Dear reviewers,

We like to thank you for your helpful comments on our paper titled “Retrieval of aerosol backscatter and extinction from airborne coherent Doppler wind lidar measurements”.

The original comments are in bold, followed by our replies. An additional file, highlighting the changes introduced in the manuscript, was also uploaded.

Reviewer #1

1. **Page 1940, line 25: the letter L used here for designating the line-of-sight of the lidar is also used for the atmospheric layers later in the document. Two different letters would facilitate the reading of the article.**

   The line of sight vector letter was changed to I to improve the readability.

2. **Page 1941, line 17: the equation linking $\mathbf{f_D}$ and $\mathbf{v_{LOS}}$ (the Doppler equation) has no minus sign. The usual practice for lidars is to count $\mathbf{v_{LOS}}$ positive when the wind is blowing away from the lidar. In that case, there is a minus sign. The authors should clarify which sign convention is used here for $\mathbf{v_{LOS}}$.**

   Because the system was mounted in an aircraft and pointing downwards, the sign convention was chosen to match the usual meteorological convention (positive for upward winds).

   Clarification introduced:

   A positive frequency shift f_D indicates a positive relative speed v_{LOS}, which, in turn, indicates that the scattering aerosols are moving towards the lidar. For the case of an airborne downward pointing lidar, this sign convention leads to positive relative speeds for upward winds and negative relative speeds for downward winds.

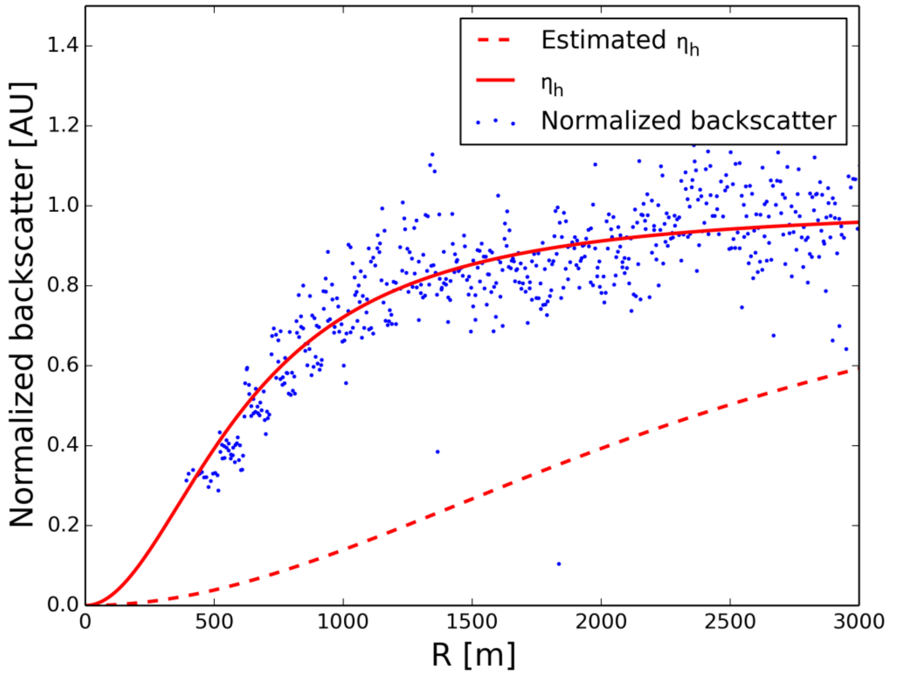
3. **Page 1944, line 10: the value of $\mathbf{K}$ should be given as the width of the rectangular window used for estimating $\langle \mathbf{P(R)} \rangle$ from the $\langle \mathbf{\hat{P}_S(R,k)} \rangle$ may have an impact on the quality of the estimation. The width must be large enough so the entire return spectrum is inside the window.**

Clarification introduced:

The optimal value for the integration window width K is the one that exactly matches the return pulse spectral width. A shorter integration window will lead to an underestimation of the backscattered power, while a longer integration window increase the estimation error due to the integration of measurement noise. Base on this facts, the integration window width K was set to be 6 (approximately 6 MHz).

4. **Page 1945, lines 4-5: considering equation (9), one can easily see that the heterodyne efficiency varies significantly in the range domain covered by the airborne lidar (see figure below). The argument that "no effects produced by the range dependency of the heterodyne efficiency were observed in the received atmospheric signal" shall be more elaborated. How the authors did come to this conclusion? Why not just correct lidar signals from the range dependence of the heterodyne efficiency with eq. (9)?**

This conclusion was achieved based on the calculation of the expected heterodyne efficiency using the DWL system parameters listed in Table 1 and the measurements obtained during changes in the aircraft altitude. For a matched Gaussian afocal monostatic lidar with a beam radius (1/e2) of 4 cm (Henderson et al., 1993), the heterodyne efficiency is expected to change over the whole measurement range. If this is true, the range corrected backscatter signal corresponding to a range gate on the top of the SAL would significantly change if its distance to the lidar changes from 3500 m to 3000 m (for example due to changes in the aircraft altitude). This change was not observed. Motivated by this comment and extending the same concept, we used the backscatter signal corresponding to the top of the SAL (which is quite homogeneous) to estimate a new heterodyne efficiency function based on real measurements for a range variations between 500 m and 3000 m. The results gave us the following heterodyne efficiency function:



Because most of the measurement were conducted flying at more than 3500 m above the top of the aerosol layers (Table 2), this new result doesn't change the previous results. Nevertheless, it can be useful to extend the retrieval process to lower aircraft altitudes or even for ground based measurements.

Clarification introduced:

The heterodyne efficiency function was moved to the Instrumental corrections section and an estimation based on measurements was added. Further details about the introduced changes can be seen in the attached file.

5. **Page 1946, line 11: To me, the overlap is contained in the heterodyne efficiency. I do think it should be mentioned here.**

   The nomenclature was changed to include the overlap in the heterodyne efficiency.

6. **Page 1955, lines 6-8: the authors should explain a little bit more how the different aerosol layers were determined from POLIS data? If it is a manual determination, they should clearly say it. If there is a chance this separation between different layers can be automated, they should write it. In practice, the**

need to distinguish the aerosol layers could be a strong limit for the proposed method.

The classification of the different aerosol layers from POLIS lidar measurements is based on a classification scheme described by Groß et al., 2013 using the lidar ratio and the particle linear depolarization ratio. Based on the different value pairs of the lidar ratio and the particle linear depolarization ratio different aerosols can be classified.

Clarification introduced:

This classification is based on measurements of the lidar intensive properties, the lidar ratio and the particle linear depolarization ratio. The classification scheme is described by Groß et al., 2013.

7. **Page 1955, lines 19-22: From figure 12 and table 3, it is not obvious that different values shall be considered for the aerosol layers. Did the authors tried to use a single value? What were the consequences on the result? If it happens that a single value can be considered, it is a good news for it facilitates the practical use of the method.**

Although the calculated values are similar, it seems to be casual. The calibration constants include depolarization and wavelength dependence effects which are dependent of the aerosol type. The extinction correction requires also an assumption of the lidar ratio which is also dependent of the aerosol type and layer.

Clarification introduced:

Although the calculated calibration constants $\boldsymbol{k}$ *for each aerosol type are very similar, this result seems to be just casual. Each calibration constant (Eq. 20) includes depolarization effects* $\boldsymbol{k_\delta}$ *and the wavelength dependency of the backscatter coefficient* $\boldsymbol{k_\beta^{532\to2022}}$ *which are strongly dependent of the aerosol type. On the other hand, the retrieval of extinction corrected backscatter coefficients profiles still requires the definition of aerosol layers with different lidar ratios to perform the extinction correction. For these reasons, and even though the retrieved calibration constants are similar in this case, the use of different calibrations constants and layers cannot be avoided.*

8. **Page 1966, table 3: the units of μ and σ should be given.**

   Units for the mean and standard deviation of the difference were included in the description of Tables 3 and 4.

Reviewer #2

9. **I might recommend a slight change in title. When I started to read the article I expected the focus to be on retrieval of backscatter and extinction coefficients in the infrared where the Doppler lidar operates. I have no problem with the focus on the visible – that's where satellite and many ground-based lidar operate, so augmenting that data set with airborne Doppler measurements is certainly worthwhile.**

   We would like to keep the title as is. A clarification of the wavelength of the retrieved profiles was added in the abstract.

10. **Page 1937, line 29: Coherent detection is not inherently insensitive to spectrally broad signals – rather coherent Doppler lidars are designed to have a narrow bandwidth matched to spectrally narrow aerosol signals to increase sensitivity for that regime. One could design a coherent lidar for signals of any bandwidth.**

    This is correct. Coherent detection was changed by coherent DWLs to take this difference in account and a small sentence was added.

    Clarification introduced:

    The latter is a consequence of the DWLs design to match the spectrally narrow aerosol return signal to increase the SNR.

11. **Page 1938:, line 11: It seems to me that the concept of overlap is more generally applied to direct detection lidars, where the design of the system typically isn't diffraction limited. A more appropriate terminology might be "antenna efficiency", which more appropriately describes the loss of sensitivity, relative to ideal, resulting from backscattered signal/BPLO coherence mismatches at ranges in the near field of the telescope. This is the nomenclature used by Henderson et al, 2005. I note that later on in the text the authors use "heterodyne efficiency" as reflecting phase and amplitude mismatches between**

the signal and LO. In this nomenclature heterodyne efficiency would include antenna efficiency plus other effects (e.g., LO beam heterogeneities).

The reference to the overlap function was removed and included as part of the "heterodyne efficiency". Further details about the heterodyne efficiency issue are already discussed in the answer 4. Full overlap regions are now named far field and non-full overlap near field.

12. **Page 1944, line 5: The sentence on subtracting the noise floor calls to mind that there is no mention of uncertainties resulting from the estimation process. Coherent lidar returns are characterized by speckle; with LO shot noise adding an additional noise term Although the high prf of the system probably minimizes the speckle noise term, a paragraph or two discussing uncertainties in estimation of power from coherent lidar returns, and why this is or is not a factor in these measurements, would be helpful, I think.**

    A short mention to the speckle problem is now introduced before Eq. (6) to justify the averaging of independent pulses. This gives also some context to Fig. 2, which shows that the received power follows an exponential PDF. A second short paragraph discussing the reduction of the power mean estimation standard deviation as a function of the average number of shots was also included.

    Comment introduced:

    Because each power spectra $\langle \hat{\boldsymbol{P}}_{\boldsymbol{S}}(\boldsymbol{R},\boldsymbol{k})\rangle$ *is calculated based on the average of several shots and the received power for a single shot follows an exponential probability density function, the mean received power* $\langle \boldsymbol{P}(\boldsymbol{R})\rangle$ *can be modeled as a gamma function. If 500 shots are averaged, the resulting average received power relative standard deviation is lower than 5%.*

13. **Page 1948, line 10: Again, I think overlap is better described as antenna efficiency.**

    Corrected according answers 4 and 11.

14. **Page 1952, line 18: As I understand it, the Raman returns are used as an HSRL to estimate extinction and backscatter. Perhaps some discussion on the effect**

of using 607 nm to characterize optical properties at 532 nm would be appropriate here.

Comment introduced:

A possible wavelength dependence between the Raman-shifted wavelengths and the elastically backscattered wavelengths is considered into this methodology, but as both the Saharan dust aerosols as well as marine aerosols are large compared to the lidar wavelength, the wavelength dependency can be neglected in this study.

15. **Page 1949, Eq. 17: Is there an extra R term in this equation?**

    The extra R term was removed.

16. **Page 1954, lines 15-30: The text states, without much substantiation, that "the contribution of the marine aerosol to the measured total AOD is lower than the contribution of the mixed layer" and later, that contribution of the dust layer to the total AOD is much larger than the contribution of the other two layers". Since this inability to separate the layers for this calculation is an inherent weakness of the technique, I think these statements require more discussion and substantiation.**

    The presented method for the estimation of the wavelength dependency of the extinction is a first approach to the problem and it is independent of the rest of the inversion method. Other approaches based on the simultaneous measurements of AERONET and POLIS or the use of the OPAC model/database in combination with in-situ measurements are planned to be used in future applications of this inversion method.

    On the other side, a sensitivity analysis was performed in order to analyze the effect of the high uncertainties of the wavelength conversion coefficient. The results, after using conversion coefficients 20% higher and 20% lower, are still meaningful for cases in which the aerosol layering and AOD is similar to the calibration condition (Table 4).

17. **Page 1955, line 22: It isn't clear to me why S is constant over each layer. Is there some averaging or filtering that I'm missing?**

The lidar ratio is an intensive lidar property and thus only dependent on the aerosol type. If the aerosol type does not change within an aerosol layer it should be constant throughout the layer.

18. **Page 1956, line 6: The bulge in the layers in Figure 13 at 00:10 is interesting – looks like some upward transport of the whole layer caused by convection. Maybe some comment here would be interesting.**

    According to the suggestion, a short paragraph was added. It’s nice to illustrate the potentials of synchronous backscatter and wind retrieval.

    Comment introduced:

    For the measurements corresponding to the time period between 00:05 UTC and 00:20 UTC, a perturbation of the internal structure of the SAL can be observed in coincidence with the presence of clouds on the top of the mixed layer. The vertical wind speed, also available from the DWL, shows a relatively constant upward wind flow with a mean speed of 0.3 m s-1 above the cloud layer, which is likely to be associated with convection processes.

19. **Page 5.5, line 20: If the mixed layer was a combination of a marine and Saharan dust layer, one might expect the histogram to fall between that of these two layers, but in fact it is more skewed to the right than either of the other layers. Any thoughts on this?**

    The histogram, which represents the difference between the estimation from the DWL and the measurements from the POLIS lidar for each layer, includes the errors from many different sources. For example, if the extinction of the dust layer is underestimated and the extinction of the mixed layer is overestimated (uncertainties in the lidar ratio or in the wavelength conversion coefficient), the systematic deviation of the mixed layer will not fall between the other two.