Dear reviewers and editor:

We would like to thank you for taking the time to review our manuscript and suggest such valuable information in order to improve its quality. Even the most trivial suggestion helped in the progress of the text's value.

Some main and relevant changes have been made to the manuscript, as summarized below:

- While submitting the first version of our manuscript, a relevant study related to Aeolus optical products were made. Thus, for the sake of completeness, the following published paper has been included now in the reviewed version of our work:

  Baars, H., Radenz, M., Floutsi, A. A., Engelmann, R., Althausen, D., Heese, B., Ansmann, A., Flament, T., Dabas, A., Trapon, D., Reitebuch, O., Bley, S., & Wandinger, U. (2021). Californian Wildfire Smoke Over Europe: A First Example of the Aerosol Observing Capabilities of Aeolus Compared to Ground-Based Lidar. Geophysical Research Letters, 48(8), e2020GL092194. https://doi.org/10.1029/2020GL092194

- Furthermore, a reprocessing of Granada's ground-based dataset was made right after the manuscript submission, to increase the quality of the depolarization-related products. No relevant differences were found between the results obtained with the former and the reprocessed products. However, the charts included in the manuscript have been updated accordingly. These changes affect Figures 5, 12, 13, 14 and 15 (a new figure is included following the Reviewer#2 suggestion, so these figures correspond now to Figure 5, 13, 14, 15, and 16, respectively).

Apart from these main changes, a set of minor changes and typos has been corrected throughout the document. Following this introduction, you may find a detailed answer to every Reviewer#1's comment. All of the listed changes and the ones suggested by the reviewers can be seen in the new version of the manuscript, marked in red.

Reviewer 1

1) *In the section 3.1-Database and intercomparison methodology (page 7) the authors describe the measurement protocol for each station - For Granada, a 1.5 hour interval for the morning overpass time and a 1-hour interval for the evening overpass time (i.e.17:30 - 18:30 UTC) were chosen. For Barcelona, a 1-hour range centered at the overpass time was considered. For Évora, a 1.5-hour interval containing the overpass time was considered to take into account the larger distance between the Aeolus ground track and the lidar site. I recommend the authors discuss in more details the uncertainties or issues that 1 or 1.5-hour time differences can cause in the comparative analyzes between the ground-based lidar and the Aeolus satellite.*

Each station is managed by independent research groups, which means that some minor differences can be found in the processing of the products, although all of them are processed according to the EARLINET protocols, processing chain and standards. In particular, each group produces its products under different temporal periods (i.e. 30 minutes or 1 hour). This fact, plus the need of an adequately long averaged period of measurements, made us use different temporal averages for each station. Regarding the temporal evolution of the meteorological conditions, it has been checked that during the 1 hour or 1.5 hour interval the atmospheric conditions had not changed significantly, so that the temporal averaging of the signals is coherent in the interval. Thus, we get to choose the most convenient temporal average for each station as no uncertainties nor artifacts are introduced from the fact that two sets of averages have been chosen. Thus, the next clarification has been added in the new version of the manuscript:

The temporal evolution of the meteorological conditions and layers has been checked individually, so as to ensure that the selection of 1 hour or 1.5 hour as average interval for the lidar signals does not introduce any artifact in the comparison.

2) *In section 3.2-Aeolus-like conversion of ground-based lidar particle backscatter coefficients - The authors proposed a method to estimate the linear particle depolarization ratio at 355 nm* ($\delta^{part}_{linear,355}$) *from the linear particle depolarization ratio at 532 nm. However, it is stated the third lidar station in Barcelona, does measure both depolarization ratios but for the sake of consistency of the data processing, calculated the same way than the other two stations. In figure 2c (page 39 and discussion on page 10) the autors presented a scatter plot of dust and non-dust aerosol particles obtained from dual-polarization measurements in Barcelona, together with spectral conversion factor Kδ=0,76±0,01 and the spectral conversion factor from literature results for dust and non-dust types, equals to 0.82 ± 0.02. It is not very clear why the authors prefer to use the literature values instead of spectral conversion factor retrieved from measurements.*

On one hand, we set up a state-of-the-art database of dual-polarization measurements covering all the aerosol types reported in the literature up to the date, offering a complete insight worldwide.

On the other hand, Barcelona station provides independent measurements that undergo a set of validation criteria but have not been published in any peer-reviewed journal yet. Moreover, the Barcelona depolarization ratios measurements at 355 nm do not cover the complete period analyzed in our study, and are mainly biased to dust cases.

Considering all these arguments, we prefered to use the spectral conversion factor derived from the literature in order to base our results on measurements of different stations (i.e. different environmental conditions and aerosol types), which themselves have been reviewed through peer review process. Thus, we included the following explanation in the mentioned paragraph:

The literature-derived factor is implemented in order to base the results in previous measurements reported for different environmental conditions and aerosol types, which are subsequently used as reference in other studies.

3) *In section 3.3 Statistical parameters, the authors stated "The resolution of these bins depends on the altitude range: 500 m between 0 and 2 km asl, 1 km between 2 and 16 km asl and 2 km between 16 and 30 km asl. Because the ground-based lidars present a much finer resolution, of the order of a few meters, the resolution of each ground -based profile has been degraded*

to the Aeolus vertical resolution." The authors' choice to downgrade the data quality of the groundbased handles to perform a bin-to-bin comparison is understandable, however, the decrease in signal quality doesn't seem to make much sense when comparing the Aeolus and the groundbased lidar, especially when taking into account the different nuances of the atmosphere in the region closest to the surface. How do the authors understand that this loss of quality, or the lack of a finer resolution in the Aeolus data, can affect the application of the data to the study of optical properties of aerosols?

This question is particularly interesting because of several facts. On the one hand, the vertical resolution of Aeolus profiles for aerosol products is really coarse compared to the vertical resolution of other satellites, like CALIPSO. Thus, the readers have to bear in mind that Aeolus vertical resolution does not provide a detailed characterization of the atmospheric optical properties. Aeolus products do provide valuable information for the detection of significant layers and clouds, as it can clearly be seen in the case studies provided (Section 4.2). On the other hand, the atmospheric layer with more relevant nuances happens to be in the lowermost troposphere, where Aeolus is proven to fail (at the very bottom). Therefore, once again, we can say that Aeolus provides valuable information for the characterization of significant layers and clouds. Thus, the following explanation is now included in the conclusions section:

However, as it can be noted from the results, Aeolus vertical resolution is too coarse (especially compared to other satellites) for a detailed characterization of the nuances of the atmospheric optical properties. Thus, Aeolus provides valuable information in the detection and characterization of significant aerosol and cloud layers.

4) *In section 4.3 Case studies, the authors stated "Sun -photometer measurements are taken into account for the sake of completeness aerosol typing, through the study of the aerosol optical depth at 675 nm (AOD 675)". Why was the 657 nm wavelength chosen? Why not choose the AOD values in the UV region as 340 or 380 nm, instead?*

The 675 nm wavelength was chosen (among all of the Sunphotometer wavelengths) as it is one of the wavelengths most used as reference by the scientific community. Furthermore, Barcelona's Sunphotometer only worked with 440, 500, 675, 870 and 1020 nm on the 2nd July 2019, so the 340 or 380 nm channels could not be used. Thus, for the sake of homogeneity, the 675 nm channel was chosen as a common reference for the three stations.

5) *"The location of the stations is highly interesting due to their proximity to the Sahara Desert and mainland Europe, so frequent events of mineral dust and anthropogenic particles could be detected by the satellite. In addition, Barcelona lies just in the coastline, and both Barcelona and Granada present high concentrations of anthropogenic aerosol, while Évora aerosol concentrations could be classified as rural. Thus, Aeolus operation can be tested under a complete set of atmospheric scenarios." How was the difficulty of comparing the layers closest to the surface taken into account that Barcelona station is located just in the coastline and is influenced by the mixture of anthropogenic aerosol and/or dust and marine aerosol? How might this difficulty in comparing the layers closest to the surface have influenced the statistical results?*

With the quoted paragraph we attempted to express that we tried to assess Aeolus performance under different situations although the focus was not on the differences this set of scenarios might produce. Indeed, the objective of this study is to test Aeolus performance, and luckily we are allowed to work with different scenarios and not just one (i.e. only coastline/marine or flat/rural settings). Thus, the aerosol mixture state does not statistically affect the results. As presented in Section 4.3, the geographical differences of the stations may affect the statistical results, with notorious differences in the lowermost regions, affected by the surface. Consequently, the statistical analysis presented in Section 4.3 was performed independently for each station so as to detect these differences. Furthermore, as previously mentioned in the answer#3, Aeolus vertical resolution is too coarse (when compared to other satellites as CALIPSO) limiting the detailed characterization of the nuances of the atmospheric optical properties at the lowermost atmosphere.

Regarding Barcelona station location just in the coastline, as the satellite overpasses the station at a close distance we can assume that both instruments (ground-based and space-borne lidars) detect the same air masses. Consequently, both instruments will register the same effects that the geographical layout might produce, so no special considerations have to be taken into account in the statistical analysis.

6) *In Page 13-lines 385-386 - "The HYSPLIT model indicates that the 12:00 UTC air masses over Évora at 1.7 and 2.7 km agl (equivalent to 2 and 3 km asl) are coming directly from lower altitudes in Northern Africa (Figure 7a)." This sentence is slightly confused, please, rewrite the sentence.*

The following sentence will be included instead of the quoted one:

The HYSPLIT model indicates that at 12:00 UTC the air mass located over Évora at 1.7 and 2.7 km agl (equivalent to 2 and 3 km asl) originates from surface-level of Northern Africa (Figure 7a).

7) *Page 15 - Lines 455 and 456 - "First, the satellite presents a satisfactory agreement with the ground-based lidar in the whole available profile under both SCA and SCAmb (Figure 11c)." For the first atmospheric layers up to 2,5 km asl it seems there is an underestimation of Aeolus particle backscatter signal, and for the layers from 2.5 to 6 km asl, it seems to have an overestimation of Aeolus particle backscatter signal. Considering that Barcelona is the station with the most complex scenario, with several layers coming from different sources and containing different optical properties, the comparison analysis seems to be much more sensitive. I believe the authors could explore more this aspects in order to improve the manuscript discussion.*

The development of a more detailed comparison analysis for this particular and interesting case study could be really rewarding. Unfortunately, regarding Aeolus limitations it is not possible to increase the detail of the comparison, as we have to work with Aeolus fixed vertical resolution.

Furthermore, as stated in Section 4.1, a single conversion factor K_δ is considered in the intercomparison, in order to minimize the uncertainties and the effects that different aerosol types might cause. However, hereunder we explore the dependency of the Aeolus-like profile depending on the K_δ, this is, depending on the aerosol types considered in Section 4.1. Thus, a set of K_δ values have been taken: 0.82, for bibliographic (considering all aerosol types); 0.76 for the whole set of Barcelona cases; 0.72 for the set of dust cases in Barcelona; and 0.90 for the set of non-dust cases in Barcelona. The following plot is a zoom of the results in a way that the differences are somehow visible. We can see slight differences between the profiles for different K_δ values. In fact, the largest difference can be seen between the K_δ= 0.90 and the K_δ= 0.72 profiles, and they differ only in 3%.

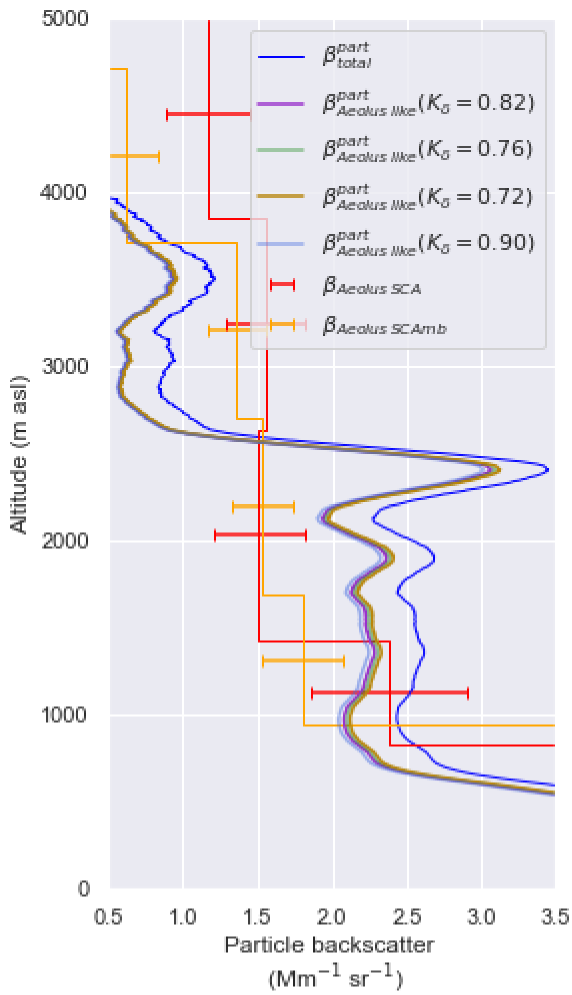

Figure. Aeolus SCA and SCAmb co-polar particle backscatter coefficients (without quality flags) and the corresponding ground-based Aeolus-like backscatter coefficient (considering a set of different K_δ values) for the case study in Barcelona on the 2nd July 2019.

Thus, a more detailed study of the sources and optical properties of the aerosol layer will not improve the intercomparison assessed in the study.

8) *Page 15 - lines 467-469 - "101 B10-overpasses for Granada, 51 for Évora and 52 for Barcelona, and after applying the set of requirements, the intercomparison has been performed with 24 cases for Granada, 15 cases for Évora and 16 cases for Barcelona, leading to enough statistical significance." What criteria were considered by the authors to reach the conclusion that this number of cases is statistically significant?*

We wanted to express that the set of different scenarios considered with the mentioned dataset of cases allows us to test Aeolus performance under different circumstances. No particular criteria is considered to reach the conclusion that the number of cases is statistically significant. Now, the quoted sentence has been restated as:

101 B10-overpasses for Granada, 51 for Évora and 52 for Barcelona were considered, and after applying the set of requirements, the intercomparison has been performed with 21 cases for Granada, 15 cases for Évora and 16 cases for Barcelona, leading to a wide dataset of cases.

9) *Page 15 - lines 474-475 - "With the implementation of the quality flags (Figure 12c and 12 d), all of the sets range from 0 Mm-1sr -1 .". Please, consider correct this sentence. All the sets range from 0 to which value?*

The quoted sentence has been restated and completed as:

With the implementation of the quality flags (Figure 13c and 13d), all of the sets range from 0 Mm^{-1} sr^{-1} onwards. Actually, the maximum values mentioned are still flagged as valid, 86 Mm^{-1} sr^{-1} and 68 Mm^{-1} sr^{-1} in the case of the SCA and SCAmb, respectively.

10) *Page 16 - lines 490-491 - "Aeolus backscatter coefficient uncertainties (known as Aeolus error estimates) are addressed through the biases between satellite and ground-based measurements. Figure 14 reveals that the larger the Aeolus uncertainties, the larger the bias." Just for improve the understanding, is the bias mentioned in this sentence and presented in axis Y in figure 14 the same values calculated in equation presented in page 9-line 286? If yes, I would recommend the authors rewrite the sentence.*

The bias mentioned in lines 490-491 and presented in Figure 14 is indeed the parameter mentioned in line 286. In fact, it is the only bias mentioned in the manuscript. The quoted sentence has been restated and completed as:

Aeolus backscatter coefficient uncertainties (known as Aeolus error estimates) are addressed through the biases between satellite and ground-based measurements (as presented in Section 3.3).

11) *Page 23 - line 710 - Please, consider correct the reference "Córdoba-Jabonero, C., Sicard, M., López-Cayuela, M.-A., Ansmann, A., Comerón, A., Zorzano, M.-P., Rodríguez-Gómez, A., and Muñoz-Porcar, C.: Aerosol radiative effect during the summer 2019 heatwave produced partly by an intercontinental Saharan dust outbreak. 1. Shortwave dust-induced direct impact, Atmospheric Chemistry and Physics, 21, 1–25, https://doi.org/10.5194/acp-2020-1013, 2021." since the DOI is leading to the pre-printed version of the article.*

The quoted citation has been corrected as:

Córdoba-Jabonero, C., Sicard, M., López-Cayuela, M.-A., Ansmann, A., Comerón, A., Zorzano, M.-P., Rodríguez-Gómez, A., and Muñoz-Porcar, C.: Aerosol radiative effect during the summer 2019 heatwave produced partly by an intercontinental Saharan dust outbreak - Part 1: Shortwave dust-induced direct impact, Atmospheric Chemistry and Physics, 21, 1–25, https://doi.org/10.5194/acp-21-6455-2021, 2021