Comment on acp-2022-412

Anonymous Referee #2

Referee comment on "Validation of the TROPOMI/S5P Aerosol Layer Height using EARLINET lidars" by Konstantinos Michailidis et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-412-RC1, 2022

The authors present a quantitative evaluation of the accuracy of the aerosol layer height (ALH) product derived from the satellite-based Sentinel 5P-TROPOspheric Monitoring Instrument (TROPOMI-SP5) using ground-based lidar observations submitted to the European Aerosol Research Lidar Network (EARLINET) database. The study is focused on the Mediterranean Basin in which observations from 7 EARLINET stations are selected, taking into consideration their proximity to the sea and the presence of absorbing aerosols. Within a 3-year time frame the authors have found 34 suitable cases for the comparison which shows the challenge in satellite validation attempts but also marks the importance of networks such that of the EARLINET. Given the importance of the ALH information in radiative forcing calculations, UV aerosol index, aviation safety etc. the study has scientific interest and therefore it is worth publishing. The work is overall sound, and I have only a few points to raise.

We would like to thank the reviewer for his/her fruitful comments that led to the improvement of the manuscript. In the following, answers to comments are reported just below each related comment. When needed, the part of the manuscript we modified or added to the old version, is reported.

General changes to the manuscript:

- In the revised version, new collocated cases have been identified and added in the analysis. We have added twenty-nine (29) more validation cases providing additional statistical significance in our validation results. Now the final collocated cases are 63, extending the time period to July 2022.
- In the revised manuscript, we separated the comparison between S5P and EARLINET for satellite pixels over sea and land.

The manuscript is well structured, but I miss a thorough discussion on the findings. Why the comparison is worse over the land/ocean dataset compared to the ocean? Mention previous studies which have already found this known feature of TROPOMI retrievals and make comprehensive conclusions.

The main reason for the underestimation of layer height by the TROPOMI sensor over land is the surface reflectivity assumed in the forward model primarily due to the high surface reflectance in O2A-band that is not favorable for aerosol retrievals. In the revised manuscript we improved the discussion of the finding taking into account the reviewer's comments/suggestions. Concerning the poor performance of TROPMI over land we added in the discussion the following sentence in **Section 2**: "The main reason for the strong underestimation of the aerosol layer height retrieved by the current algorithm from TROPOMI over land is the surface reflectivity climatology used in the forward model, leading in biased or non-convergent retrievals over land. Sensitivity studies showed that the observed large bias over land is reduced when fitting of the surface albedo as estimated from TROPOMI itself was included in the retrieval procedure. This will be further investigated in the near future and is

intended to be implemented in future versions of the ALH L2 product"

Concerning previous studies, we added a new table in the end of the discussion section (Sect. 4) summarizing the outcomes of this study, including the findings of other works (Griffin et al., 2020; Nanda et al, 2020). The present analysis provides important additions to the existing validation studies that have been performed so far for the TROPOMI S5P ALH product, which were based only satellite-to-satellite comparison (e.g. CALIOP and MISR) and confirms in a consistent way the effect of surface albedo in the retrieved ALH. Moreover, the use of high-resolution ground-based lidar data makes feasible a better characterization of the biases found

What is missing from this comparison which would be beneficial in the community?

The distribution of the EARLINET stations allows us to study the temporal, regional and continental-scale representativeness of the observations and to compare these findings with the results of spaceborne passive instruments. However, the limitations of the current version of ALH algorithm over land does not allow a full exploitation of its potential. The inclusion of more stations from continental Europe will improve the significance of the results and will allow to study the impact of different aerosol types (e.g urban, rural, etc) on the comparisons. In addition, it will make feasible to examine possible geographical dependencies.

The Section 4 in the manuscript has been modified highlighting the main points of the above discussion. We think that pinpoints the importance of this work.

In this study, the geometrical features from ground-based lidars were compared against the ALH product from TROPOMI. I was wondering since most of the ground-based lidars used in this study have a lower detection limit at around 700-800m, if not higher, how was this taken into consideration when calculating the ALH? How did the authors tackle the overlap issue in the ground-based lidar observations and what is the error from ground-based lidar overlap limitation to the calculated ALH? For example, in the smoke episode the lidar signal is cut below 1km (Figure 11 and Figure 14). I assume that the bias is not big (equation 1) but given that the attempt is to validate a satellite product this effect should be discussed.

The reviewer raises a crucial issue regarding how the overlap altitude can affect the lidar-based ALH estimates. In the initial submission we did not consider the effect of the overlap in the estimation of the ALH from a lidar measurement.

To overcome this issue, in the revised version, we rely on certain assumptions. To calculate the ALH_{bsc} from the lidar backscatter profiles using equation 1 (Sect. 2.3), for the height range between the surface and the full overlap height we assumed a constant backscatter coefficient (height-independent) equal to the one measured at the full overlap height. This is general acceptable since the Planetary Boundary Layer (PBL) is characterized as well-mixed aerosol conditions (Siomos et al., 2018). This assumption obviously affects the calculation of the lidar aerosol height (ALH_{bsc}) compared to the ones shown in the initial submission since it also considers the contribution of the aerosol load in the lowermost part of the atmosphere. The ALH_{bsc} estimates, when considering this part of the profile, are therefore smaller and the bias with TROPOMI is reduced.

In order to quantify the uncertainty of the ALH_{bsc} estimates due to assumption we made for tackling the incomplete overlap issue we performed a number of sensitivity tests. An example case over Thessaloniki on 15 June 2022 is presented in the **Figure RC1-1** which demonstrate the effect of the overlap altitude on the lidar ALH calculation for different scenarios. Dashed colored lines correspond to the different indicative assumptions for the backscatter coefficient profile below the overlap height. The horizontal-colored lines indicate the corresponding lidar

weighted height. The scenario with a negative slope of backscatter in lower part is indicated in green, the case for vertical extension (zero signal slope) in red, and the third with a positive slope in blue. After applying the above sensitivity tests to all lidar measurements (N=63) used in the study, we can concluded that the effect of the different assumptions shown on the calculation of ALH_{bsc} is of the order of 100 - 400m, depending on the technical characteristics of each lidar system.

It should be also noted that in the calculations described above, the altitude of the EARLINET stations is considered for the calculation of the ALH_{bsc}. Most stations are located at low altitude in coastal areas, so it does not play a significant role, in contrast to stations located at an altitude > 600m such as Granada and Potenza, where the effect is significant.

In the revised manuscript, a new paragraph has been added to present the main points of the discussion above.

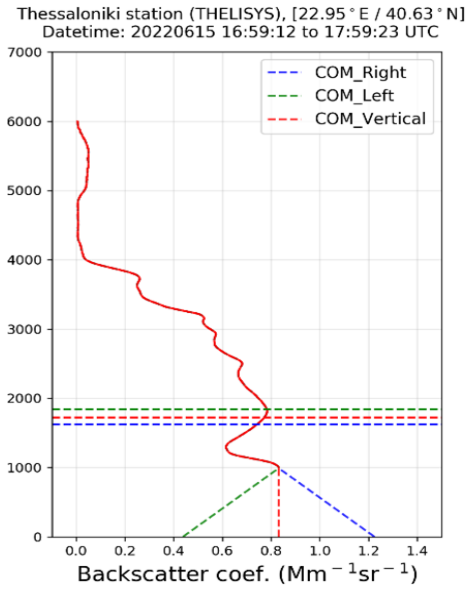

Figure RC1-1. Sensitivity test to define the effect of overlap altitude on ALH_{bsc} calculation under different scenarios (Thessaloniki 15 Oct 2022)- Dashed lines correspond to the different overlap assumption.

During daytime, the Klett method was used for the retrieval of the particle optical properties. I was wondering how the selection of a single lidar ratio (LR) for the whole profile can skew the ALH calculation in the presence of multiple aerosol layers in which the aerosol type is not the same since it affects the particle backscatter coefficient value and therefore the ALH calculation? How many of the 34 cases were Raman cases? Was there any difference in the bias between the Raman cases and the Klett cases?

We use the lidar backscatter coefficient profiles at 1064 nm (or 532 nm), analyzed by the Single Calculus Chain (SCC; https://scc.imaa.cnr.it/) algorithm (D'Amico et al., 2016) for quality-assured measurements. Raman measurements are not used in the study. The Klett-Fernald-Sasano (KFS) inversion is applied (Klett,1981; Fernald,1984; Sasano and Nakane,1984) to retrieve the height-resolved aerosol backscatter coefficients with selection of constant lidar ratios in most of cases, based on the climatology obtained from each station. Following the reviewer's suggestion, we studied the effect on the ALH_{bsc} estimates based on lidar backscatter profile retrievals for different Lidar Ratio (LR) values (20-60sr). We assumed: (a) constant values of the lidar ratio with height and (b) LR height dependent profile. The results show that in both cases the effect of the different lidar ratio values on the weighted height calculation (ALHbsc) is small, lower than 40m.

A relevant phrase has been added in the manuscript.

Technical corrections:

In general, the writing of a scientific article should be impersonal, therefore, I would recommend rechecking these places in the manuscript were the word 'we' has been used. To this direction, there are a few typos and misuse of English language in several places in the manuscript. A careful review is required.

Rephrased

L38: Repentance of the text 'over the Mediterranean basin'. Please, correct.

Corrected.

L43: '…illustrates that TROPOMI ALH is consistent with EARLINET'. A satellite product (ALH) cannot be consistent with a network (EARLINET). Rephrase the sentence.

Rephrased.

L54: 'Aerosol properties are one…' à 'Aerosol properties present one…'

Rephrased.

L77: Nanda et al., (2020) presented a comparison between TROPOMI ALH and CALIOP observations therefore not relevant in the context of this sentence.

We added this reference here as, to our knowledge, it is the most comprehensive paper currently available that describes the TROPOMI ALH. We did not add it as a validation reference.

L80: The EARLINET acronym is already defined in P2/L63. Similarly, in P4/L33.

Corrected.

L139. 'EARLINET measurements…' - > '. Observations submitted to EARLINET database follow….'

Corrected.

L149: 'On the other, during nighttime' -> 'During nighttime,…'

Entire phrase was deleted in the revised manuscript.

L166-L175: I suggest removing this paragraph as its context is not relevant to EARLINET. Section 2.2 or/and the discussion are more relevant candidates.

The paragraph has been removed following the reviewer suggestion. A new re-formulated paragraph added in the discussion section and in Section 2.2

L203: Please, provide the acronym for TOA.

Added.

L208: Provide a reference for OMI/Aura and their corresponding acronyms.

Acronym and reference provided.

L220: 'To construct…TROPOMI observations. Please, rephrase the sentence.

Rephrased.

L251: 'In addition, $ALH_{ext.....}$'. Do the authors refer to the weighted-extinction height? Please, specify.

The authors refer to the "weighted-backscatter height ALH_{bsc} ". Text corrected.

L270: The acronym Z_{COM} is already defined in L246. Please, go through the manuscript and carefully correct the usage of the acronyms. Define an acronym once and then use in the rest of the manuscript. Also give the acronyms that are missing e.g TOA, OMI etc.

Zcom has indeed been previously defined, but in these lines all the different Zs are given explicitly. So as to help the reader follow our work without unnecessary back-and-forth in the text, we opted to also give Zcom a second time.

TOA and OMI are now provided.

L278: 'In the case where more than one layers with a significant contribution to the optical thickness of the profile, an average value.…retrievals' à 'In case more than one layers with significant contribution to the optical thickness of the profile are present,…..'

Rephrased.

L286: 'two selected' à Two or three? In some places it is mentioned 2 in some others 3. I assume three is the correct answer.

It is indeed three, one dust and two smoke episodes

L388: Add UTC next to the time and the corresponding Fig. 4a.

UTC added next to the time.

L559: 'All the input datasets considered in the study have been previously pre-processed at high resolution'. What is high resolution referring to?

The EARLINET lidar profiles (e.g. backscatter profiles in our study) are provided with a height resolution of a few tens of meters (7.5 to 60m) and a temporal resolution of a few minutes.