Response to Referee #1:

Blue italic font means the authors' answer, standard, *black text symbolizes the reviewer's comments, and red italic* font *means manuscript changes, while black italic font symbolizes the original content of the manuscript.*

I appreciate all the revisions. But I still think the manuscript still failed to meet the requirement of publication in ACP. Firstly, although better than previous version, the manuscript was still written poorly and should further undergo extensive English revisions. Many measurement platforms were used in this article such as ground in-situ measurements, ground lidar and airborne in-situ measurements to present the comparison of lidar retrieved optical properties with airborne in-situ measurement based ones, but the details in the comparison are very rough, which might lead to lots of errors in either the calculation of model optical properties or the lidar retrieved optical properties. How can they be comparable? Does the comparison make sense? Besides, I would suggest authors to address other major weaknesses, which are listed below.

Response:

We gratefully thank the anonymous referee for his comments and input. The mentioned points are addressed below. We updated lengthy sentences and revised them in terms of grammar.

1) Line 15: In the abstract, the author use "The study highlights the complexity of …", which is a well-known issue and is what we really want to solve. Through the whole manuscript, I can only find the author highlight the complexity without an actual solution, and there are few innovations for both results and the method during the comparison process.

Since the manuscript type is changed to "measurement report", we argue that providing a definite answer to resolve the uncertainty is not required for the scope of this journal's manuscript type. Also, due to the complexity of aerosol and missing information of the atmospheric column, no definite answer can be given. Nevertheless, the data set and manuscript are valuable sources for planning future campaigns and highlight, e.g., the know dependence of the LR to RH or addressing the mentioned issues by selecting appropriate instrumentation and shows demand for new instrumentation that can resolve these issues.

2) Line 300: replace "2.1.2 Ground-based remote sensing" by "3.1.2 Ground-based remote sensing

We updated as requested.

3) Line 316: "During the daytime, the signal-to-noise ratio in the Raman channels is too weak …" and the author use constant LR to retrieve aerosol optical properties, which will lead to huge errors especially for multi-wavelength lidar, different observation sites and experiment dates. The lidar data at night are free of the noise problem. Why not try to using these data to calculate the LR?

Yes, determining the LR during the night could have been a method to derive the LR for the column. However, at 1064 nm, the lidar ratio could not be measured at this time (first approaches have just evolved, e.g., Haarig et al., 2018); thus, we used corresponding values provided in the literature (45 ± 15 sr; Mattis et al., 2004).

Furthermore, if possible, the utilized lidar ratios have been validated by comparing the height-integral of the aerosol particle light extinction coefficient at 532 and 355 nm with the corresponding AOD measured by the deployed sun-photometer. However, for June 26 and 17, 2015, no sun-photometer data is available due to cloud coverage during these days. For June 28, 2015, the integral of the mean aerosol particle light coefficient between 0 and 2500 m and 8 to 10 UTC (below the overlap height, the values are linearly extrapolated to the ground) is 0.13 at 355 nm and 0.072 at 532 nm. The corresponding AOD(355 nm), extrapolated with the Angstrom between 340 and 380 nm, is 0.14 and 0.097 at 532 nm (extrapolated between 500 and 675 nm). However, since no sun-photometer data for validation is available for June 26, and 17 and the lidar ratios provided by Mattis et al. (2004) seem to fit reasonably well for June 28, 2015, we decided to stick with the lidar ratios provided by Mattis et al. (2004).

Hence, we added (342 - 349): "Directly deriving the LR from nighttime observations with the Raman-Lidar would also have been a feasible approach. However, as the atmospheric conditions between night and daytime were not homogenous and quite variable, we could not apply the nighttime finding to our daytime observations. However, we used AERONET AOD data to validate our extinction profiles and found good agreement whenever atmospheric conditions allowed. E.g., for June 28, 2015, the integral of the mean aerosol particle light coefficient between 0 and 2500 m and 8 to 10 UTC (below the overlap height, the values are linearly extrapolated to the ground) is 0.13 at 355 nm and 0.072 at 532 nm. The corresponding AOD(355 nm), extrapolated with the Ångström exponent between 340 and 380 nm, is 0.14 and 0.097 at 532 nm (extrapolated between 500 and 675 nm). Thus, we believe the used lidar ratio values are well justified."

References: Haarig, M., Ansmann, A., Baars, H., Jimenez, C., Veselovskii, I., Engelmann, R., and Althausen, D.: Depolarization and lidar ratios at 355, 532, and 1064 nm and microphysical properties of aged tropospheric and stratospheric Canadian wildfire smoke, Atmos. Chem. Phys., 18, 11847–11861, https://doi.org/10.5194/acp-18-11847-2018, 2018.

4) Line 344: replace "2.1.3 Airborne in-situ measurements" by "3.1.3 Airborne in-situ measurements"

We updated as requested.

5) Line 479-480: "… and slightly lower than… That indicates different aerosol populations in these layers" These might also result from errors during the calculation, such as the determination of

refractive index and lidar ratio and the particle size distribution range used in Mie model, which can't be ignored and determine whether the comparison results were meaningful.

Yes, the different aerosol population is part of the uncertainty. However, based on the measurement setup, some important aerosol properties, e.g., the refractive index, could not be determined within the atmospheric column. As a result, the considerable uncertainty in the LR should cover a broad range to cover changes in aerosol type. Also, the lidar ratio can be ignored when comparing the backscatter coefficient, and only uncertainties in the aerosol particle number size distribution and chemical composition impact the modeled values. Hence, the measurement and comparison are meaningful since the report highlights the challenges of vertical in-situ measurement-based studies with these considerable uncertainties.

6) Line 460, 515: The figure c), d) and e) are not easily readable. Some legends can't be found in the figure. And 7) Line 515: The same as Line 479-480.

We updated the figures and added corresponding legends in panel a) and b). Accordingly, we updated Figures 7 and 8.

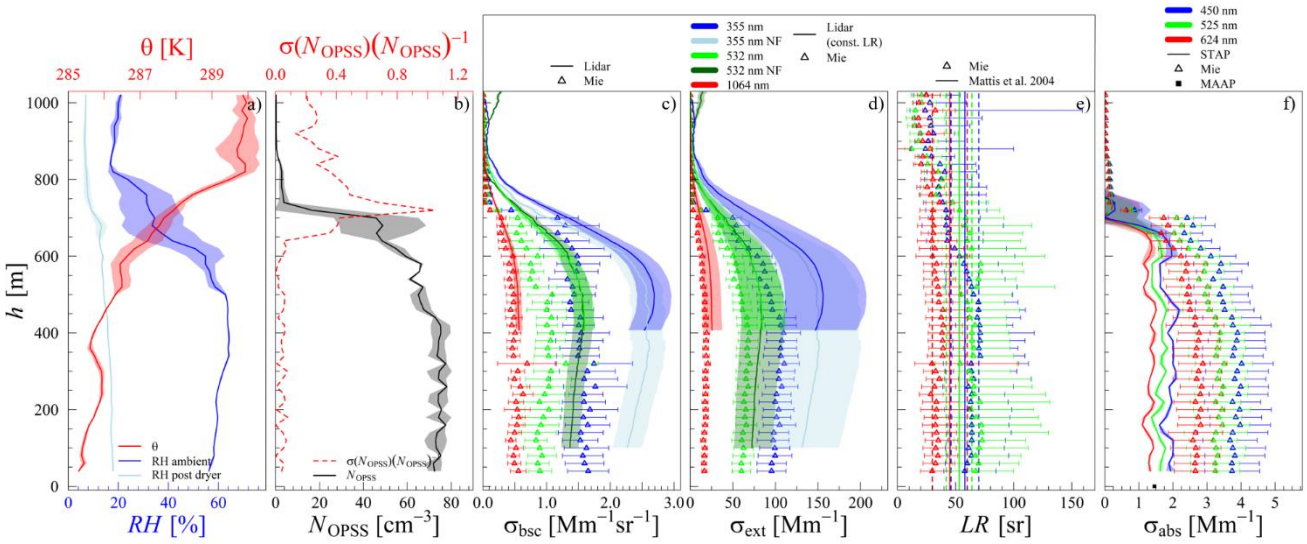


Figure 1: New Figure 8.

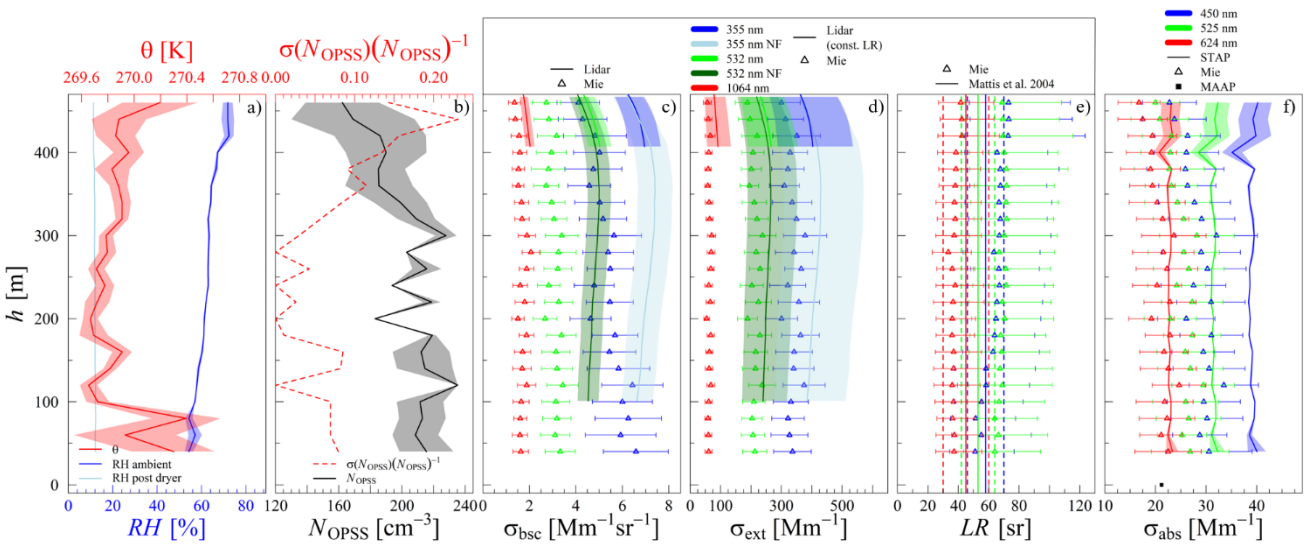


Figure 2:New Figure 7.

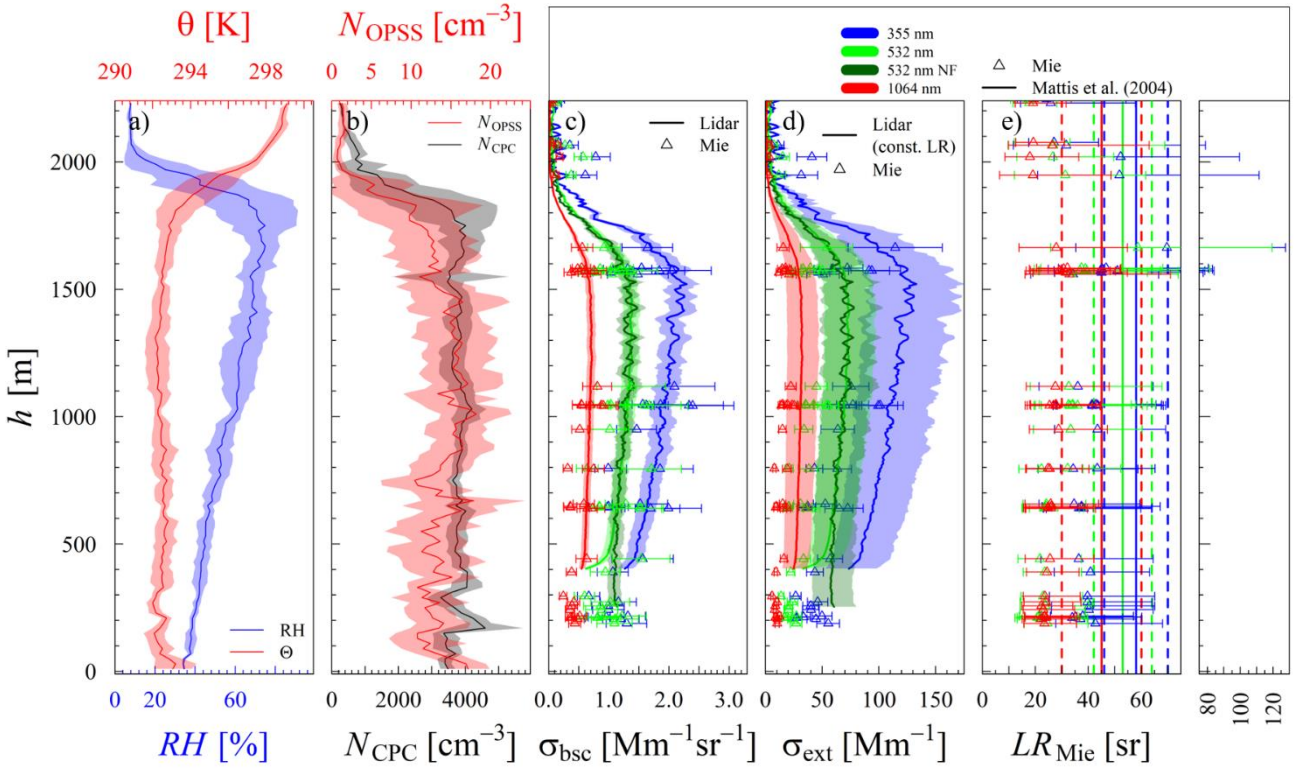
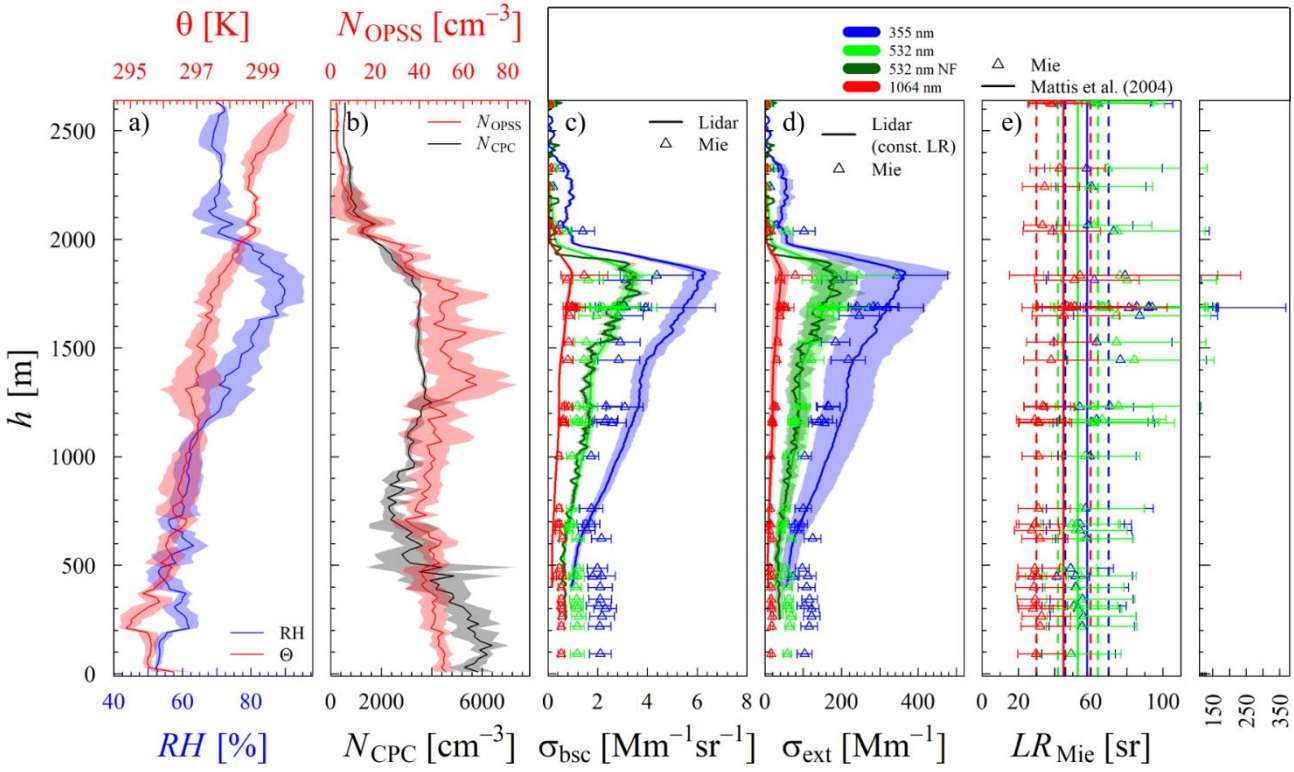


Figure 3: New Figure 5.



95 **Figure 4:New Figure 3.**

8) Line 627- 630: How long did the flights 20150617b, 0626a, 0628a, 0628b last? Which heights and locations were chosen for the LR calculation? If there the particle size distributions were influenced by air mass transported from other regions, how can you guarantee the changes in calculated LR were merely resulting from the relative
100 humidity?

Typical flight duration and altitude coverage are stated in the Supplementary Material. All flights were planned to last around 2 hours and were conducted up to altitudes of 2700 m. In detail, flight 20150617b lasted from 12:43 to 14:19 UTC, 20150626a from 08:08 to 09:58 UTC, 20150628a from 09:07 to 11:04 UTC, and 20150628b form 13:05 to 14:48 UTC. The exact flight time for flights
105 *20150626a and 20150617b is stated in lines 471 and 518.*

All data points from the four selected flights are part of the LR calculation. For this, each available aerosol particle number size distribution is taken within and above the PBL. In particular, the altitude range of flight 20150617b is 0 to 2280 m, 20150626a, 0 to 2660 m, 20150628a 0 to 2380 m and 20150628b 0 to 2670 m. Hence, we agree that we cannot guarantee that the increase in relative
110 *humidity solely causes the increase of the aerosol particle light extinction coefficient when we talk about the vertical profile of the LR.*

However, Figure 5 displays the profile of the LR in the ambient and humidified state and the RH profile of the given flights. The profile of the RH influences the LR more than the aerosol population indicated by relatively constant LR with height at the dry state. Nevertheless, the vertical aerosol distribution is not necessary for the investigation of the LR-RH-behavior because the LR-RH-dependence was calculated for each particle number size distribution itself. At most, the type of aerosol can influence the fitting exponent γ(λ) of the LR-RH-fit and is already discussed in the manuscript (lines 655 – 658, manuscript after first round of revision).

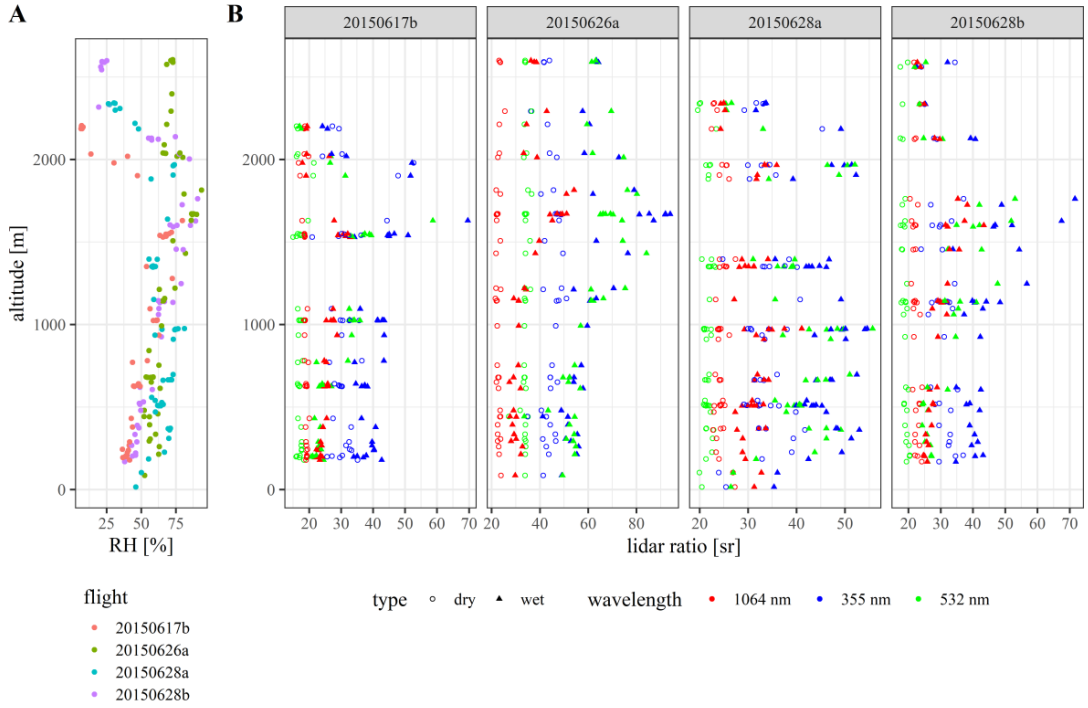


Figure 5: Profile of *RH* (A) and lidar ratio modeled in dry (circles) and ambient state (triangles) for the four investigated flights of the study.

However, to clarify the direct relationship between the LR and RH and to examine the influence of the dry state aerosol, we added (line 635 - 637): "Aerosol changes with height probably cause some changes in the LR too. However, a comparison of the LR profile in the dry state with the LR profile in the ambient state shows that the LR increases more with increasing RH than it does with a change in the aerosol itself (see Figure S7).", and referred therein to a new section of the supplementary material:

LR at dry state vs. ambient state

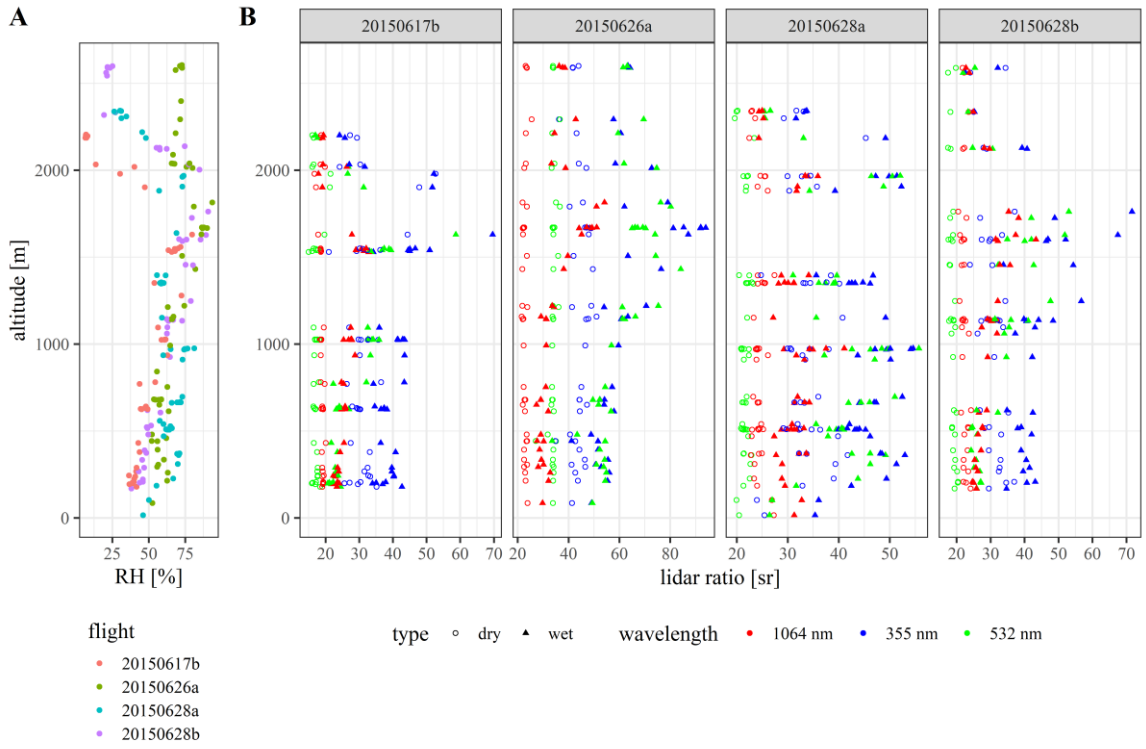


Figure S7: Profile of RH (panel A) and lidar ratio (panel B) modeled in dry (circles) and ambient state (triangles) for the four investigated flights of the study.

Figure S6 displays the profile of RH (panel A) measured during the four flights of interest of the summer campaign. Panel B of Figure S7 shows the profiles of the LR at 355, 532, and 1064 nm calculated with the Mie model in ambient (triangles) and dry (circles) states. While the LR changes only slightly, if at all, in the dry state, the change is more pronounced in the ambient state. Hence a direct relationship between LR and RH is visible.

9) Line 929 - 942: Is the conclusion here necessary?

Yes, we think it is a neat way to summarize the conclusion. Some articles have been presented with that kind of summary, and since the first reviewer does not criticize, we keep it as it is. However, we shortened the conclusion and combined point a and b: It states now:

" In conclusion:

a) Conducting comparison studies of aerosol optical properties, e.g., to validate lidar-based $\sigma_{abs}(\lambda)$, requires a precise determination of the aerosol mixing state, its composition, the inclusion of BrC, and the application of a wavelength-dependent complex refractive index. Information on size- and height-resolved aerosol composition is needed.

~~*b) Airborne in-situ measurements of, e.g., the aerosol chemical composition, including the BrC content, would improve studies focusing on the validation of lidar-based $\sigma_{abs}(\lambda)$.*~~

~~*c)b) A wide range of aerosol particle sizes is covered in this study. However, the modeled σbsc(λ) were on average lower than the measured one. A much further extension of the observed*~~*Observing*

aerosol particles above a size ~~range beyond~~ of 10 µm would ensure that ~~this parameter~~these non-observed particles would not cause a significant ~~underestimation~~ bias based on De Leeuw and Lamberts (1987).

~~d)~~c) Knowing the connection between RH and the LR(λ), the LR(λ) enhancement can be a valuable tool to estimate the LR(λ) at ambient state when the dry state LR(λ) is known. Also, it allows calculating back the LR(λ) in the dry state, when the LR(λ) is directly measured in the ambient state, and an RH profile is known, e.g., via radio soundings.

~~e)~~d) However, long-term measurements must be conducted to verify the LR(λ) enhancement estimates for various aerosol-types and different seasons."

Response to Referee #2:

Blue italic font means the authors' answer, standard, black *text symbolizes the reviewer's comments, and red italic font means manuscript changes, while black italic font symbolizes the original content of the manuscript.*

The manuscript was improved and I thank the authors for the effort done. However, in my opinion, before publication, a further improvement is required concerning 3 aspects:

Response:

We gratefully thank the referee for the spent effort and comments. We respond to the individual points below.

1) Concerning my question on aerosol drying and the hysteresis cycle, the given answer (based on Zieger et al. 2013) is not convincing. Other measurements in Europe, in high organic mass fraction conditions (not marine aerosol) showed aerosol hysteresis (Fierz-Schmidhauser et al., 2010; Ferrero et al., 2019). However, I understand that the used apparatus cannot solve the question as at lines 639-640 is reported: "The RH measured after the dryer was at most 48.3% on flight 20150617b and reached a maximum of 35.8% on the other days". Thus my suggestion is to add a comment concerning the uncertainty expected in case in which the aerosol was not completely dried by the used apparatus.

References:

R. Fierz-Schmidhauser, P. Zieger, G. Wehrle, A. Jefferson, J. A. Ogren, U. Baltensperger, and E. Weingartner. Measurement of relative humidity dependent light scattering of aerosols. Atmos. Meas. Tech., 3, 39–50, 2010 www.atmos-meas-tech.net/3/39/2010/

Ferrero, L., Riccio, A., Ferrini, B.S., D'Angelo, L., Rovelli, G., Casati, M., Angelini, F., Barnaba, F., Gobbi, G.P., Cataldi, M., Bolzacchini, E., 2019. Satellite AOD conversion into ground PM10, PM2.5 and PM1 over the Po valley (Milan, Italy) exploiting information on aerosol vertical profiles, chemistry, hygroscopicity and meteorology. Atmospheric Pollution Research 10, 1895–1912. doi:10.1016/j.apr.2019.08.003.

Figure 6 displays the RH post the deployed dryer. During 20150617b the RH was exceeding the threshold RH of 40% just occasionally. After that measurement flight, the dryer was replaced by a new one, ensuring the drying capability. However, backward-simulations of the hygroscopic growth show that the difference in diameter at 48% is at most 3.2% compared to 35% RH (kappa =0.3, temperature = 25°C), which is the maximum RH reached during flight 20150626a. Although the scattering, backscattering, and absorption efficiency change with change in particle diameter, the resulting optical coefficients are proportional to the cross-section of the particles, which is proportional to the square of the diameter. Hence in the "above 40% RH" case, the deviation is in this regard 6.5%. Compared to 40% RH, the "dry-state" threshold, the deviation is 2.1% or 4.2%, respectively.

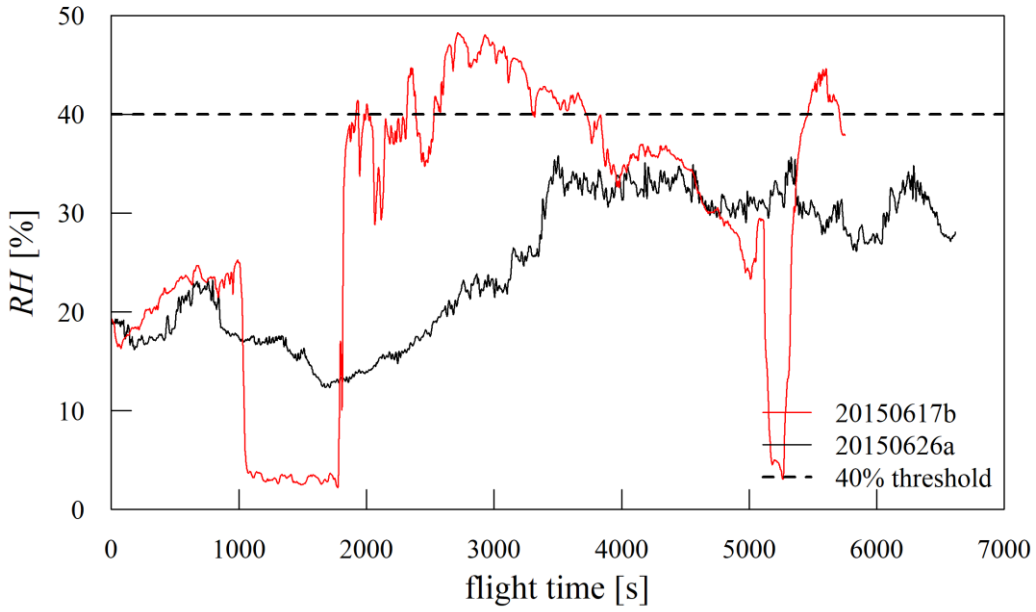


Figure 6: Post dryer RH of measurement flight 20150617b (red) and 20150626a (black).

We added the comment:" In the 48%-RH case, the difference in RH results in a deviation of 3.2% in dry state diameter. The optical coefficients from the Mie calculation are proportional to the cross-section of the aerosol particle. Hence, the dry-diameter deviation translates into a deviation of 6.5% in this regard.".

2) Concerning my question on OPSS PNSD corrections two points in the given answer needs to be clarified:

Point A) at lines 428-429 of the revised manuscript it is reported that "this OPSS PNSD is corrected with in-house software in terms of the complex aerosol refractive index". The development of an in-house software for OPSS correction in terms of the complex aerosol refractive index is yet a matter of a research paper. Here it is used with the aim to correct the size distribution to perform then a closure experiment. Thus there is the need to carefully describe the software and the related performance together with the inner geometry of the OPSS used for the correction. I strongly reccomend to add this part to the supplemental material.

We agree that the correction of OPSS PNSD is still challenging since, besides the refractive index; also the aerosol particles' shape and mixing state are essential. As a best guess, the underlying software calculates with Mie theory, based on the geometry and features of the device, the intensity of the sideward scattered light for particles of a given refractive index. In detail: The software adjusts the particle diameter that the intensity of its sideward scattered light matches the intensity of the calibrated PSL particles. Hence for each size bin, the bin borders are shifted and adjusted to the new refractive index. For the TSI OPSS (mod. 3330), the opening angle of the aperture of 60° (±60°) is used. The wavelength of the device is 660 nm, and the refractive index of the PSL particles is set to 1.581+i0.

Calculations are conducted under the assumption of unpolarized light. The software was used already in Alas et al. (2019).

We added (line 438 - 444): "Briefly, the used software utilizes Mie theory to calculate the intensity of sideward scattered light with a given wavelength of aerosol particles with a complex refractive index and a given diameter D within an angular range. The next step shifts the diameter up to the intensity that matches the intensity of the calibration aerosol (here PSL) of a specific diameter and refractive index. As a result, the size bins are remapped to a new diameter array. For the calculations, the specific characteristics of the device have to be known. In this case, the sideward angular range is ±60°, the wavelength is 660 nm assuming unpolarized light and a refractive index of the calibration aerosol at this wavelength of 1.581+i0."

Point B) as asked, "the altitude correction factor (fh) in eq. 6 does not correct the OPSS optical equivalent aerosol size-bin (i.e. the size of particles) which is, instead, the right parameter needed for proper Mie calculations". The answer does not shed a light on fh, probably it was a my fault to present a vague question. Eq 6 reports fh=N_OPSS(h)/N_OPSS(<x m), where N_OPSS is the number concentration. Thus it is necessary to clarify if the authors are using the total number concentration, the one from each size bine, or a normalized one just to correct the size distribution shape with altitude

In the summer campaign, the PNSD measured by the OPSS was not corrected because of the presented reasons. In the winter campaign, we used in-house software to do so. We lacked information on non-observed size ranges (summer from 800 nm on; winter, up to 300 nm). Hence, these missing parts were corrected by the corresponding PNSD measured at ground level normalizing its shape (each size bin with the same factor) with the ratio of the aerosol particle number concentration detected close to the ground and the one detected at higher altitudes.

The respective part (line 462 - 464) states now: This factor normalizes the ground-based PSND (each bin equally) with the number concentration ratio of the aerosol particles detected by the OPSS at altitude h ($N_{OPSS}(h)$) and the mean in a layer near ground below an altitude x ($N_{OPSS}(<x\ m)$). The altitude-correction factor $f_h(h)$ is calculated according to Eq. (8):

$$f_h(h) = \frac{N_{OPSS}(h)}{N_{OPSS}(<x\ m)}. \quad (8)$$

.

3) The revised version of the manuscript presents at line 151 equation 8 as the first equation. It is necessary to careffully check all the paper, the equation numbers and the related reference along the text

We updated as requested. Changes are highlighted in the marked-up version.