

Atmos. Meas. Tech. Discuss., referee comment RC2
<https://doi.org/10.5194/amt-2022-205-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2022-205

Anonymous Referee #2

Referee comment on "First assessment of Aeolus Standard Correct Algorithm particle backscatter coefficient retrievals in the eastern Mediterranean" by Antonis Gkikas et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2022-205-RC2>, 2022

The authors perform an assessment of the SCA backscatter coefficient product from the Aeolus satellite by comparison to ground based lidar observations. They have split their work in two parts. First, an analysis of four dedicated, illustrative test cases is presented, including a creditable multitude of ancillary data that provides information on the aerosol origin and type. In a second part, all available collocation cases over the chosen lidar stations contribute to a statistical analysis of bias and RMSE, spanning the current mission lifetime. In lack of a cloud mask within Aeolus' data products, the authors efficiently filter the data themselves and can thereby show moderate to good performance of Aeolus backscatter coefficients. However, the findings suggest that particularly the retrieved backscatter coefficients closest to the ground are not reliable since they suffer from low SNR. The retrieved backscatter coefficient above the ground is biased due to surface reflectance.

However, there are some substantial changes and clarifications necessary before publication of the work.

General (Major) Comments

- I agree with referee 1 that the wording and sentence structure throughout the manuscript makes it often more difficult to grasp. That is particularly because of

numerous insertions into the sentences, separated by commas or parentheses, and maybe a general trend for nouns over verbs. To provide only one example from L.468 "Under the prevalence of the Etesian winds (Tyrlis and Lelieveld, 2013), a typical pattern dominating over the broader Greek area during summer months, when winds blow mainly from NNE directions, anthropogenic aerosols from megacities (Kanakidou et al., 2011) and particles originating from biomass burning in the eastern Europe and in the surrounding area of the Black Sea (van der Werf et al., 2017) are transported southwards."

The main clause "Under the prevalence of winds [...] aerosols [...] are transported southwards." is stretched out too much.

- Some parts of the manuscript seem not to contribute to or distract from the scope of the paper. Some sections or paragraphs could potentially be shortened or omitted, by asking who the audience of this work is. E.g. the second section with the Aeolus instrument description contains very general information that is mostly not used throughout the rest of the manuscript and can therefore be referenced (see suggestions in specific comments). Also, the conclusion can be made more compact by separating it into a conclusion and an outlook section, or can be condensed in other ways (also see suggestions in specific comments).

- It is good that the authors assess the aerosol climatology via the MODIS-Aqua AODs. However, the performed analysis of concentric circles seems not well suited for the assessment of the horizontal heterogeneity, see specific comments regarding L.351-366

- throughout most of the text, the authors do not differentiate between the performance of the Aeolus satellite itself and the performance of the retrieved SCA co-polar backscatter coefficient within the L2A product. This needs to be clarified, particularly since two significantly improved optical properties products are available as of March this year (see specific comments)

- the currently implemented collocation method appears to me to have an offset of about 45 km in flight direction, since only the start of a BRC but not its center location is used for the distance calculation to the lidar ground stations. If I did not miss something, this will need to be adjusted, making necessary to reanalyse the data and update the corresponding plots

- Subsections 6.1.3 and 6.1.4: In my opinion, the descriptions and conclusions of the individual Aeolus lidar profiles in Fig. 3 may be much too detailed and flawed. I explain in my specific comment on L.502-503, that there is reason to believe that the discussed discrepancies are just noise induced and therefore the reached conclusions are not valuable or generalizable. I recommend the following procedure: As a first validation step, I encourage the authors to provide Figure 3 with all negative SCA backscatter values shown. This will provide an impression of the actual noise level encountered in the SCA backscatter in the different test cases. I expect to see values up to minus 0.5-1 Mm⁻¹sr⁻¹ in some cases in accordance with e.g. Fig. 8 in Ehlers et al. (2022, doi.org/10.5194/amt-15-185-2022). If that is indeed the case, then the discrepancies along the profiles may be mostly noise induced and the current, detailed conclusions must be reconsidered, i.e. the authors should test for the hypothesis and make accordingly changes to the text. In this case, especially the statement in the abstract L.41-43 "The level of agreement between spaceborne and ground-based retrievals varies with altitude when aerosol layers, composed of particles of different origin, are stratified (8th July 2020, 5th August 2020)." is contestable.

Specific comments

L.38-43 This could be more compact, considering it is in the abstract. Particularly the discussion of the 4 test cases seems very specific and could be condensed into a shorter sentence.

L.41-43 The "level of agreement" is not strictly defined and, hence, seems subjective. In my opinion, this statement is presented too confident. In fact, this is concluded from two single BRC of Aeolus. The authors themselves stress the issues with collocation, so I am not convinced at this point that the remaining variations in the profiles are caused only by the stratification, but can originate from horizontal inhomogeneity of the atmosphere's aerosol load. The supplementary material helps only little, since the models provide AOD only.

L.74-80 This sentence is too long and the last part seems not to fit in gramatically.

L.82 "as well as the geometric features of the particle's layers" What are the "geometric features", if not the already-mentioned vertical structure? Please omit or specify.

L.94-115 This paragraph provides an overview of the L2B wind product development and application. Considering the paper's scope of aerosol backscatter assessment, I recommend to omit/condense it.

L.116-128 This paragraph might be better placed in / merged with the second section about the ALADIN instrument.

L.169, Section 2; While reading this section I was reminded of other Aeolus related works. In fact, I was wondering whether the degree of detail is relevant for the audience of your manuscript, or if you could get away with a more high-level description of the Aeolus typical vocabulary only (as in the L2a user guide). Essentially, the information here can be looked up in the Aeolus Science Report or many other papers. But since your work focuses on validating the data rather than e.g. modifying the L2A processing chain or including so-

far unknown instrumental effects, it may be a consideration to omit most parts for brevity.

L.192-194 It is unclear with which property the angle increases. Please make the formulation unambiguous by changing the statement to something like e.g. "The 35 degree off-nadir pointing corresponds to an angle of about 37.6 degree with the Earth surface, due to its curvature".

L.205, Section 3; This description of the L2A data product is outdated at least with the start of the new baseline 2A14 from 29th March 2022. It has been decided to remove the ICA product completely and two new optical property products have been added, namely the SCA-MLE Optical Properties (Ehlers et al., 2021) and the AEL-PRO Optical Properties (from adjusted EARTHCARE algorithms). Both products are expected to bring considerable improvement over the SCA, since the inverse retrieval problem is solved not algebraically but via state-of-the-art methods (Maximum Likelihood Estimation, Optimal Estimation, respectively), see Ehlers et al. (2021) for the SCA-MLE product. These changes are tracked e.g. in the Aeolus Level 2a Processor Input/Output Data Definition available here: <https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-L2A-Input-Output-Data-Definitions-Interface-Control-Document> Please give an adequate description of the data product, in order to put your analysis in the correct context. To my knowledge the Aeolus mission data has not yet been reprocessed with the new processors, which then offers potential for future studies.

L.238 NITWT is not the name of the method, but the name of the variable that allows for simpler notation.

L.243-250 This paragraph is a perfect introduction to then mention the SCA-MLE and AEL-PRO optical properties data products, which aim to mitigate such problems to a big part. A brief description could be added hereafter to update the section. It must be stressed that

also the backscatter profits from the processing update!

L.250-252 This is not a primary reference for the zero-flooring. The primary reference is Flament et al. (2021) or the L2A Algorithm Theoretical Baseline Document, section 6.2.2.1, see here: <https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-L2A-Algorithm-Theoretical-Baseline-Document>

L. 351-366, Criterion for spatial homogeneity; The authors want a measure for spatial homogeneity of the atmosphere's aerosol load on an instantaneous base. However, the presented, concentric, climatological analysis is not suited for these needs for at least two reasons. Also, the description lacks some detail. The two main points below:

i) Figure 1 provides one AOD value per concentric circle and location. So the reader has to assume that in addition to the spatial average a temporal average over the 10 year period has been performed. This average is not mentioned in the text and the word "climatological" appears only in the figure caption. Therefore, the word "Annual" in Fig. 1i and 1ii is misleading.

Now, averaging the AOD pattern over time will potentially smoothen out most of the horizontal heterogeneity of AOD that is present on a daily basis. However, the latter is the desired property in order to assess the quality of collocation. An (oversimplified) counter example goes as follows: Assume AOD pixels follow a chess board pattern (with changing locations over time due to wind). This would show a lot of heterogeneity, hampering collocation. But due to the two averages, one over the rings and one over time, the developed criterion would indicate perfect homogeneity.

ii) Another shortcoming is the possibly very location specific outcome of this analysis: In my opinion, there is no reason to either favour Aeolus' frequent observation location or the ground-based lidar location as a center for the concentric circles. However, if the center was chosen e.g. 80 km away from pollution sources such as Athens or Thessaloniki, then their increased AOD pixels would be averaged with all pixels from a whole ring of unrelated locations, including ocean, meadows and villages 160 km away. This way, pollution sources will be hidden by averaging, if the circles' center is not coincident with them, making the presented analysis little robust.

Standard tools such as 2D autocorrelation functions of the AOD "images" would not suffer from such shortcomings.

L.374 You take the beginning of the scan as the location of an Aeolus BRC, however, its middle is more representative as centerpoint of the measurement but lies about 45 km further away. When considering Figure 2(i), imagine now that by random chance, Aeolus had started scanning each BRC 5 km earlier. Then the BRC that is now red would not be considered in the analysis at all, though still closest.

This means, your collocation criterion is currently offset by about 45 km in flight direction, which is quite a lot! This must be fixed and can be done, e.g. approximately, by applying the running average filter ([0.5 0.5]) over the current latitude and longitude arrays, or by assuming the satellites' speed and direction. Otherwise, the location of the center measurement within the BRC can be extracted from elsewhere in the L2A product, to the best of my knowledge.

Fig.2 The tips of the orange arrows are barely visible, please enhance.

L.380-383 The collocation criteria should be objective, so please quantify by up to how much time they have been relaxed, another hour? 2 hours?

L.395-397 When reading the manuscript, my burning question was, how many of the above mentioned cases/BRCs remained after filtering. I only found this information much later in the text. Could the authors please consider moving this information up here?

L.408-410 Can the authors motivate here why explicitly these cases were chosen? Also, I wonder which criterion was applied to choose a single BRC out of each case, presented in Fig. 3. The spatially closest? The visually most representative? I presume that at least for one of the cases there was more than one BRC to consider.

L.442-444 The word "ideal" is exaggerated.

L. 484-486 As the authors report themselves earlier, the backscatter coefficient in SCA and SCA midbin is essentially identical, just averaged onto two different scales. Hence, I do not support this argument of overestimation/underestimation and find it misleading. Also, with a quick look it seems that both are overestimating. Do the authors mean that the layer reaches too far up in SCA midbin? Please specify.

L.502-503 This sentence is stated with a suggestion, while in fact the information should just be that Aeolus and PollyXT do not agree over the entire profile, and where. As you are well aware this mismatch can have various reasons but a lack of performance of Aeolus.

At this point I want to also mention, that the error bars on the L2A products are not found

to be accurate, and hence suggest a wrong sense of precision, see the recent work of Adrien Lacour from Meteo France and e.g. Fig. 8 in Ehlers et al. (2022): In this test case, the MLE retrieval brings the optical properties much closer to the ground truth than SCA and SCA midbin. So the gaping disagreement between the SCA or SCA midbin and the ground truth is apparently due to noise in the cross-talk corrected particulate (Y) and molecular (X) signals ($\beta_p = Y/X \cdot \beta_m$). The true magnitude of this noise can also be illustrated with the magnitude of the negative backscatter values in almost clear atmosphere, which unfortunately are not shown in Figure 3. However, Figure 6 gives a good idea of the spread of negative values in SCA, indicating a ballpark value of 0.5 up to 1 Mm⁻¹sr⁻¹ around the GROUND $\beta = 0$.

Now, the discussion following in L.502 to 507 focuses entirely on the value of two Aeolus bins between 2 and 4 km altitude, the errors of which are most likely underestimated. However, the discrepancy with the ground truth is not much bigger than the approximate noise amplitude estimated from Figure 6 above. Therefore, it is very much possible that these discrepancies in just these 2 bins are indeed caused by noise! Therefore, it is not reasonable to generalize from these results an altitude dependent performance and to conclude a contradiction between the observations on such a weak fundament. This can only be done statistically.

L.530-535 I see no to little reason to underline that SCA midbin is "better" in this particular case. As you point out, SCA midbin has just worse resolution, which helps to reduce the mismatch here, because averaging consecutive bins also reduces the noise. This is no characteristic of this particular lidar profile but follows from the math: In general, SCA midbin is worse at high SNR because of the resolution loss, but appears better in low SNR due to the additional smoothing that the average implies.

L.576-577 Can you specify how the ground profiles have been rescaled to match vertically the Aeolus bins? Were the ground-based observations averaged onto both different scales or simply sub-sampled? The latter is not preferred. Please provide an explanation or formula.

L.579-585 The range bin index is a tricky reference to perform the analysis on, but I see the need for this implementation. However, it should be stressed in a separate sentence that, this way, one may mix up bins of e.g. 250 m size with bins of 1 km size, which have different noise properties. This is important for interpreting the reliability of RMSE and bias.

L.586-596 Reading the text while looking at the figure, I cannot follow the choice of the authors to discuss the groups of bins 1-3, 4-12 and 12-23 separately. Can the motivation be explained? The bias and RMSE within the group 4-12 is anything but homogeneous.

L.593 "the most important finding is that Aeolus is not capable to reproduce satisfactorily the backscatter profiles" I find this a bold statement to make here. It is not Aeolus but specifically the current Aeolus SCA product in absence of cloud flagging. The cloud-flagged observations, presented some lines thereafter, let you draw a very different conclusion!

L.600-603, point ii); It is not motivated how increased noise causes bias in backscatter coefficients (I assume that "overestimation" is used synonymously to bias, if so, please use "bias" throughout the manuscript whenever appropriate). This is explained in Sec. 4.1 & 4.2 in Ehlers et al. (2022), so maybe reference here as well?

L.615 It should be already explained here, that this low positive bias is due to omitted negative backscatter values (this can be seen in the scatterplots), as you do later in L.662-665. To my knowledge, the corresponding L2A processor parameter has been adjusted so that negative backscatter values in SCA midbin are not just omitted in the newer baselines!

L.622-625 It should be made clear that this statement regards only the bin closest to the surface. Also, "level" should be replaced by "bias".

L.644-646 I would not use the word contradiction. I can simply not be said based on metric 1,2 and 3 which product is "better". However, it should also be mentioned that the SCA midbin scatterplot contains less data due to the inherent flagging of negative values, see scatterplots, and hence the analyses are not strictly comparable.

Also, the discussion whether SCA backscatter or SCA mb backscatter is "better" depends simply on SNR, as has been addressed in my comment on L.530.

L.672 please use "bias" as in Table 1, instead of "overestimation", in order not to confuse the reader whether or not these are two different statistical properties.

Table 1&2; provide units!

L.668-678 This paragraph describes the statistics of the unfiltered data in detail. At this point, it has already been made clear to the reader that the unfiltered data is not suited for statistical analysis. Hence, Table 1 may as well be moved into the Appendix and may be kept for the sake of completeness, including a hint in the text. Instead, a bit more detail about the metrics for the filtered backscatter profiles would be appropriate in L.678-681.

L.722 What does "they" refer to?

L.748 It's rather "SCA backscatter coefficients" to be specific.

L.763 Please specify and write "Aeolus' SCA backscatter product" instead of "Aeolus".

L.781-790 Partly repeats the cross-polar misdetection mentioned above in L.757. Also, this paragraph does not contain a conclusion from your analysis seems detached. It resembles more of an Aeolus-2 future mission outlook? Maybe move into a separate section "outlook", if the information is crucial in your opinion?

L. 793-797 The content of this text should be moved into Section 3, since these products have been released already at the end of March 2022. Specifically, it needs to be clarified that not only the extinction but also the quality of the backscatter coefficients (especially precision) is significantly increased with the Maximum-Likelihood Estimation (MLE), making new Cal-Val studies worthwhile once there processed data is available.

L.804-810 This also reads as a mission outlook rather than as a part of your conclusion and may be dropped or moved into a separate section "outlook".

Technical corrections

L.230 Refer to the "C coefficients" as cross-talk coefficients as above. In general, using words as "so-called" and setting words in quotation marks should be avoided. It suggests little reliability.

L.241 "downwards", same comment as in L.230

L.343 The formulation seems odd. Just write "in Section 5" and omit the part in parentheses.

L.369-370 This sentence is wordy/bulky. Better: "The Aeolus L2A backscatter profiles are compared to the measurements of three PANACEA lidar stations."

L.384 replace "rest" by "remaining"

L.481-483 The information in the parentheses is different from the information in the text (SCA vs. Ground and Ground vs. Aeolus-like Ground observations).

L.586 This should be Fig. 4 instead of 5.

L.653 replace "not any" with "no"

Ref list: Ehlers et al. (2022) is not included though cited in the text?