulent fluctuations in the PBL is of great interest to the community. In the paper, only a short sample of the dataset acquired during HOPE (2 hours) under "ideal" meteorological conditions is presented. I am convinced that such data acquired in more complex situations would be extremely worthy, and that obtaining from an instrument such as BASIL the diurnal evolution of profiles of higher-order moments of turbulent fluctuations would be invaluable for understanding PBL processes associated with night-day transitions, PBL development under cloudy situations and model validation.

We agree with the referee about the enormous value and potential scientific impact of having measurements of the diurnal evolution of higher-order moments of turbulent fluctuations in clear sky conditions, as well as in more complex meteorological situations. For this motivation, after the publication of the present manuscript aimed to demonstrate the capability of our Raman lidar to characterize turbulent processes within the convective boundary layer, we plan to possibly extend the analysis to a variety of clear-sky and cloud-topped convective boundary layer conditions from different field deployments and include these additional results in a forthcoming paper.

At the end of the paper, the authors provide some leads concerning future studies to be undertaken with the BASIL dataset along these lines. I think it would be worthwhile discussing in the paper (maybe in the final section, 1 short paragraph) to what extend the temporal resolution of the flux profiles retrievals could be reduced to hourly (or less) temporal resolutions and if a realistic diurnal evolution of these variables can be extracted from the impressive HOPE dataset.

We agree with the referee about the opportunity to discuss in the paper the feasibility of measurements of turbulent processes with increased temporal resolution which can be obtained by reducing the time window for the application of the auto-covariance approach from two hours to one hour or less, possibly allowing to follow the diurnal evolution of the convective boundary layer. So far attention has been focused only on those time segments characterized by a well-mixed and quasi-stationary convective boundary layer, characterized by a stable or almost stable boundary layer height, which corresponds to the period of its maximum development around local noon. In this regard, it is to be noticed that the selection of a shorted time interval (one hour or less) with respect to the one presently considered (two hours) would lead to an increase of the sampling error, i.e. the error incurred when the statistical characteristics of a measured population are estimated from a subset of the real population (the smaller is the sub-set, the larger is the sampling error). In the future this can be overcome by averaging turbulence profiles measured under similar meteorological conditions. This requires continuous measurements during field campaigns and even more the application of our methodology to data acquired by operational lidar systems such as those present at the ARM Southern Great Plains (SGP) site in Oklahoma (USA), at DWD in Lindenberg

(Germany), and at Meteo-Swiss in Payerne (Switzerland). An example for a pioneering campaign applying these techniques is the Land-Atmosphere Feedback Experiment (LAFE) to be performed in August 2017 at the SGP site (see www.arm.gov/publications/programdocs/doe-sc-arm-16-038.pdf).

Additionally, the selection of a well-mixed and quasi-stationary convective boundary layer characterized by a stable height allows for the use of a simple data de-trending approach based on the application of a linear fit to the atmospheric variable time series and the following removal of the mean value. This approach cannot be applied in the presence of high-frequency variations associated with synoptically induced transients such as frontal passages or dry-lines. However, we believe that a modified version of this de-trending approach will be effectively applicable to also remove influences associated with large-scale advection, synoptic processes, and, ultimately, the diurnal cycle and the presence of a cloud-topped convective boundary layer.

All these aspects are now clearly discussed in the final section of the paper, where the following paragraph has been introduced: “As a final remark, we need to specify that we foresee the possibility to apply this approach to characterize the diurnal evolution of turbulent processes within the convective boundary layer, by monitoring the changing patterns of water vapour and temperature higher-order moments during its different evolution phases, including day-to-night and night-to-day transitions, possibly with increased temporal resolution. However, measurements of turbulent processes with increased temporal resolution obtained by reducing the time window for the application of the auto-covariance approach to one hour or less would lead to an increase of the sampling error. This can be overcome by the analysis of continuous measurements such as those carried out at observatories like the ARM Southern Great Plains (SGP) site in Oklahoma (USA), the DWD Meteorologisches Observatorium in Lindenberg (Germany) and the Meteo-Swiss Centre for Meteorological Measurement Technology in Payerne (Switzerland). Furthermore, more data will become available via new field campaigns. An example for a pioneering campaign applying these techniques is the forthcoming Land-Atmosphere Feedback Experiment (LAFE, Wulfmeyer and Turner, 2016) to be held at the SGP site (see www.arm.gov/publications/programdocs/doe-sc-arm-16-038.pdf). Finally, in order to apply the auto-covariance approach to characterize the diurnal evolution of turbulent processes in clear sky conditions, as well as in more complex meteorological situations, i.e in the presence of large-scale advection, synoptic processes and cloud-topped convective boundary layer, modifications in the de-trending approach considered in the present paper are required.”

P6, l21: I am aware that moisture profiles can be extracted from GPS tomography. Can this also be done for temperature? Or is it just a cut/paste from the water vapour paragraph above?

Temperature profiles cannot be extracted from GPS tomography in the lower troposphere and reference to this capability was only an oversight. The corresponding sentence has been corrected and now reads: “These two calibration constants can be determined through the comparison of the lidar signal ratio with simultaneous and co-located temperature measurements from different sensors (e.g., from radiosondes, microwave radiometers, etc.)”.

P17, l26: “starting”

Corrected. Now the sentence reads: “… with cirrus clouds starting from 15 UTC”.

P18, l18: “aerosols property”

Corrected. Now the sentence reads: “… by exploiting aerosols property to act as atmospheric tracers”.

P19, l24-25: “[: : :] as in fact: : :. column.” This sentence is unclear, please rephrase.

The sentence has been rephrased as follows: “This property makes lidar systems much more suitable for studying turbulence statistics than in-situ sounding systems. In fact, because of the capability of the formers to monitor the vertical air column above the station as opposed to radiosondes, undergoing a horizontal drift during their ascent and consequently a deviation from the vertical, lidar systems guarantee the capability to measure turbulence statistics within the turbulent eddies involved in the boundary layer mixing processes.”

Figure 9: it is a bit odd to see that a large part of the signal is off-scale on this figure. Why is it so? This should be improved.

We improved figure 9 by changing the full-scale value from 250 to 800 s and now values in the lower portion of the CBL are within the scale. The text describing the figure has also been slightly changed as follows: “The integral time scale of both water vapour mixing ratio and temperature fluctuations is found to have large values (also in excess of 200 s, up to 500 s for water vapour) in the lower portion of the CBL up to @ 750 m (i.e. z/z_i <0.6).”

Figure 13: same comment as above, especially for Fig. 13b.

We improved figure 13 by changing the full-scale values both in panel a and b. Full-scale values are now -15/+15 for panel a and -5/+25 for panel b.

Referee # 2

In the abstract is missing a sort of "punch line". Why the measured values in terms of water vapor and temperature are important and who is going to use those measurements?

We agree with the referee that a clear statement of the relevance and potential impact of these measurements, already present in the introduction, should also appear in the abstract. For this purpose, the following sentence was introduced in the abstract: “These measurements, in combination with measurements from other lidar and in-situ systems, are important to verify and possibly improve turbulence and convection parameterization in weather and climate models at different scales down to the grey zone (grid increment ~ 1 km) (Wulfmeyer et al. 2016)”.

The retrieved planetary boundary layer height in the manuscript is 1290+/-77. I know that this number comes from the standard deviation calculation, but which is the physical meaning of this value as the lidar resolution is discrete? Doesn’t make more physical sense to round up to the closest available range bin?

The approach used in the present research effort to estimate the planetary boundary layer height is based on the detection of the strongest gradient in the particle backscatter coefficient at 1064 nm. The signal strength of the elastic signal at 1064 nm is much larger than the strength of the vibrational and rotational Raman signals, which have used to determine water vapour mixing ratio and temperature profiles. As a result of this larger strength, the elastic signal at 1064 nm is acquired with a much higher vertical resolution (7.5 m) than the one considered for the water vapour and temperature measurements, i.e. 90 m and 30 m, respectively.

The standard deviation of ±77 m associated to the mean boundary layer height (1290 m) is aimed to indicate the boundary layer height variability associated with atmospheric fluctuations, which is larger than the vertical resolution. Based on these considerations, we don’t see any physical reason for having the standard deviation and the mean of this measured quantity rounded up to the closest available range bin, as long as these values are expressed with the same number of significant digits as the vertical resolution. By the way the mean boundary layer height (1290 m) is an exact multiple value of the range resolution (7,5 m), so for this mean value rounding up to the closest available range bin is not required; on the other hand, the boundary layer height standard deviation (±77 m) is 2 m off an exact multiple value of the range resolution (75 m). However, in order to remove any residual skepticism by the referee, we corrected the standard deviation value from ±77 m to ±75 m both times this was appearing in the text (i.e. page 2, line 6 and page 19, line 4), this change having negligible effects on the results reported in this paper.

Pag 2 Line 30 please read "unstable"

Corrected as requested by the referee.

Pag 7 Line 1-3. It is not clear if the calibration coefficients were calculated globally or for each single radiosonde launch.

As already specified in the text, during HOPE the Raman lidar was calibrated based on the comparison with the radiosondes launched approximately 4 km away. We considered all clear sky radiosonde launches coincident with lidar operation (sixty in total), thus determining sixty values for each calibration coefficient. A mean value for the calibration coefficient was then estimated from these sixty values. This is now more clearly specified in the text, where the corresponding sentence has been modified as follows: "All clear sky radiosonde launches coincident with lidar operation (sixty in total) were considered, thus determining sixty distinct values for each calibration coefficient. A mean value for each calibration coefficient was then estimated and used throughout the HOPE period."

Page 11 Line 25, and in all the manuscript: please change leakage with "cross-talk".

We agree with the referee that the term "cross-talk" is more appropriate than the term "leakage" to describe the optical effect taking place in our Raman lidar receiving system. In fact, as most Dictionaries confirm (among others, Oxford Dictionaries, Collins English Dictionary, American Heritage Dictionary), the term "leakage" is meant to indicate the "undesired escape or loss of a quantity, such as electromagnetic radiation, electric charges, etc.", the lost quantity not being necessarily transferred to a different channel, while "cross-talk" indicates the "unwanted transfer of signal from one channel to another". In our case, we don't have an escape or loss of a portion of the elastic signal photons entering the elastic channel, with a consequent reduction of the strength of the sampled elastic signal, but the undesired entrance of a small portion of the elastic signal photons (which would not be entering the elastic channel in any case) into the low-quantum number rotational Raman channel. As a result of this, the amplitude of the sampled elastic signal is not reduced, but indeed there is an undesired transfer of elastic signal to the low-quantum number rotational Raman channel. The term "leakage" has been substituted with the term "cross-talk" throughout the text.

Page 13 Line 18 Most of the backscattering values are less than 1 10-6 m-1 sr-1

The referee is right: most particle backscattering values within the boundary layer are less than $1x10^{-6}$ m^{-1} sr^{-1} and not in excess of this value, as erroneously stated in the text. Values were erroneously reported as a result of an oversight in reading the color scale. Instead, particle backscattering values are between 0.3 and 1.4 $x10^{-6}$ m^{-1} sr^{-1}. This oversight has been corrected in the text, where the corresponding sentence now reads: "The figure reveals the presence of a significant aerosol loading within the boundary layer (with values of β_{par} in the range 0.3-1.4$x10^{-6}$ m^{-1} sr^{-1}), tracing the presence …".

Page 18 Line 21 I suggest adding two more works that will give an exhaustive view of PBL retrieval methods by lidar means. The works I suggest are: Haeffelin, M., Angelini, F., Morille, Y. et al. Boundary-Layer Meteorol (2012) 143: 49. doi:10.1007/s10546-011-9643-z and Conor Milroy, Giovanni Martucci, Simone Lolli, et al., “An Assessment of Pseudo-Operational Ground-Based Light Detection and Ranging Sensors to Determine the Boundary-Layer Structure in the Coastal Atmosphere,” Advances in Meteorology, vol. 2012, Article ID 929080, 18 pages, 2012. doi:10.1155/2012/929080

The two additional papers suggested by the referee are now cited in the manuscript and have been added to the reference list.

Pag 18 Line 27 see remarks in the abstract section

Here the referee is pointing to the mean and standard deviation values for the CBL height published by other authors (Muppa *et al*., 2016) and which are cited here only for comparison with ours. We believe these values should appear here as they appeared in the literature.

Pag 24 Line 10 Due to the intrinsic properties of lidar measurements, are the retrieved variables (water vapor and temperature) normal distributed (and then does it make sense to define 3rd and 4th moment?) Please clarify

A geophysical time series such as water vapour and temperature fluctuations are generally not normal distributed because these are determined by non-linear turbulent processes. It is exactly our goal to specify the distribution of the turbulent fluctuations because they give us unique insight in the turbulent process understanding and its parameterization in models. The third- and fourth-order moment are variables closely related to skewness and kurtosis, respectively. Specifically, the third-order moment quantifies the degree of asymmetry of the distribution of the considered atmospheric variable fluctuations, with positive values indicating a right-skewed distribution (with the mode smaller than the mean) and negative values indicating a left-skewed distribution (with the mode larger than the mean). Analogously, the fourth-order moment indicates the steepness of the distribution and the width of its peak. While the vertical profiles of skewness and kurtosis for water vapour mixing ratio and temperature fluctuations are reported in the paper (page 26 and figure 13), the paper primarily focuses on the discussion on the third- and fourth-order moment of these atmospheric variables fluctuations. The choice of focusing the discussion primarily on the third- and fourth-order moment is motivated by the fact that the third- and fourth-order moments have more regular vertical profiles - characterized by a smaller number of spikes and spurious fluctuations - than skewness and kurtosis. This is because skewness and kurtosis are obtained from the third- and fourth-order moment through normalization by the atmospheric variance and, consequently, spikes and spurious fluctuations may occur in those altitude regions of the convective boundary layer where atmospheric variance has very small values. However, in the entrainment zone, atmospheric variance for both water vapour mixing ratio and temperature is characterized by large values and this makes skewness and kurtosis for both atmospheric variables very stable and accurately determinable within this altitude region. In the entrainment zone values of skewness for both humidity and temperature fluctuations are found to be large, as expected for fluctuations strongly deviating from a normal distribution. This aspect is now more clearly specified in the text, where the corresponding sentences have been changed as follows: in page 26, lines 7-12, “Values of water vapour mixing ratio skewness are in very good agreement with those reported by Wulfmeyer *et al*., 2010, with positive values (up to 1.5) in the lower portion of the CBL (up to 800 m, i.e. z/z_i =0.65), with negative values (down to -1) in the middle and upper portion of the CBL ($800<z<1290$ m, i.e. $0.65<z/z_i<1.00$). Large positive values are found within the entrainment zone and just above the CBL top (with a maximum of approx. 5 at 1400 m, i.e. z/z_i =1.15), which testifies the presence of humidity fluctuations strongly deviating from a normal distribution”; and in page 26, lines 20-26, “Temperature skewness has a negative peak (~ -4) at 500 m, i.e. z/z_i =0.4, positive peaks (~ 8 and 4) at 890 and 1100 m (i.e. z/z_i =0.7 and z/z_i =0.85), respectively, and negative values within the entrainment zone and just above the CBL top (with a peak value of ~ -7 at 1460 m, i.e. z/z_i =1.13), in in good agreement with measurements reported by Behrendt *et al.* (2015, 40-120 s) for the nearby site of Hambach, but a different case study. Again, values of skewness within the entrainment zone and just above the CBL top are found to be large, as expected for temperature fluctuations strongly deviating from a normal distribution". Furthermore, as specified in the Abstract (page 2, line 16) and in Section 4.3 (page 26-27, lines 27-3), in the upper portion of the CBL, water vapour and temperature kurtosis are found to have values close to 3, which is an indication of normally distributed humidity and temperature fluctuations (mesokurtic-Gaussian distribution), while in the entrainment zone and above the CBL top water vapour and temperature kurtosis show values as large as 18 as a result of the presence of humidity and temperature fluctuations strongly deviating from a normal distribution. The corresponding paragraph has been changed as follows: "Values of kurtosis in the upper portion of the CBL (in the height interval 1160-1280 m) are in the range 2.76-3.83, with a mean value of 3.36, for water vapour mixing ratio fluctuations, while they are in the range 2.68-3.45, with a mean value of 3.17, for temperature fluctuations. These values indicate normally distributed (mesokurtic-Gaussian distribution) humidity and temperature fluctuations in the upper portion of the CBL (Wulfmeyer *et al.*, 2010; Turner *et al.* 2014a, Behrendt *et al.*, 2015, Muppa *et al.*, 2016). In the entrainment zone and above the CBL top, values of water vapour mixing ratio and temperature kurtosis are found to be large (up to 18) as a result of the presence of humidity and temperature fluctuations strongly deviating from a normal distribution."

For what concerns the reference to the "intrinsic properties of lidar measurements", here we believe the referee refers to the application of Poisson statistics to lidar signal photon counts. In fact, photon counts are measured by our Raman lidar through a photon counting unit, while "virtual" counts are measured through an analog module (e.g. Newsom *et al.*, 2009). For lidar measurements carried out in the convective boundary layer, as those presented in this paper, the number of photon counts and "virtual" counts detected by the receiving system is large enough to allow the approximation of count statistics with a Gaussian distribution. In fact, for an increasing number of photons, Poisson distribution becomes more and more symmetric, and begins to resemble a Gaussian shape. The Gaussian distribution is found to adequately approximate the Poisson distribution when the mean number of photons is in excess of 10 (among others, Russo, 2007). Higher moments of the noise distribution can also be studied with the tools introduced in Lenschow et al. 2000. Our current results do not indicate a significant deviation of the noise distribution from Gaussian.

Figure 4, 5 and 7 are not very clear and the resolution should be improved.

In these figures we improved the graphical resolution to make them clearer.

References

Newsom, R. K, Turner, D. D., Mielke, B., Clayton, M., Ferrare, R., Sivaraman, C.: The use of simultaneous analog and photon counting detection for Raman lidar. Appl. Opt., 48, 3903–3914, doi: 10.1364/AO.48.003903, 2009.

Russo, F.: An Investigation of Raman Lidar Aerosol Measurements and their Application to the study of the Aerosol Indirect Effect, PhD dissertation, Physics Department, University of Maryland, Baltimore County, 2007.

Wulfmeyer, V., Muppa, S., Behrendt, A., Hammann, E., Späth, F., Sorbjan, Z., Turner, D.D., and Hardesty, R.M.: Determination of convective boundary layer entrainment fluxes, dissipation rates, and the molecular destruction of variances: Theoretical description and a strategy for its confirmation with a novel lidar system synergy, J. Atmos. Sci., 73 (2), 667-692, doi: 10.1175/JAS-D-14-0392.1, 2016.

Wulfmeyer, W., and Turner, D. D.: Land-Atmosphere Feedback Experiment (LAFE) Science Plan, DOE/SC-ARM-16-038, www.arm.gov/publications/programdocs/doe-sc-arm-16-038.pdf, 2016.