We thank all reviewers for valuable comments and suggestions. After considering them, we see that the quality of the manuscript has improved. Work is planned on further development of the retrieval algorithm based on the validation results reported in the manuscript. Some of the comments/suggestions will be considered during the evaluation of the next version of the Sy AOD product.

Response to anonymous reviewer #1

We thank the reviewer for her/his very helpful suggestions

RC1: 'Sogacheva et al. (2022) amt-2022-101', Anonymous Referee #2, 03 Jun 2022 reply

Review for Atmospheric Measurement Techniques

Title: Extended validation and evaluation of the OLCI-SLSTR Synergy aerosol product (SY_2_AOD) on Sentinel-3

Authors: Larisa Sogacheva, Matthieu Denisselle, Pekka Kolmonen, Timo H. Virtanen, Peter North, Claire Henocq, Silvia Scifoni and Steffen Dransfeld

General Comments:

This manuscript presents a very thorough and detailed validation of the SY_2_AOD and related Angstrom Exponent products by comparison to AERONET and MODIS data sets. This analysis provides the user community with the statistics that are required to intelligently utilize these datasets. What is somewhat lacking in many sections (see some specifics below) are explanations and/or reasons for poor performance in the satellite retrieval AOD products versus AERONET measured AOD in some specific regions. This contrasts with much better performance in other regions yet there is little to no discussion on why some regions are much better than others.
There are common reasons why the performance of retrieval algorithms is worse at certain conditions (e.g., cloud and snow contamination), in specific regions (e.g., bright surface), and for specific instrument-related reasons (e.g., influence of the viewing geometry, as for S3). Those reasons are mentioned in the text (e.g., lines 154, 205, 865 as in AMTD)
I think the authors should include much more discussion on the likely algorithmic and/or physical reasons for the discrepancies in the problem regions, much as they did in the last paragraph of the Conclusions section.
As suggested by all three reviewers, more discussion on the likely algorithmic and/or physical reasons for the discrepancies between Sy_2 and reference products was included.
Additionally I feel that this paper is too long with too many multi-panel figures for most readers. I suggest that the authors select a significant fraction of the figures (maybe one third) and associated text and move them to an appendix section. This would significantly improve the readability and clarity of the paper.
One figure and five tables are moved to the Supplement

Specific Comments:
***Lines 28-30**, Abstract: "The retrieval of Angstrom exponent, related to aerosol size distribution, shows good spatial correlation with expected sources but generally overestimates AE for cases where AERONET Angstrom is low, resulting in overall high bias." I think this somewhat overstates the accuracy and utility of the satellite retrieved AE. The regional AE comparisons in Figure 24 show very poor accuracy for most regions in the satellite AE product. I suggest removing this sentence from the abstract or making a more quantitative statement on the retrieved AE accuracy.*
The statement on the AE is re-formulated.
Similar comments can be applied to the poor retrieval accuracy of the satellite FMF in Figure 22, except for good agreement at the highest AOD levels.
Conclusions on the FMF and FMAOD are added to the abstract
***Line 172-173**: Please describe somewhere in the text how is AE computed from FMF.*
The section 2.2.3 is now clarified: "During post-processing, further aerosol outputs are derived from the retrieved AOD_{550} and FM AOD. This includes spectral variation of AOD, which is given using pre-computed look-up table

from the retrieved FM AOD and aerosol mixture. The Angstrom exponent is computed based on a pair of spectral AOD values. Here we choose 865nm and 550nm."

***Line 176**: Typo, I assume 'duct' is supposed to be 'dust'.*

The typo is corrected

***Line 196-197**: Please provide a brief explanation as to why the back scatter at the TOA is more critical in the northern hemispheres versus the southern. Is this just because the percentage of land in the SH is much lower? This is an example of a general lack of physical/algorithm explanations for anomalies and/or comparison results in this manuscript.*

The text for this at line 181 has been rewritten: "In the NH, the SLSTR oblique scan generally samples backscattered radiance, which has a weaker aerosol contribution than the corresponding forward scattering sampled in the SH (e.g., https://www-cdn.eumetsat.int/files/2021-09/SARP_Report_Option_1_final.pdf). This leads to reduced quality in AOD in the Northern Hemisphere (NH) compared with Southern Hemisphere (SH) for the SLSTR products, which has been revealed earlier (https://climate.esa.int/media/documents/Aerosol_cci_PVIR_v1.2_final.pdf). For this reason, SY_2 AOD products from the NH and SH were validated separately."

***Line 250**: 'was be' should be 'has been'*

corrected

***Line 265-266**: It might be noted that the MAN instruments are calibrated against the same reference instruments as utilized in AERONET. These reference instruments are calibrated by Langley method at Mauna Loa Observatory to an accuracy of 0.002 to 0.005 in the visible and near IR and ~0.009 in the UV.*

The sentence is added, as suggested

***Line 287**, Section 6.1: Since AERONET does not measure at 550 nm, please note the spectral interpolation method used. Note that the quadratic or 2nd order fit of AOD versus wavelength is more accurate than the linear or Angstrom fit.*

AE fit was used for interpolation; clarification is added

***Line 295-296**: It seems the word 'error' or 'bias' may be missing here. How could 91% of AOD be < 0.04? This AOD level is too low for the majority of the earth.*

The typo (0.04) is corrected to 0.4

***Line 311**: Please define the acronym GCOS here.*

GCOS acronym is added

***Line 380**: Please provide some reasons or explanation for the smaller retrieval errors in the SH.*

This has been now summarised at line 196 (see earlier comment).

***Line 396-397**: An obvious missing region is the Pacific Ocean since oceans dominate the Earth's surface (70%). The Arctic Ocean, Indian Ocean, Southern Ocean, are also very important. Why were these regions not included?*

Validation was performed over Pacific and Indian islands where AERONET stations are located. However, the number of the matchups is critically low over those ocean regions to provide solid conclusion.

***Line 408**: It is surprising that the performance is poor for Europe. An explanation of the reason is warranted here.*

For these three comments (408, 409, 414) we feel the correct place to address these discussing is in the conclusion/discussion section, which has now been extended and addresses these points.

***Line 409**: The scatter and results for the boreal forest region are very poor. This is surprising since the surface is dark (green forests) and the aerosol type is dominated by fine mode (biomass burning smoke). Please explain/discuss the causes of the poor accuracy retrievals in this region.*

This comment is answered above

Line 414: *An explanation is certainly needed/expected for the large regional differences in the fraction of pixels in EE.*

This comment is answered above

***Line 417-419**: The Aus and AOb regions both had very low AOD, none>0.3 so that is a major factor. This should be mentioned in the text otherwise it is somewhat misleading to the reader.*

Details suggested are added to the text

***Lines 444-448**, Section 6.1.4: This is an awkward writing style to have a section consist of mainly one line equations and short statements, with no full sentences. I suggest trying to expand it a little to make more readable.*

Section 6.1.4 includes three sub-sections. In the introduction to this section (lines 444-448) we provide only a definition of the relative offset, which is analysed and discussed with respect to different variables (latitude, surface reflectance, ets.) in sub-sections 6.1.4.1-6.1.4.3.

***Line 467-468**: In Figure 9 I am missing the separation of NH and SH data that you suggest here. Is there a missing label or legend in this figure?*

There was a typo in the text. The analysis was performed not for the globe, NH and SH, but for dual, singleN and singleO matchups. The sentence was revised

Line 521-523: *The AOD decreases significantly as wavelength increases (except for dust). This may be part of the reason for the offset and rms to decrease as wavelength increases.*

Clarification is added to the text

There is almost consistently a lack of explanation for the observations/comparisons in this manuscript.

The main goal of this work (performed in the frame of ESA LAW project) was to evaluate SY_2 AOD product, reveal problems in the retrieval algorithm and notify algorithm developer and potential users about algorithm performance and product quality in different conditions. We also showed that quality is different for different approaches (e.g, dual or single). In case reasons for limited quality were clear (e.g., back scattering contribution, cloud/snow contamination, bright surface), they are mentioned in the text. However, often a throw revision of the retrieval algorithm is needed to find a reason for a limited performance. This work is planned.

Line 642-643: *This is too vague, it does not really say how the AERONET fine mode AOD from SDA was estimated at 550 nm from the 500 nm product. Please provide more detail here.*

A link for the $aAOD_{500}$ to $aAOD_{550}$ conversion is provided

Line 733-734: *A bias in AE of ~1 and rms of 0.5 effectively renders the satellite retrieval of AE as almost useless for most applications. This should be discussed or summarized in the text.*

The AE in table 7 (which is now moved to the supplement) shows consistent positive correlation with AERONET values, albeit with low R values. We see similar patterns in the retrieval of FMF by MODIS as with SLSTR (new Fig. 28), and SLSTR uniquely gives continuous retrieval over land.

Line 735: *By what metric is this syAE considered 'good' quality? I cannot agree with your assessment unless you define 'good' more clearly.*

We move the description of these as ‘good’, and more simply report the performance.

Line 740: *Validation over ocean: Why are the AE retrievals not compared for over ocean? This would be a useful comparison/validation to include.*

MAN AE (mAE) is provided for 440-870 nm only; Direct comparison between mAE and syAE is not possible

Line 793-794*: Any ideas or explanation about this large difference between MODIS and Sentinel S3A retrievals over Nigeria? This is a striking gradient in large AOD differences, both positive and negative. Which one is more likely to be closer to reality? This is another example of the lack of analysis in giving some explanations in this paper.*

A reason for the large difference is still unclear. We looked at syFMF and modFMF products (new Fig. 27 in the revised version), but modFMF (provided in MOD04_L2 product) is often missing over land. The reason for the luck of explanation is mentioned above (after comment to line 521-523)

Line 815-816*: The way this sentence is written is confusing and does not make too much sense. Please rephrase and clarify.*

The sentence is rephrased

Line 884-889*: This type of analysis and reasons for biases and differences, while good, are mostly lacking in the main text of this paper. It is strange to wait until the Conclusions section to provide this type of analysis.*

We expanded discussion on the reasons for biases and other differences, where reasons for those were clear. To explain some biases, a through revision of the retrieval algorithm is needed.