

Atmos. Chem. Phys. Discuss., referee comment RC3
<https://doi.org/10.5194/acp-2022-412-RC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2022-412

Anonymous Referee #3

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-RC3>, 2022

The paper entitled "Validation of the TROPOMI/S5P Aerosol Layer Height using EARLINET lidars" aims to investigate the ability of the Sentinel-5P Tropospheric Monitoring Instrument (TROPOMI) to derive accurate geometrical features of aerosol layers, through implementation of ground-based observations from the European Aerosol Research Lidar Network (EARLINET). The article falls within the scope of "Atmospheric Chemistry and Physics". The authors utilized the database of EARLINET in order to present a statistical analysis of the ALH retrievals of TROPOMI. I would suggest publication, for the results are of interest for the scientific community implementing and working on S5P, however following major revisions.

Comments:

1. "The purpose of this study is to investigate the ability of the Sentinel-5P Tropospheric Monitoring Instrument (TROPOMI) to derive accurate geometrical features of lofted aerosol layers on a continental scale".

Considering EARLINET, the network spans on a continental – European scale, with stations established and operated even beyond the continental boundaries in recent years. The study focuses on a part of the continent, the Mediterranean Sea region. Thus, I would suggest to the authors to replace the term continental whenever used in this concept. Moreover, it is not clear the reason behind focusing on the Mediterranean Sea region, for EARLINET offers a unique wealth of ground-based observations based on more than 30 established and regularly operated and maintained stations. In this way the correlative dataset of TROPOMI-EARLINET would provide a significantly more extended number of collocated cases, offering the potential to provide more robust results.

Even in terms of a study focusing on the "Mediterranean Sea" region, the Evora station is considered less a Mediterranean station than an Atlantic Ocean station, with a significant number of EARLINET stations falling within the Mediterranean and not included in the study.

Please justify in the study the selection of the stations used and not used in more solid way, for the selection significantly affects the conclusions, due to the effect on the number of cases in the intercomparison.

2. "... key component in the validation of passive satellite aerosol product ...". Please add "passive and active ...".

3. The introduction has to include quantitative outcomes of studies related to previous evaluation or validation of TROPOMI ALH. Moreover, the manuscript will benefit by a table in the end summarizing the outcomes of the study, including lines/rows with outcomes of previous studies (i.e. Nanda et al, 2020).

4. prominent ->key.

5. The manuscript could benefit from references to support the context, for very few are used. Though some examples are provided here of missing references, I would strongly suggest the authors to strengthen the manuscript with references, and go through the text carefully and enrich it.

- ... relatively short life ... (add reference).
 - ... variety of their natural and anthropogenic sources ... (add reference).
 - ... and their different formation mechanisms ... (add reference).
 - ... aerosols exhibit highly variable spatio-temporal distributions around the globe ... (add reference).
 - ... strongly affect both air quality ... (add reference).
 - ... the delicate balance in atmospheric chemistry ... (add reference).
 - ... is essential for understanding the impact of aerosols on the60 climate system ... (add reference).
 - Both active and passive remote sensing methods have been developed, from both ground-based and space-borne systems, in order to estimate the aerosol layer height ... (add reference).
 - ... influenced by the Sahara Desert on the South and the highly populated and industrialized European countries to the North ... (add reference).
 - ... this relatively high aerosol load in the region can have strong effects on the regional radiative budget (add reference), climate (add reference), and ecosystems (add reference) ...
 - ... with frequently observed events of mineral dust and smoke particle (add reference) ...
- And more. Please go through the manuscript.

6. The English language of the manuscript is acceptable for publication; however, it is rather poorly used. The manuscript is characterized by a large number of non-formal, non-highly scientific approach of describing the core of the context of the manuscript. I suggest to the authors to improve the language of the manuscript. Only some examples are provided here, phrases that do not read well..

... to reduce uncertainties in our understanding ...

... Space-based instruments are able to fill this gap ...

... In order to trust and use the space-based products ...

... Mediterranean Sea basin ...

... The lidar technique is the most predominant tool ...

... The large majority of the involved stations is based on multi-wavelength ...

... and are equipped with depolarization channels EARLINET measurements follow absolute

accuracy standards ...
... ground-based lidar measurements from first need to be collected and collocated ...
... , 34 coincident cases were found, checked and flagged ...
... The total available dataset is on the small side but suitable for the comparison study
and general representativeness of the TROPOMI ALH product...
... Only a few data satellite points are available over the land and so a meaningful direct
comparison over land only is not possible...
... Recall that ...
... This example amply demonstrates that when ...
... The smoke arrived over the Iberian Peninsula (IP) in southwestern Europe on 24
October (Figure 6a), just in time for a regular overpass of the TROPOMI over Iberian
Peninsula ...
... the TROPOMI ALH whose reasons warrant further investigation in the future...
Please go through the manuscript.

7. Lines 77-79. Both validate and evaluate terms are used in a sentence. The two terms
are different. The objective is to validate or evaluate TROPOMI ALH?

8. Line 92: strategy -> methodology.

9. Half of Section 1.1. has nothing to do with providing information on the Mediterranean
Sea, for it discusses EARLINET. I would suggest moving the first part to the introduction
and the second (EARLINET) to section 2.1.

10. Frequently the authors provide the extended acronym explaining the terms (or
providing information on last access) two or more times in the manuscript, not necessarily
during the first time when the terms are used. Please elaborate on the issue.

11. Line 138: ... and Raman-shifted signals, and ...
Please check for spaces between words and characters.

12. Line 145: attenuated backscatter or backscatter coefficient?

13. Major comment:

Line 145: A major consideration is that though EARLINET is a high-level network, and is
considered as such the ground-truth for such studies, the present study has done a poor
job in addressing sources of discrepancies in the comparison related to the characteristics
of the reference dataset.

I would ask the authors to extensively discuss:

- 1) The errors of the EARLINET measurements in the study and how they are used and
affect the comparison.
- 2) The strong effect of the overlap. In many cases, the stations are at relatively high

altitude amsl, while an additional overlap may push the comparison even higher, or even over the PBL. Lower aerosol layers detected in the PBL/MBL are observed by S5P, not fully by EARLINET. Which is quantitatively the effect of overlap in the study?

3) Topography plays a key role in satellite-ground based intercomparison of measurements/products. The different stations are characterized by different topographical features, affecting the homogeneity of aerosols in the comparison. A nice example is illustrated in Gkikas et al., 2022, "First assessment of Aeolus L2A particle backscatter coefficient retrievals in the Eastern Mediterranean" – Figure 1. In the present study, the same collocation criteria are used for stations of different characteristics.

14: "In this study, the lidar data were analyzed using the KF method whenever the weather conditions were adequate and the signal quality was sufficient for deriving high-quality backscatter vertical profiles".

However equation 1 is based on backscatter coefficient (?). Therefore, which is the reason why KF is used to compute LRs and extinction coefficients? Please provide more information and explain in the manuscript. If not needed, remove the KF/Raman sub-paragraph.

15. Lines 152-160. Move to end of introduction, where the manuscript provides EARLINET implementation for passive sensors.

16. Lines 161-168. Non-relevant to the study.

17. Section 2.1. Provide a table of the QA procedures applied to EARLINET.

18. Section 2.2. Provide a table of the QA procedures applied to S5P.

19. "The profiles from different types of lidar instruments have to be interpreted in terms of their ALH profile parameter (e.g. height of the assumed single aerosol layer) in a consistent way to reduce mismatch errors due to the significant different horizontal sensitivity between TROPOMI and lidar measurements" -> not clear, please re-write.

20. A better justification on the selection of the collocation criteria should be used. For instance, in Pappalardo et al., 2014, the key connection link was based on meteorology. In Gkikas et al., 2022, more strict criteria are applied, while the air masses to be compared are related to the station measurements in-time through trajectories. The +-4 hours times window raises questions on the homogeneity of the air masses that are compared, for the atmospheric scene may have well changed in an eight-hours' time-window. How did the authors check/ensured that the cases were homogeneous enough to compare?

21. Section 2.3 – number 4.: we use the lidar backscatter coefficient profiles mainly at 1064 nm (or 532 nm), analyzed by the SCC. Which is the reason that the authors did not use only 1064nm or 532nm, and used in some cases the 1064nm and in other cases the 532nm. How did this selection affect the study, due to the different detection and scattering properties of aerosol in 532nm and 1064nm?

22. Does S5P provide pixels “aerosol-free”? If yes, are these cases included or removed from the analysis?

23. “... parameters such as the layer base (ZBASE), layer top (ZTOP), layer thickness (LTH) and center of mass (ZCOM), can be also calculated from the lidar signals...” Please provide the values, including the number of layers.

24. “The total available dataset is on the small side but suitable for the comparison study and general representativeness of the TROPOMI ALH product.” Although the dataset may be sufficient to reach some conclusions, in my opinion, strong statements on the performance of S5P should be avoided or used with caution when the dataset is characterized by a low availability of intercomparison cases.

25. Line 317: add relative bias.

26. Major comments:

A major question in algorithms is not merely “how good are the QA datasets”, but also “how but are the non-QA datasets”. In this case the question is which is the performance of “land only”? The authors should provide some metrics on this major category, as done in the “ocean only” and “ocean-land” categories and comment/discuss although the dataset is not extensive..

Moreover, the authors should provide a separate analysis for the cases that only one layer was detected by EARLINET, and for the cases where two-or-more layers were detected by EARLINET (two clusters), for the structure of atmospheric aerosol layers is a challenging task for a passive sensor affecting ALH, as also discussed in the section of S5P.

Finally, frequently, it is mentioned that it is a challenge of S5P to detect ALH over land due to surface albedo. As an evaluation study, I would expect to provide through a surface-type database an assessment on the effect of different surface types. For instance it is mentioned: “Many factors can play a role in this apparent disagreement between TROPOMI retrievals over land and sea including that high surface albedos negatively influence the ALH, biasing the ALH towards the surface.” Though this is known, and frequently mentioned in the manuscript, as a validation study this is something that should be more extensively addressed.

27. Figure 2a: The datasets are high correlated, according to the results. Could the authors provide to the figure or in the manuscript, the correlation coefficients of the “dust” and “smoke” cases also separately?

28. Figure 5: "TROPOMI is in excellent agreement with the calculated ALH_{bsc} from the lidar profile". However the two layers are ~ 0.5 km retrieved differently. Please avoid so strong statement, not supported.

29. Line 514: In this part should also be discussed the effect of the 760m elevation of the EARLINET station in the intercomparison.

30. "The statistical results show the ability of the TROPOMI instrument to detect aerosol layers under cloud-free atmospheric conditions with significant aerosol load, such as dust and smoke plumes". Following the cases presented, and discussed, avoid a so define statement. Moreover, the number of cases prohibits statements such as "Overall, our results testify that the TROPOMI product complies with the S5P mission requirements" and "This work confirms that the TROPOMI ALH product is within the required threshold accuracy and precision requirements of 1 km."