Review of manuscript 'Validation of the TROPOMI/S5P Aerosol Layer Height using EARLINET lidars', (acp-2022-412) by Michailidis et al. (2022)

Summary

This manuscript presents results of comparisons of aerosol layer height (ALH) derived from Oxygen-A-band (O_2A) observations by the Sentinel5-Precursor Tropospheric Monitoring Instrument (TROPOMI) to ALH's inferred from the vertical distribution of aerosols by European Aerosol Research Lidar Network (EARLINET). Lidar observations at seven EARLINET stations along the Northern Mediterranean coastline yielded 34 Lidar-TROPOMI coincidences. A coincidence defined as the averaged TROPOMI ALH inside150 km radius circles centered at the EARLINET site within an 8-hr window (± 4hr). The lidar inferred ALH is calculated as the 1064 nm backscatter-weighed aerosol altitude.

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.

General Comments

Although the authors have generally carried out a carefully planned validation analysis of TROPOMI ALH, a few details need additional explanation before the article is acceptable for publication.

-The 8-hour temporal window, dictated by the need of getting enough information for the analysis, is probably too long to capture the variability of the dust plumes' structure. If the frequency of EARLINET observations allow it, I suggest adding a figure documenting the typical variability of aerosol load vertical structure during the passage of a dust storm at a representative site. This analysis could provide important information to characterize the uncertainty associated with the adopted ALH definition.

We present in detail in our response to Reviewer 3 (comment 20) the time difference between the lidar measurement and the TROPOMI overpass. The majority of the measurements used are within +/- 2 hours and only two of them extend to +/- 4 hours. However, these cases correspond to persistent dust episodes and were considered useful for the comparisons. Following the

reviewer's suggestion, we present in detail such an intense dust episode originating from Saudi Arabia, on the 24 April 2022 over Eastern Mediterranean, closely to Cyprus. The temporal evolution of the range-corrected signal (RCS) and volume linear depolarization ratio (VLDR) over Limassol and illustrated in the **figure RC2-1**. The plots confirm that presence of a persistent dust layer located in the height range between 2 to 4km, above sea level, for many hours before and after the S5P overpass. The averaged backscatter profiles at 1064 nm ($m^{-1} \cdot sr^{-1}$), calculated with an integration time of 60 min, for the time windows from 07:00 to 08:00, 09:00 to 10:00 and 12:00 to 13:00 UTC (see the yellow vertical dashed boxes), are presented on the upper part, revealing the differences in the vertical structure of the atmosphere between the different time windows. As it is evident from this figure the ALH_{bsc} ranged from 3.1 – 3.3km.

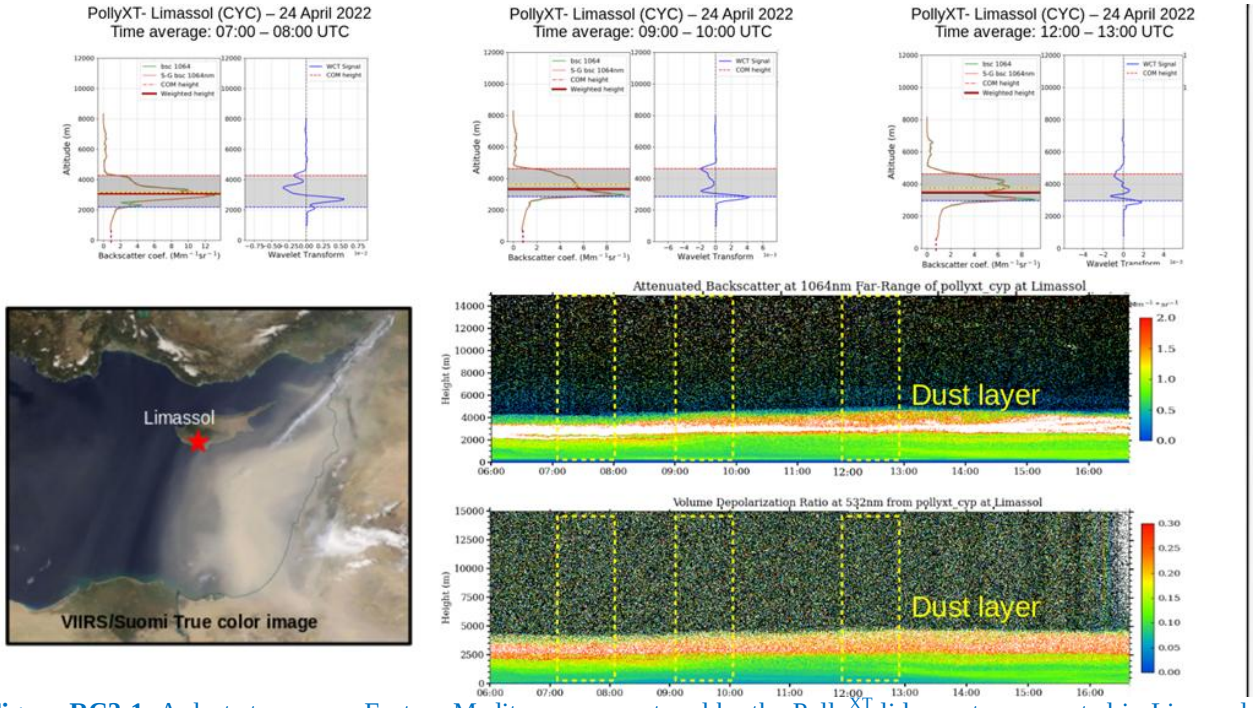

Figure RC2-1. A dust storm over Eastern Mediterranean captured by the Polly^{XT} lidar system operated in Limassol, Cyprus on April 24, 2022 (ADD better description). VIIRS/Suomi NPP True color image on this day also illustrated, generated from NASA Worldview Snapshot (https://wvs.earthdata.nasa.gov/)

-The authors mentioned having considered the UVAI as an indicator of the presence of absorbing aerosols. Please explain the manner the UVAI was used considering the UVAI dependence on height, in addition to the known dependence on AOD and composition. Just having a positive UVAI value is not enough to assume aerosol presence because there are other non-aerosol related sources of positive UVAI values.

An in detail explanation at this point is included in the revised manuscript.

- Given the limited set of coincidences as well as the localized nature of the analysis, I disagree with the authors' over-optimistic conclusion that the TROPOMIALH product meets the expected 1 km threshold requirement of either accuracy or precision. A previously published evaluation analysis (Nanda et al, 2020) demonstrated that the TROPOMI ALH product is systematically lower than CALIOP over both land and oceans.

Taking into account the reviewer's comment, in the revised version of the manuscript we reformulate the text accordingly so that it clearly reflects our claims. The reviewer can also refer to our response to comment of reviewer RC1 & RC3. In the revised version the estimated ALH_{bsc} are smaller and reduce the bias to 0.51±0.77 km as seen in the revised Figure RC3-2 (Rev. comment

-There was not mention of the possible role of calibration. It is well known that thecalibration of the sensor has been drifting. The authors should address this issue and explain how calibration effects could (or could not) explain the observed levelof disagreement in retrieved ALH.

We think that calibration issues are clearly beyond the scope of this manuscript. The technical documentation of the TROPOMI ALH provides error analyses and among many other issues, it describes the effects of instrument errors on the retrieval. The L1B radiances and irradiances are calibrated but will be corrected for degradation effects, in the foreseen v3.0. For this new version of the data an updated validation will be performed, which will allow such an assessment. More details about the calibration aspect can be found in Section 5 on the ATBD; http://www.tropomi.eu/sites/default/files/files/publicSentinel-5P-TROPOMI-ATBD-Aerosol-Height.pdf.

Specific Comments

Line 46. The statement 'This work confirms that the TROPOMI ALH product is within the required threshold accuracy and precision requirements of 1 km' should be removed. The temporarily and spatially limited analysis presented in this work does not provide any basis for such a general and over-reaching.

This comment has been replied to previously in this review.

Line 75. Add TEMPO to the list of upcoming AQ satellites

TEMPO is added in the text. Also, the reference below has been added in the manuscript.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O'Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J., and Chance, K.: Tropospheric emissions: Monitoring of pollution (TEMPO), J. Quant. Spectrosc. Ra., 186, 17–39, https://doi.org/10.1016/j.jqsrt.2016.05.008, 2017.

Line 87 TROPOMI aerosol *height* products

Added

Line 210 Over the oceans positive UVAI also result from non-aerosol sources such as sunglint

and ocean color effects. Negative UVAI can also result from optically thin clouds and aerosols over both land and oceans. Ocean color effects associated with chlorophyll absorption yield negative values over the oceans. Surface spectral dependence over arid and semi-arid regions also generate non-aerosol related UVAI signal. Include original references [Herman et al., 1997; Torres et al 1998] and recent references documenting improvements in the treatment of water clouds in UVAI calculation [Torres et al, 2018]

In the revised manuscript a relevant phrase was added at this point.

Line 244 Elaborate on the cloud screening applied to TROPOMI observations for the comparison to EARLINET observations.

The ALH is very sensitive to cloud contamination. However, aerosols and clouds can be difficult to distinguish. We follow the PRF recommendations to identify possible cloud-contaminated pixels for aerosol retrievals. Cloud masks are available from VIIRS and FRESCO and is strongly recommended to filter for residual clouds. For the current algorithm versions, an operational cloud mask for the off-line processing mode is based on observations by VIIRS aboard Suomi-NPP. S5P flies in formation with Suomi-NPP and observes within approximately 5 min from Suomi-NPP's overpass. These and other sources of uncertainties are indicated with the "qa_value". Use of pixels with a "qa_value" below 0.5 is not recommended. Moreover Cloud flags are available and are strongly recommended to filter for residual clouds. For cloud filtering, the "cloud_warning" flag is the preferred flag for removing possibly cloudy pixels. The flag of "cirrus_reflecatnce_viirs_filter" for residual cirrus clouds is also used. In our case we apply all the above flags to the satellite retrievals.

An extended list of applying cloud flags to filter out possibly cloudy pixels is also provided in detail in Nanda et al., 2020 (Table 1). A full description of the pixel selection scheme is provided in product ATBD and PRF.

Line 245 The term *real* aerosol height is not appropriate in this context. Perhaps *effective* is a better choice.

Altered.

Line 246 Use 'backscatter-weighted' instead.

Altered

Line 445 Add the Torres et al [2018] reference that specifically addresses the treatment of clouds in the UVAI parameterization.

A relevant phrase was added at this point. In addition, the suggested reference has been added in the revised manuscript.

Torres, O., Bhartia, P. K., Jethva, H., and Ahn, C.: Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products, Atmos. Meas. Tech., 11, 2701–2715, https://doi.org/10.5194/amt-11-2701-2018, 2018.

Line 538 It is reasonable to assume that the TROPOMI AER_LH algorithm development team are familiar with the thermodynamic nature of the Earth's atmosphere. So, either remove the sentence '…The AER_LH algorithm was not created to retrieve AER_LH at such low air pressures' or add a specific reference in support of that statement.

Obviously, the statement here needs rephrasing to avoid any misunderstanding. and in addition relevant references have been added in support of the statement, already mentioned in the article. The main point here is (see details below) that the limitation for low pressures exists not because of the thermodynamics but due to the fact that the current algorithm has not been properly trained for very high altitudes (i.e.. low pressures). A relevant explanation is added in the revised manuscript.

The current implementation of the algorithm is based on a neural network forward model and an optimal estimation scheme in the retrieval for spectral fitting with various aerosol layer pressures and aerosol optical thicknesses in the oxygen A-band. The ALH is reported in both altitude and pressure. The limitation for low pressures (i.e. high altitudes) exists because the algorithm has not been trained for very high altitudes. This is perceived by the cases of elevated layers (e.g. smoke, volcanic ash, sulphates) detected by TROPOMI. Currently, the ALH neural network is trained for ambient pressures between 1000 and 75 hPa, and plumes above these heights cannot be resolved.

The discussion above is included in the TROPOMI ATBD. The relevant reference is provided into the manuscript.

Line 542 The analysis presented in this paper shows a real limitation of the TROPOMI AER_LH algorithm, not just a *possible* one.

Agreed, slightly rephrased.

Line 578 Accuracy and precision are two different concepts. Based on the presented results, this reviewer is not convinced that the 1 km threshold requirement of either accuracy or precision has been met. I believe additional analyses are needed.

This comment has been replied to previously in this review.

Line 579 The use of 'testify' is mainly a legal term. It is not appropriate in thisscientific context. It could be replaced with 'show'.

Altered.