We would like to thank the Reviewers for the thoughtful comments and suggestions. We have modified the manuscript according to these suggestions. Our replies are below (Reviewer's comments in *Italic*, response in normal face).

Other changes than those suggested by the Reviewers were applied to the manuscript during the revision.
The merging approaches have not been modified, but the number of the products used for merging was revised. TOMS, OMI and EPIC products, which were reported at other than 0.55 µm, were removed from merging. However, we keep those products in the inter-comparison and evaluation with AERONET exercises.
Another product, AVHRR NOAA (over ocean) was added.
In the merging approach 1, the reference to estimate the average offsets with individual products was re-considered: ATSR_ens was replaced with Terra DT&DB. With this change, the overlapping period exists between the reference and all individual products, thus the direct inter-comparison is possible (in the version submitted to the ACPD the offsets between ATSR_ens and VIIRS and ERIC were calculated in two steps, with estimating intermediate offsets to MODIS Aqua).
Section on the estimation of uncertainties in the L3 merged AOD product was added; the spatial and temporal uncertainties are shown and discussed.
In Sect. 6 (revised version), when we discuss the results from different methods for merging annual time series, we now show TOMS (over land) and AVHRR NOAA (over ocean), both shifted to the merged time series (shifted to the reference in the version submitted to the ACPD).
We thoroughly revised the paper, which required an input from a new co-author.

Interactive comment on "Merging regional and global AOD records from 15 available satellite products" by Larisa Sogacheva et al.
Anonymous Referee #2
Received and published: 23 July 2019

Review report of the ACP manuscript Merging regional and global AOD records from 15 available satellite products (https://doi.org/10.5194/acp-2019-446), by Sogacheva et al.
This manuscript discusses approaches to merge satellite AOD data sets from a large number of datasets derived from various instruments. The analyses start at the monthly AOD L3 products at a low spatial resolution (1x1◦ lat-lon). An extensive intercomparison of the various datasets is performed and different merging techniques are discussed.
The strongest part of this manuscript is the section 4, where the different datasets are compared. This could be a publication on its own. The weaker parts are the sections 5 and 6, which should be significantly improved in structure and readability.

Main comments

One of the goals of the manuscript is to present a merged dataset. However, different merging methods are described, and no clear recommendation is made for a merged dataset. Also, a description of the final dataset is lacking. Therefore, the claim made in the abstract that a merged dataset is introduced is not fulfilled. If a dataset is presented its contents should be described, including on the technical level (in an appendix). Also, the dataset should be made available, preferably on one of the large datacentres, and with a doi.

According to the Reviewer's comments, we revised considerably sections 5-7, where methods for merging are described. The scheme for the merging approaches was added to the introduction for merging approaches (Sec. 4 in the revised version).

All merged products are now described and validated (Section 5 in the revised version). Section on the pixel-level uncertainties for the final L3 merged products is added. The recommendations are given on the final merged product. The merged data set will be openly available at Finnish National Satellite Data center, http://nsdc.fmi.fi/ ; the full link will be provided in the manuscript accepted for publication.

The intended audience for the manuscript is not clear to me. If the intention is to describe a merged dataset, the intended reader is a potential user of that dataset. This user group is probably not an expert in the aerosol field and is probably not (so) interested in the performance of the individual underlying datasets (section 4, which is the largest part of the manuscript), but rather in a description of the performance and caveats of the merged one. This needs to be taken into account in the sections 5 and 6, which should be written at the right level, and more or less separate from section 4.

So far, individual products, which have certain limitations discussed in the manuscript, have been used in the air quality and climate studies. We expect that the potential users, if not experts, have a good knowledge on aerosols. For the merged product offered here, which has the main advantage in better temporal overage with similar or better quality, we expect the same audience, which use other individual satellite products, but since the merged product allows looking at the longer period, the climate researchers will benefit from having access to the longer data set.
The interest to the AOD merged product was shown by the AeroCom community. Several request from modelers have already been obtained for evaluation of the modelled AOD products.

To make the manuscript useful for experts in different fields, the discussion of the performance of the merged products has been considerably enlarged by including the evaluation results for all tested merged product and inter-comparison of the selected merged product with individual products.

To summarize my main comments:

- Make the merged datasets available and include a technical description.

- Rewrite the sections 5 and 6 with the intended user of the dataset in mind as the audience. More comments on section 5 and 6 are found below.

- Make clear what the final advertised merged data set is.

The technical description has been revised and supported by the results (Section 5 and 6 are combined).
The scheme for the merging approaches was added to the introduction for merging approaches (Sec. 4 in the revised version).
New section, there the merged L3 monthly products are introduced, evaluated, inter-compared is added.
The main merged product is chosen and inter-compared with individual products.
Section on the pixel-level uncertainties for the final L3 merged products is added.

Section 5

This section should be rewritten, to clarify what was done and limited to methods that are used in the further analyses.

Section 5 and 6 are combined. The technical description has been revised and supported by the results

Section 5.1. This section is too brief and starts with a statement why the mean is not a good statistical indicator, whereas the it is one of the parameters that is calculated. What is missing is information on which data it is applied (to the monthly mean L3, or also to the seasonal and/or annual L3?).

The "mean" approach has been removed from the manuscript.
It was clarified in the text, that the merging has been applied to L3 monthly dataset and annual/seasonal/monthly time series.

Section 5.2. This section is too brief and unclear. With the information contained in this section I would not be able to reproduce the results. The ATSR_ensemble is not available for the entire dataset. How do you deal with this? Clarify all the steps of the method.
Section 5.2 was combined with Section 5.1. The offset correction method was supported by the offset correction results.

Section 5.3. This section describes to methods: RM1 and RM2. However, RM1 is -as far as I can tell- not used in the rest of the paper. Therefore, it should be removed from this section, so approach 3 is limited to RM2 (in the remainder of the manuscript reference to RM2 should be changed to Approach 3). Furthermore, I propose to add one or more equations to clarify the procedure. Also, it should be clarified on which datasets it is applied, because if I understand section 6 correctly, there are also some sub-methods here (e.g. regional weights, monthly weights versus time-series weights, aerosol type weights).

In the revised version, results for RM2 are shown and inter-compared with the results from other merging approaches.
The equation for the weighted mean is added.

The datasets, on which the approaches and sub-methods are applied, is discussed in the new Sect. 5, where the merged L3 products are introduced, evaluated, and inter-compared.

Section 5.3 In approach 1 and 2 the mean, median and standard deviations are calculated. Why is this not done for approach 3 (see for example https://en.wikipedia.org/wiki/Weighted_median)?

The conclusion on the choice of one final merged product is based on the fact, that there is a very small deviation between L3 products and time series merged with different approached (1 or 2) and sub-methods (aerosol types), except for approach 1 applied to the shifted products. Thus, we do not see the need to calculated average (median) from all tested merged product, which, if done, makes the analysis more complicated with no significant improvement.

Section 6

Section 6 should in my opinion describe the quality and caveats of the merged data using method 1, 2 and 3. It should not describe the performance of individual datasets. I think that part of my confusion seems to come from what is called the "merged product". As a reader, I think that methods 1,2 and 3 all yield a merged product but using different merging methods.

In Sect.6 (ACPD version), which is combined with Sect.5 in the revised version, it is important, in our opinion, to show the results for individual products because those results contribute to the final ranking of the products. In the revised version, the results are shown for two regions only, Europe and ChinaSE, as an example. The results for all regions are moved to Supplement.

Section 6.1.1 I think this section doesn't belong in section 6. It describes the rationale for merging approach 2 and therefore should be moved to section 5.2.

Section 6.1.1 was moved to Section 4.1 in the revised version

Section 6.1.2: The title of this section is not covering the contents: in the current manuscript it is comparing the Merging Methods 1 and 2. However, I don't understand why Method 3 is left out in this section. Instead, I propose to describe the comparison of all three minutes, using figure 13 and to drop figures 9 and 12.

Approaches 1 and 2, as in the ms submitted to the ACPD, are combined in the revised version. Annual time series, calculated with the approach 1 (uncorrected AOD) and approach 2 (weighted AOD, former approach 3) are now introduced in Sect. 6; differences in the results from different approaches and different steps (time series from merged L3 product and merged time series) are discussed.

Section 6.2.1: This section described the weights; it doesn't assess the merged data quality. I strongly suggest moving this section to section 5.3, which also increases the readability of that section.

Section 6.2.1 was moved to Sect.4.2, where both method and results (weights) are discussed

Section 6.2.2: first paragraph. This describes sub-methods of approach 3 and should be described in section 5.3. Figure 11, I like this figure, but why include it only for method 3? I think it should also be generated for methods 2 and 3 and the differences discussed.

In the revised version, the evaluation results for all products are shown and discussed (Sect. 5.1). The evaluation results for the chosen product (former Fig.11) are summarized in Sec. 5.2.1. Uncertainties for the final merged product (M) are introduced and discussed in Sect.5.2.2

Section 6.2.3. In the first 2 sentence 2 sub-methods are introduced of approach 3. This is not the right place, this should be done in section 5.3. The remainder of section 6.2.3 should be moved to 6.1.2, and also differences with the other methods should be described.

Section 6.2.3 was partly combined with Sect.4 (revised version) and partly moved to Sect.6 (revised version). The differences in the time series merging results are discussed in Sect 6; results are summarized in Table 3.

Section 7. I don't really see the need for this section. Line 1-22 would fit with the comparison of the time series of the three methods (e.g. 6.1.2). The last paragraph should be moved to the conclusion.

Section 7 (Sect.6 in the revised version) has been considerably revised. The merged time series are introduced, the difference between them is discussed.
The last paragraph was moved to the conclusion.

Specific comments
I strongly suggest adding a figure with timelines of the availability of all the products as part of the section 2. This information is also in Table 1, but a graphical overview would be a great help.

The information on the availability of the products use for merging is now summarized in Table 2. The availability of the other products (TOMS, OMI, AVHRR NOAA) is mentioned in the text.

Table 1 presents the datasets, but the doi's and url's in Table 2. For each dataset the reference of doi (or url if doi is not available), should be included in Table 1, and the reference doi's and urls should be included in the list of references. Table 2 can be removed.

According to the ACP rule, that data availability should be included as a separate section. In the title for Table 1, we added the information that the data availability is summarised in Table 4.

Page 14, lines 1-11. In the discussion of the comparison of AERONET with AOD L3 data, instead on the more common comparison with L2 data, one argument is missed.

We discuss the L3 validation in Lines 3-11. To make it more clear, we now mention L3 specifically.

When L2 data is compared with AERONET with strict temporal and spatial criteria, the L2 data is implicitly cloud-cleared, because the AERONET data is only available under these conditions. This does not hold when comparing the L3 data. If the cloud clearing is not optimal, this would lead to difference in the comparison results of L2-AERONET versus L3-AERONET.

The problems related to the difference in cloud screening are mentioned in abstract, Introduction, Sect.2 and conclusions, briefly or with some details.

Page 15, line 5 "manuscript" -> "work"

Corrected

Page 16, line 16-17. It is not clear what is meant here. What does "different surface treatment" mean (compared to what?).

Different approaches for surface treatment in different products. Clarification is added.

Page 19, section 4.3. Define how the ATSR_ensemble is computed.

The definition for the ATSR_ensmble product is added to Sect. 2.2

Caption Figure 5, In light of my comments on section 5-6, I don't understand which merged product is shown as "M" in this figure.

The caption was revised by including the reference to the merged product

Page 22, section 5.1. I would suggest to not only compute the standard deviation, but also percentiles, for example the 10th, 25th, 75th and 90th, because the standard deviation is very sensitive to outliers.

The standard deviations were low, thus the contribution of the outliers, if existed, was negligible.

Page 22, line 5: "AOD weighted" is not clear. I suggest "Weighted mean, where the weights are derived from the comparisons with AERONET.

The whole paragraph was re-written

Page 23 line 26: "ATRS" should be "ATSR".

Corrected, as suggested by the Reviewer

Page 29, line 1: "aerosol particles" should be "aerosol types"

Corrected, as suggested by the Reviewer

Review report #3, acp-2019-446

The current study is divided in two parts. In the first one, an intercomparison among the most widely used aerosols datasets (obtained from spaceborne passive sensors) is performed while at the second one, various merging techniques are applied towards the development of a unified AOD product. To realize, monthly AODs derived by 15 satellite databases are analyzed over the period 1995-2017. The submitted manuscript is too long thus making difficult to the reader to get the information in a straightforward way. Moreover, the authors should make an effort to provide a better description of their merging methodology and their interpretation of the associated findings. Between the two parts, the first one (up to Section 4) can be a stand-alone paper without making substantial modifications while the second one (from Section 5) needs a lot of improvements regarding the description, interpretation and figures (e.g., add legends wherever do not exist, better description in the captions). Therefore, my recommendation to the authors is to split the current version of the manuscript in two separate works thus helping any potential reader to understand the overarching goal of the study as well as its components and the obtained scientific results. Below are listed my comments that should be addressed prior the publication of the submitted text.

The main reason for keeping those two parts of the analysis together is that the evaluation results for the individual products are used for merging. To shorten "part 1", Sect. 4.1 (AOD spatial distribution and diversity) and 4.4 (AOD annual cycles) and some figures from other sections were moved to Supplement; the discussion on the quality of the individual products was also shortened a bit.

1. ***Page 4; Lines 14-16:*** *Is not clear what the authors want to say here.*

The sentence was replaced with "Whereas a lack of diversity among data sets does not mean that they have converged on the true value e.g., AErosol RObotic NETwork (AERONET, Holben et al., 1998) AOD, which is a recognized standard for instantaneous AOD reference, the existence of unexplained diversity does imply they have not."

2. ***Page 8; Lines 2-3:*** *According to Table 1, there are not available data for 2000 from the AVHRR. Why don't you use 2001 in order to have full temporal coverage also from MODIS-Terra and MISR?*

TOMS is not reliable in Nov-Dec 2001. Thus, year 2000 was chosen.

3. ***Section 2:*** *Is there any criterion applied in the monthly products aiming at improving their quality (e.g., temporal representativeness, best quality retrievals) or just the raw products are utilized?*

Monthly data from the open sources or obtained from the data providers was utilized, except for MISR and AVHRR NOAA (included into analysis during the revision), for which the monthly products were reported at lower resolution. For those, a simple averaging to 1° was applied to match the other products (details in Sect. 2.2). Most of the products do not include the quality flags.

4. ***Page 9; Lines 14-15:*** *Has any significance this threshold?*
In the revised version, the actual number of the maxima offset (0.011) is given.

5. ***Section 4.1:*** *It would be useful to add a table with the AOD averages over continental and maritime surfaces as well as for the whole globe.*

The absolute median AOD numbers for land/ocean/globe for years 2000, 2008 and 2017 are in the upper panel of Fig.2; Offsets from global/land/ocean averaged AOD is given for all individual products, when available. Thus, the actual AOD for each product can be easily calculated.

6. ***Page 10; Lines 26-27:*** *Where exactly? In the storm track zone (emission of marine aerosols due to strong winds) of the Southern Hemisphere or in the Southern Atlantic Ocean attributed to the transport of biomass aerosols from the central/south parts of Africa?*

We added a short discussion on the possible contribution of the storm track zone to the elevated AOD over Southern ocean and provided the reference.

7. ***Page 12; Lines 25-26:*** *Similar diversity levels are also encountered in the US, Mexico, S. America and Tibetan Plateau. Is there any explanation for that?*

In S. America (Amazon), the difficulties/differences in cloud screening might be an issue. Big events of the forest fires might be screened as cloud in some products. To check that, L2 AOD and cloud screening results should be intercompared, which is out of the scope of that manuscript. Same for thick dust events. Different assumptions in bright surface treatment might cause another offset in AOD. However, the AOD diversity is changing there, related to the time period and availability of the products, while over Australia the deviation remains constant along the time.

8. ***Page 14; Lines 23-24:*** *The defined thresholds of Ångström exponent must be modified in order to create a buffer zone between fine and coarse aerosols modes. For example, fine and coarse particles can be "identified" when Ångström is higher and lower than 1.2 and 0.8, respectively. Even though the proposed limits are not the optimum, they are more realistic than the selected ones. An another solution could be the selection of representative AERONET stations for specific aerosol types or aerosol mixtures.*

First, we wanted to be consistent with previous studies (Sayer al., 2018a, Sogacheva et al., 2018 a, b). The other reason is that in monthly aggregates the aerosol types are defined as a median from the 1-month period, while the presence of other types is possible and often obvious. Thus,

such a strict differentiation of the aerosol types, suggested by the Reviewer, is not of great importance in the current study. We included to the Supplement the figure (Fig. S5), where for each AERONET station the prevailing annual and seasonal aerosol type has been estimated based on the chosen criteria. There is a sense in the results obtained, which confirms the applicability of the aerosol type classification suggested in the current study. The results also show that aerosol types differ from one Aeronet stations to another in the same region, thus the prevailing aerosol type can't be defined with a high confidence for the chosen regions.

9. ***Page 14; Lines 31-33:*** *I don't agree with the regional averaging of AERONET observations. Instead of giving equal weight on each AERONET site, it would be more correct (representative) to calculate the statistics on the whole AERONET dataset for each region.*

We tested the approach suggested by the Reviewer, when the study was planned. The validation results for specific areas were often similar with two approaches. However, following the logic that the weight of the validation results might be biased toward the longest time series from a few AERONET stations in the particular area, which are not fully representative for the big region, we chose the other validation approach, explained in the manuscript.

10. ***Page 15; Lines 3-4:*** *How has been defined the spread envelope? Why don't you use only the uncertainty limits defined by GCOS?*

The results for the spread envelope, defined in Sect 4.2 of the version submitted to ACPD, are removed in the revised version.

11. ***Section 4.2.1:*** *The authors should guide better the reader by adding colors corresponding to aerosol groups in Figure 4. Also, rephrase the sentences in lines 13-15 and 25-27. In Figure 4, in the y-axis write that the difference is defined as satellite-AERONET, add a legend and rewrite the caption. Moreover, which is the background AOD? Are there available results for the total AOD without considering different aerosol classes?*

The legend with the explanation for the colours was added

The explanation to the background AOD was given in the text and now added to the figure caption. The evaluation and the following merging were performed also for all aerosol types (total AOD).

12. ***Figure 5:*** *Clarify that the offset is defined as satellite-AERONET.*

Clarification was added to the text and y-label caption.

13. ***Page 22; Lines 8-10:*** *Rewrite this sentence because it is not clear.*

The whole paragraph was revised.

14. ***Section 5:*** *Definitely, a better and more analytical description of the applied merging approaches is needed explaining the benefits and the drawbacks of each methodology.*

The scheme for the merging approaches was added in the introduction for merging approaches (Sec. 4 in the revised version).

The description of the applied merging approaches has been expanded and supported by further discussion of the results.

Section on the pixel-level uncertainties for the final L3 merged product is added.

15. ***Section 5.3:*** *In the RM2, why the levels are 10 and not 9 according to the discrimination of the computed statistics? For example, for the correlation coefficient they have been defined equalrange bins between 0.5 to 1 with a 0.05 step. If I have understood correctly this corresponds to 9 groups of R values instead of 10.*

With the 0.05 step, 10 bins (groups) exist between 0.5 and 1

1	0.50	0.55
2	0.55	0.60
3	0.60	0.65
4	0.65	0.70
5	0.70	0.75
6	0.75	0.80
7	0.80	0.85
8	0.85	0.90
9	0.90	0.95
10	0.95	1.00

16. ***Section 6:*** *The overarching goal of the current study (stated clearly in the title) is to merge different satellite databases. However, it is not clear to me which is the optimum methodology that should be followed. Also, I fully agree with the rearrangements proposed by the Reviewer #2.*

The manuscript has been revised considerably by adding the results from the intercompariosn between the products merged with different approached and considering different aerosol types. Based in the inter-comparison results, one merged product was chosen. The summarised validation results for that product are shown in the new section, which also now includes the inter-comparison between the merged and individual products.

17. Page 23; Lines 26-27: *Please explain better this sentence.*

Terra DT&DB was chosen as a reference for offset correction in the revised manuscript. The text was revised accordingly.

18. ***Figure 8:*** *Check if the shaded area corresponds to ±0.04.*

Checked. The shaded area corresponds to ±0.03

19. ***Section 7:*** *It is not clear why this Section is important.*

In the revised version, the merged annual/seasonal/monthly time series are introduced in Sect.6. Difference between time series merged with different approaches is discussed.

20. ***Page 37; Line 7:*** *What do you mean "…AERONET monthly mean gridded dataset…"?*

"gridded" is removed.

Merging regional and global AOD records from major~~15~~ available satellite products

Larisa Sogacheva[1], Thomas Popp[2], Andrew M. Sayer[3,4], Oleg Dubovik[5], Michael J. Garay[6], Andreas Heckel[7], N. Christina Hsu[8], Hiren Jethva[3,4], Ralph A. Kahn[8], Pekka Kolmonen[1], Miriam Kosmale[2], Gerrit de Leeuw[1], Robert C. Levy[8], Pavel Litvinov[9], Alexei Lyapustin[8], Peter North[7], Omar Torres[10], Antti Arola[1]

[1] Finnish Meteorological institute, Climate Research Program, Helsinki, Finland
[2] German Aerospace Center (DLR), German Center for Remote Sensing (DFD), Oberpfaffenhofen, Germany
[3] Goddard Earth Sciences Technology And Research (GESTAR), Universities Space Research Association, Columbia, MD, USA
[4] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[5] Laboratoire d'Optique Atmosphérique, CNRS – Université Lille, France
[6] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[7] Dept. of Geography, Swansea University, Swansea UK
[8] Climate and Radiation Laboratory, Earth Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[9] Generalized Retrieval of Atmosphere and Surface Properties SAS, Lille, France
[10] Atmospheric Chemistry and Dynamics Laboratory, Earth Science Division, NASA Goddard Space Flight Center, MD 20771, USA.

Correspondence to: Larisa Sogacheva (larisa.sogacheva@fmi.fi)

Abstract. Satellite instruments provide a vantage point ~~to~~ for studying aerosol loading consistently over different regions of the world. However, the typical lifetime of a single satellite platform is on the order of 5-15 years; thus, for climate studies, the usage of multiple satellite sensors should be considered. ~~This paper assesses some options for creating merged products from an ensemble of 15 individual aerosol optical depth (AOD) data records produced from a broad range of institutions, sensors, and algorithms.~~

Discrepancies exist between aerosol optical depth (AOD) ~~AOD~~ products due to differences in their information content, spatial and temporal sampling, calibration, ~~retrieval algorithm approach, as well as~~ cloud masking, and other algorithmic approach/assumptions. Users of satellite-based regional AOD time-series are often confronted with the challenge of choosing ~~the~~ an appropriate dataset for the intended application. ~~In this study AOD products from different sensors and algorithms are discussed with respect to temporal and spatial differences.~~

In this study, 16 monthly AOD products obtained from different satellite sensors and with different algorithms were inter-compared and evaluated against Aerosol Robotic Network (AERONET) monthly AOD. Global and regional ~~comparison of AOD monthly aggregates with ground-based AOD from the Aerosol Robotic Network (AERONET)~~analysis indicates that ~~different~~ products tend to agree qualitatively on the annual, seasonal and monthly time scales, but may be offset in magnitude.~~agree qualitatively for major aerosol source regions on annual, seasonal and monthly time scales, but have regional offsets.~~ Several approaches ~~are~~ were then investigated to merge the AOD records from different satellites and create an

optimized AOD dataset.~~, based on evaluation and inter-comparison results.~~ ~~Global and regional comparison of AOD monthly aggregates with ground-based AOD from the Aerosol Robotic Network (AERONET) indicates that different products agree qualitatively for major aerosol source regions on annual, seasonal and monthly time scales, but have regional offsets.~~ ~~All merged regional AOD time series show highly consistent temporal patterns illustrating the evolution of regional AOD.~~ With few exceptions, all merging approaches lead to similar results, indicating the robustness and stability of the merged AOD products.~~reassuring the usefulness and stability of the merged products.~~

~~Here~~In this paper, we introduce a gridded monthly AOD merged product for the period 1995-2017~~, which provides a long-term perspective on AOD changes over different regions of the world~~. We show that the quality of the merged product is as least as good as that of individual products. Optimal agreement of the AOD merged product with ~~the~~ AERONET further demonstrates the advantage of ~~the~~ merging ~~of~~ multiple products. This merged dataset provides a long-term perspective on AOD changes over different regions of the world, and users are encouraged to use this dataset.

1 Introduction

Interactions of atmospheric aerosols with clouds and radiation are the largest source of uncertainty in modelling efforts to quantify current climate, and predict climate change (IPCC, 2018). To reduce such uncertainties, we need observations to constrain ~~these~~ climate models. However, these observations must be accurately calibrated and validated, have consistent or at least well-characterized uncertainties, and provide adequate temporal and spatial sampling over a long period of time.

With their ability to cover the globe systematically, satellites provide this global and temporal perspective. Satellite observations have produced major advances in our understanding of the climate system and its changes, including quantifying ~~by retrieving~~ the spatio-temporal states of the atmosphere, land and oceans, and aspects of the underlying processes. However, ~~since~~ as the typical lifetime of a single satellite platform is on the order of 5-15 years, a single sensor data record may not be long enough to discern a climate signal (WMO, 2017). Moreover, aerosol products from different satellites and algorithms all have limitations regarding their spatial and temporal coverage and vary in their accuracies depending on environmental conditions (aerosol loading and type, surface brightness, observation geometry). Thus, the application of satellite observations for climate change studies requires using products from multiple sources to derive consistent regional conclusions.~~Moreover, aerosol products from single satellite sensors are often not sufficient due to limited spatial and temporal coverage and the need to avoid contamination by clouds. Thus, the application of satellite observations for climate change studies requires using products from multiple sources.~~

The key parameter used for ~~various~~ aerosol-related studies is the aerosol optical depth (AOD), which is the vertical integral of extinction by aerosol particles through the atmospheric column. Over the last several decades, AOD remote sensing ~~of AOD from space~~ has been performed from space using a wide variety of sensors having different characteristics: passive and active, ultraviolet (UV) to thermal infrared (TIR) spectral regions, single-view to multi-view, single-pixel to broad swath, sub-km to tens of km resolution, intensity-only and polarimetric, different orbits and observation time(s). Table 1 lists the data sets used in the current~~is~~ study, together with key references. Except for the Earth Polychromatic Imaging Camera ~~EPIC~~ (EPIC, orbiting at L1 Lagrange point directly between Earth and the sun on the DSCOVR ~~in L1 orbit~~satellite), all ~~other~~ sensors are in polar-orbiting sun-synchronous low-earth orbits (~600-800 km). Only a few of these sensors were optimized for accurate ~~retrievals of~~ aerosols properties retrievals, and for many, AOD at one or more visible wavelengths is the only quantitatively-reliable aerosol parameter they provide. Thus, we expect significant differences in AOD products retrieved from those sensors. Table 1 is not exhaustive for available AOD ~~; other~~products. Other ~~platforms~~ AOD products ~~include~~ such from active sensors such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) and imaging radiometers on geostationary satellites are not considered here~~. These~~, as they have very different sampling characteristics (~~i.e~~e.g.,~~.~~ CALIOP profiles a 60 km wide ~~line~~swath, with areas either viewed twice daily and twice during the night during a month, or not at all; geostationary sensors sample a constant disk, typical ~~with~~ at a ~~typical~~ frequency of 10 minutes to 1 hour); ~~this means~~thus their monthly mean products are conceptually very different from polar-orbiters.~~, so they are not considered here.~~

Differences between products exist whether using the same algorithm on multiple sensors or multiple retrieval algorithms on the same sensor. Examples of using the same basic principles, to several instruments having similar but not identical characteristics (Sayer et al., 2017, 2019; Li et al., 2016b, Levy et al., 2015). Even between "identical sensors", such as Moderate resolution Imaging Spectrometer (MODIS) on Terra and Aqua, ~~Even with one algorithm, using common basic principles applied to several instruments with similar but not identical characteristics, differences between products exist (Sayer et al., 2017, 2019; Li et al., 2016b). Further, even between the MODIS Terra and Aqua sensors, essentially identical instruments to which the same AOD retrieval algorithm is applied, there remain~~ differences remain due to calibration and time-of-day differences between the sensors (Sayer et al., 2015; Levy et al., 2018). Using different retrieval algorithms ~~between~~ for products retrieved from the same instruments introduces additional discrepancies (Kokhanovsky and de Leeuw, 2009; Kinne, 2009). ~~Likewise, different algorithms applied to the same data set, such as~~As an example, the three algorithms applied to AATSR, provide similar but slightly different results (de Leeuw et al., 2015; Popp et al., 2016). Retrieval assumptions ~~may~~ might work well in certain conditions, but not globally. Thus, there are regional differences in ~~the consistency between~~ AOD products ~~exist~~ (Li et al., 2014b).

An important factor ~~behind the~~contributing to differences ~~could be~~is related to the ~~strictness of~~approach to cloud masking, which affects the ~~ing which~~ pixels ~~are~~ selected for ~~processed~~ processing by retrieval algorithms~~,~~ and ~~propagating~~ propagates into different~~ing~~ levels of clear-sky bias in daily and monthly aggregates (Sogacheva et al., 2017; Zhao et al., 2013; Li et al., 2009). Escribano et al. (2017) estimated the impact of choosing different AOD products for a dust emission inversion scheme and concluded that the large spread in aerosol emission flux over the Sahara and Arabian Peninsula is likely associated with differences between satellite datasets. Similarly, Li et al. (2009) concluded that differences in cloud-masking alone could account for most differences among multiple satellite AOD datasets, including several for which different algorithms were applied to data from the same instrument. Due to these discrepancies, ~~none of the~~no two satellite AOD products give~~s~~ identical values of aerosol properties, and none ~~or~~ is uniformly most accurate (de Leeuw et al., 2015, 2018; Kinne et al., 2006).~~ In other words, there is no single "best" AOD satellite dataset globally. However, d~~ Different techniques have been applied to reveal the spatial and temporal differences between AOD monthly products, e.g., principal component analysis (Li et al., 2013; Li et al., 2014b) or maximum covariance analysis (Li et al., 2014a, b). They~~,~~ show that there are key similarities among the AOD products tested~~, including MODIS, MISR and SeaWiFS~~.

There is no single “best” AOD satellite product globally. Merging multi-sensor AOD products holds the potential to produce a more spatially and temporally complete and accurate AOD picture. With multiple observational datasets available, it is important to examine their consistency in representing ~~the~~ aerosol property variability in these dimensions.~~,~~ ~~which~~ This is useful for constraining aerosol parametrizations in climate models (Liu et al., 2006), in the study of aerosol climate effects (Chylek et al., 2003; Bellouin et al., 2005), and for verifying global climate models (e.g., Kinne et al., 2003; 2006; Ban-Weiss et al., 2014)~~,~~ where satellite-retrieved AOD monthly aggregates are used.

However, to integrate a collection of several satellite aerosol products into a coherent and consistent climatology is a difficult task (Mishchenko et al., 2007; Li et al., 2009). There are only a few studies~~,~~ where an AOD record was merged from different satellites. Chatterjee et al. (2010) describe a geostatistical data fusion technique that can take advantage of the spatial autocorrelation of AOD distributions retrieved from the Multi-Single imaging Spectroradiometer (MISR) and MODIS, while making optimal use of all available data sets. Tang et al. (2016) performed a spatio-temporal fusion of satellite AOD products from MODIS and Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) using a Bayesian Maximum Entropy method for East Asia and showed that, in the regions where both MODIS and SeaWiFS have valid observations, the accuracy of the merged AOD is higher than those of the MODIS and SeaWiFS AODs ~~alone~~individually. Han et al. (2017) improved the AOD retrieval accuracy by fusing~~on of~~ MODIS and CALIOP data. Sogacheva et al. (2018b) combined ATSR and MODIS AOD to study the trends in AOD over China ~~during the period~~between 1995 and ~~-~~2017.

Naeger et al. (2016) combined daily AOD products from polar-orbiting and geostationary satellites to generate a near-real-time (NRT) daily AOD composite product for a case study of trans-Pacific transport of Asian pollution and dust aerosols in mid-March 2014. Li et al. (2016a) constructed a monthly mean AOD ensemble by combining monthly AOD anomaly time series from five widely used satellite products (MODIS, MISR, SeaWiFS, Ozone Monitoring Instrument (OMI) and POLarization and Directionality of the Earth's Reflectances (POLDER)) and applying an Ensemble Kalman Filter technique to these multi-sensor and ground-based aerosol observations to reduce uncertainties. Penning de Vries et al.~~,~~ (2015) examined relationships between monthly mean ~~aerosol properties (~~AOD~~, and extinction~~ Ångström exponent (AE) from MODIS, UV Aerosol Index from the Global Ozone Monitoring Experiment–2 (~~-~~GOME-2)~~,)~~ and trace gas column densities and showed the advantage of using multiple datasets with respect to ~~the~~ characterizing ~~ation of the~~ aerosol type. Boys et al. (2013) combined SeaWiFS and MISR AOD~~ data~~ with the GEOS-Chem global model to create and study trends in a 15-year time series of surface particulate matter levels.

When merging datasets, clearly identifying the limitations of each one ~~should~~ must be considered. Taking advantage of the strengths of single sensors when merging AOD products, derived from different satellite instruments, could help move toward the goal of a long-term, consistent, community AOD record. On the other hand, the spread of satellite AOD records also ~~contains added~~ adds value for constraining the uncertainty of the satellite knowledge. ~~While~~ Whereas a lack of diversity among data sets does not mean that they have converged on the true value e.g., AErosol RObotic NETwork (AERONET, Holben et al., 1998) AOD, which is a recognized standard for instantaneous AOD reference, the existence of unexplained diversity does imply they have not. Note that, as with all measurements, even the AERONET spectral AOD~~ measurements~~, which we adopt as the evaluation standard here, have limitations. For example, AERONET includes ~450 active stations in 2019, offering far more spatial coverage than in 1993~~,~~ when the network was founded, yet even now~~,~~ AERONET spatial sampling is very limited for the current application, especially in regions where aerosol gradients are large, e.g., near sources (e.g., Li et al., 2016a).

To assess their consistency, the products should be compared during overlapping periods, because interannual and shorter-term variability of atmospheric aerosols can be significant in some parts of the world (e.g., Lee et al. 2018). In the current study, AOD monthly aggregates from 16 different satellite products were evaluated with ground-based measurements such as those from the ~~Aerosol Robotic Network (~~AERONET~~, Holben et al., 1998)~~. Based on the comparison with AERONET, we estimate how well the satellite AOD monthly aggregates reproduce the AERONET AOD climatology. To reveal the spread among AOD products globally, we considered areas with different aerosol types, aerosol loading and surface types, which are the dominant factors affecting AOD product quality. Considering different regions globally, we also ~~identify~~ identified the strengths and ~~weaknesses~~ limitations of the aggregate dataset in capturing different aerosol conditions, and the performance of the individual aerosol retrieval algorithms over different surface types. This allows users to choose the AOD product of better quality, depending on the area and research objective. A verification of open-ocean monthly data using the Maritime Aerosol Network (MAN, Smirnov et al., 2009) is not possible in this way, because MAN data are acquired during cruises on ships of opportunity rather than as regular, repeating observations at specific locations.

Different ~~methods~~ approaches for merging the AOD products (~~mean,~~ median,~~ shifted,~~ weighted according to the evaluation results) are introduced in the current paper. AOD evaluation results are used to merge the L3 gridded monthly AOD data globally and AOD time series for the period 1995-2017. The AOD merged ~~datasets~~ products are inter-compared and evaluated against AERONET globally and regionally~~, and for each AOD product separately~~. Annual, seasonal and monthly regional time series obtained with different merging methods are also inter-compared.

This study grew out of discussions at annual AeroSat (https://aerosat.org, last accessed 09.05.2019) meetings about how to move forward on the difficult topic of combining distinct aerosol data records. AeroSat is a grass-roots group of several dozen algorithm developer ~~groups~~ teams and data users.~~,~~ M~~m~~eeting in person around once a year in concert with its sibling AeroCom group of aerosol modelers (https://aerocom.mpimet.mpg.de/~~index.php?id=2404~~, last accessed 09.05.2019) allows active discussion between data providers and data users~~,~~ to highlight developments,~~ and~~ discuss current issues and open questions in the field of satellite aerosol remote sensing and aerosol modelling.

The paper is organized as follows. ~~The main properties of the instruments and basic principles of the AOD retrieval algorithms are summarized in Sect. 2.~~ In Section ~~3~~2, the AOD ~~datasets~~ products and regions of interest are introduced. Main principles and results for the statistical evaluation of individual monthly AOD retrievals are presented in Sect. 3~~ 4 (and detailed results are contained in the Annex)~~. Alternative methods for merging are discussed in Sect. ~~5~~4. AOD merged products are introduced, evaluated and inter-compared with individual products in Sect. 5. A~~and a~~s a result, annual, seasonal and monthly regional AOD time series are presented and discussed in Sect. 6~~ and 7~~. A brief summary and conclusion are given in the final section.

2 Regions of interest, ~~I~~instruments~~-~~ and ~~algorithms/~~AOD products

2.1 Regions of interest

There are huge regional differences in AOD loading, types (composition, optical properties), seasonality, and surface reflectance (Holben et al., 2001; Dubovik et al., 2002; Pinty et al., 2011). Retrieval quality (accuracy, precision, coverage) varies considerably as a function of these conditions, as well as whether a retrieval is over land or ocean. Therefore, this study focuses on surface-specific (land or ocean) and regional evaluation of these diverse aerosol products.

In addition to evaluating AOD products AOD over land, ocean and globally (note, that not all sensor/algorithm combinations retrieve over both surfaces), we chose 15 regions that seem likely to represent a sufficient variety of aerosol and surface conditions (Fig. 1, Table S1). These include 11 land regions, 2 ocean regions, and one heavily mixed region. The land regions represent Europe (denoted by Eur), Boreal (Bor), Northern, Eastern and Western Asia (AsN, AsE and AsW, respectively), Australia (Aus), Northern and Southern Africa (AfN and AfS), Southern America (AmS), east and west of Northern America (NAE and NAW). The Atlantic Ocean is represented as two ocean regions, one characterized by Saharan dust outflow over the central Atlantic (AOd) and second that includes burning outflow over the southern Atlantic (AOb). The mixed region over Indonesia (Ind), includes both land and ocean. Due to documented large changes in AOD during the last 25 years (Sogacheva et al., 2018a, 2018b), we also considered the South-eastern China (ChinaSE) subset of the AsE region.

The main body of the manuscript focuses on the big-picture results (global, all-land, all-ocean, and two regions). The two regions, Europe and ChinaSE, were chosen because they are often the focus of aerosol studies. Results from the remaining regions are presented in the supplement.

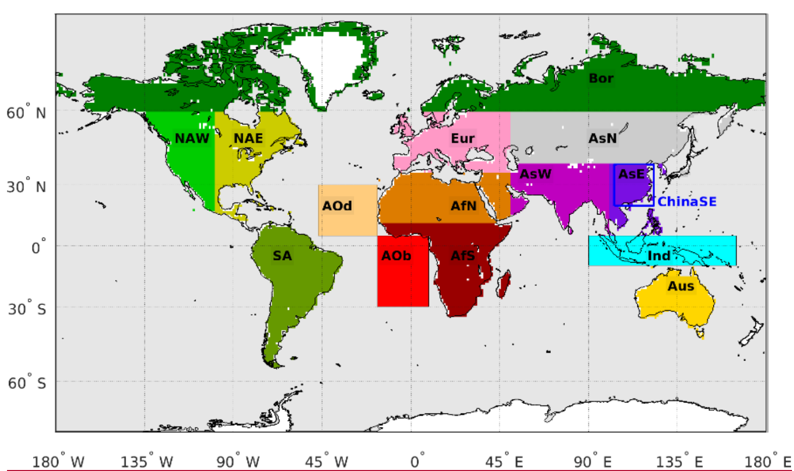


Figure 1: 15 Land and ocean regions defined in this study: Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AfN), southern Africa (AfS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outflow (AOd), Atlantic Ocean biomass burning outflow (AOb). In addition, Southeast China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.

2.2 Instruments, algorithms and AOD products

An overview of the instruments and AOD products included in this study is presented in Table 1. AOD products from the same instruments retrieved with different algorithms are named in the paper with the instrument and retrieval algorithms, e.g., ATSR ADV, ATSR SU, Terra Dark Target (DT) & Deep Blue (DB) and Terra MAIAC. When both Terra and Aqua are considered, we call them together as MODIS DT&DB or MODIS MAIAC. Note that we used the merged MODIS Deep Blue and Dark Target product (denoted "DT&DB"), rather than the results of the individual DB and DT algorithms, as this merged dataset was introduced into the product for similar purposes ~~like~~ as the one explored in this work. An ensemble ATSR product (ATSR_ens) was generated from the three ATSR products (ATRS ADV, ATSR SU, ATSR ORAC) in order to combine the strengths of several algorithms and to increase the coverage of the combined product (Kosmale et al., manuscript in preparation). The ensemble was calculated per pixel as the weighted mean of the individual algorithm values with weights as the inverse of the individual pixel level uncertainty values. The ensemble algorithm required as minimum for each pixel valid results from at least two of the contributing algorithms. The uncertainties of each algorithm were first corrected in their absolute values to agree on average with the mean error.

For some products, AOD data is available for wavelengths other than 0.55 µm. Total Ozone Mapping Spectrometer (TOMS) and OMI products include AOD at 0.50 µm, Advanced Very-High-Resolution Radiometer (AVHRR) NOAA at approximately 0.63 µm (with slight variation between the different AVHRR sensors), and EPIC AOD is available at 0.44 µm (in the dataset used in the current study). Note, if the wavelength is not mentioned specifically, 0.55 µm is implicit.

The official AOD monthly products (typically referred to as Level 3 or L3 data), which correspond to arithmetic means of daily mean data aggregated onto (typically) 1°x1° grid, have been used without further processing. The first exceptions are for AVHRR NOAA and POLarization and Directionality of the Earth's Reflectances (POLDER), which provide very high AOD values poleward of ca. 60° and over Hudson Bay (50°-70°N, 70°-95°E), respectively. The values are unrealistic, a likely a consequence of cloud and/or sea ice contamination. To eliminate those unrealistic values, all their AOD values of >0.7 have been removed. Applying that limit decreased the offset between the AVHRR NOAA product and other products but did not eliminate it (see Sect. S2 for details). Additionally, MISR Standard (0.5°x0.5° resolution) and AVHRR NOAA (0.1°x0.1° resolution) L3 AOD products were aggregated by simple averaging to 1° to match the other datasets.

Note that due to differences in instrument capabilities and swath widths (Table 1), the spatial and temporal data sampling available for calculating monthly averages varies considerably among the satellite products. The ATSRs and MISR have narrow-swaths, and generally provide only a few days with retrievals per month, whereas most of the rest (TOMS, AVHRR, SeaWiFS, MODIS, OMI, POLDER, Visible Infrared Imaging Radiometer Suite (VIIRS)) see the whole planet roughly every day or two, so that their coverage is mostly limited by, e.g., the persistence of cloud cover. As mentioned previously, EPIC is a special case, as it provides moving snapshots of the daylit portion of the Earth, up to several times per day, as distinct from overpasses at only specific local solar Equatorial crossing times for the sensors on polar-orbiting satellites. Further, TOMS and OMI have notably coarser pixels resolution than the others, so their coverage and quality (including potential cloud contamination) is more sensitive to cloud masking decisions. Some datasets provide measures of internal diversity (e.g., standard deviation), but none currently provides estimates of the monthly aggregate uncertainty against some standard, which would be a combination of (both systematic and random) retrieval uncertainties and sampling limitations. This is an area currently being investigated by the AeroSat due to the wide use of L3 products.

For the inter-comparison between AOD products, three "reference" years were chosen:

- 2000, when the AOD products from TOMS, AVHRR NOAA, SeaWiFS, ATSR-2, MODIS Terra and MISR are available (for the full year, except for MISR and MODIS Terra, which were available from March to December);
- 2008, when the AOD products from AATSR, MODIS Terra and Aqua, MISR, AVHRR NOAA, AVHRR DB/SOAR, SeaWiFS and POLDER are available;
- 2017, when the AOD products from MODIS Terra and Aqua, MISR, VIIRS and EPIC are available;

For products with incomplete or no coverage over ocean (TOMS, OMI, and MAIAC-types products (Terra MAIAC, Aqua MAIAC, EPIC)), the AOD over land only product was considered.

3 AOD data and regions of interest

3.1 Monthly AOD aggregates

AOD L3 (1°x1° resolution) monthly global products were utilized in the current study. The MISR Standard L3 AOD product (0.5°x0.5° resolution) was aggregated to 1° to match the other datasets by simple averaging. Note that because of differences in instrument capabilities and swath widths, the spatial and temporal coverage of the data for calculating the monthly average are quite different among the satellite products. Some datasets provide measures of internal diversity (e.g., standard deviation), but none currently provides estimates of the monthly aggregate uncertainty against some standard. This is an area currently being investigated by the AeroSat group due to the wide use of the L3 products.

Table 1. Overview of the sensors, data records and AOD algorithms discussed in this paper. For the products availability, see Table 4.

Sensor(s)	Coverage, ~~Resolution~~L3 grid size	Algorithm Version	Algorithm Principles	References
Total Ozone Mapping Spectrometer (**TOMS**) (*UV spectrometer*)	~~1978~~1979–1993 1996–~~2003~~ 2001, 3100 km swath, 1° ~~× 1°~~ daily and ~~/~~monthly	Nimbus-7/TOMS: N7AERUV Ver. 0.4.3 EP/TOMS: EPAERUV Ver. 0.1.3.	Enhanced sensitivity of TOA spectral reflectance in the UV to aerosol extinction and absorption	Torres et al. (1998, 2005)
Advanced Very High Resolution Radiometer (**AVHRR**)	1981-2017, 2900 km swath, 0.1°, daily and monthly	AVHRR NOAA	Single-channel retrieval of aerosol optical depth; over ocean only	Ignatov and Stowe (2002) Heidinger et al. (2002) Zhao et al. (2008)
~~Advanced Very High Resolution Radiometer (**AVHRR**)~~ (*Bispectral, single-view, broad-swath radiometer*)	1989-1991 (NOAA7), 1995-1999 (NOAA14), 2006-2011 (NOAA18), ~~8.8 km,~~ 0.5° and 1°, daily and monthly	D~~D~~eep Blue/SOAR, V. 4	Land: surface modeled using data base or NDVI. ~~Water~~Ocean: bispectral simultaneous retrieval	Hsu et al. (2017) Sayer et al. (2017)
Along-Track Scanning Radiometer (**ATSR-2**) and Advanced ATSR (**AATSR**), both called as **ATSR**	1995–2003 (ATSR-2) 2002–2012 (AATSR), 512 km swath, ~~global coverage in ~ 6 days,10 km daily~~ 1° daily and monthly	ADV/ASV V2.31	Land: spectral constant reflectance ratio Ocean: modelled reflectance	Flowerdew and Haigh (1995) Veefkind et al. (1998) Kolmonen et al. (2016) Sogacheva et al. (2017)
(*dual view radiometer in the visible and near-infrared; thermal infrared for cloud*)		SU V4.3	Iterative model inversion for continuous retrieval of AOD and FMF. ~~-~~ Land: retrieval of BRDF parameters. Ocean: prior reflectance model.	North et al. (1999) North 2002 Bevan et al. (2012)
		ORAC V4.01 (in current paper, ~~not as separate product but~~ as a part of the ATSR ensemble)	Optimal estimation Land: SU surface parametrization Ocean: sea surface reflectance model	Thomas et al. (2009) Sayer et al. (2010)
		ATSR ensemble V. 2.7	Uncertainty weighted mean of ATSR2/AATSR baseline algorithms ADV, ORAC and SU	Kosmale et al., in prep.
Sea-viewing Wide Field-of-view Sensor (**SeaWiFS**)	1997-2010	Deep Blue/SOAR V.1	Land: surface modeled using data base or NDVI.	Sayer et al. (2012a, b) Hsu et al. (2004, 2013)

(*Multispectral, single-view, broad-swath radiometer*)	1502 km swath~~, 13.5 km,~~ 0.5° and 1° daily and monthly		~~Water~~Ocean: multispectral simultaneous retrieval	
Multiangle Imaging SpectoRadiometer (**MISR**) (*Multispectral (4-band; Vis-NIR), multiangle (9-angle) radiometer*)	2000-present, 380 km swath~~,; global coverage ~ once /week; 4.4 km~~ 0.5° daily and monthly	Standard ~~Product~~ Algorithm (SA) V23	Land: surface contribution estimated by empirical orthogonal functions ~~(EOFs)~~ and assumption of spectral shape invariance ~~Water~~Ocean: Two-band (red, NIR) retrieval ~~with known surface wind speed~~ using cameras not affected by sun glint Both: Lookup table with 74 mixtures of 8 different particle distributions	Martonchik et al. (2009) Garay et al. (2017) Witek et al. (2018) Kahn et al. (2010) Garay et al. (2019)
Moderate Resolution Imaging Spectroradiometer (**MODIS**) Terra and Aqua (*Multispectral, single-view, broad-swath radiometer*)	Terra: 2000-present Aqua: 2002-present 2300 km swath ~~global coverage ~ two days, 10 km and,~~ 1°, daily, 8-day, and monthly	DT&DB C6.1	DT: ~~Lookup table approach, s~~Surface is "known" function of wind speed (ocean), or parameterized spectral relationship ~~at different wavelengths~~ (land – vegetation/dark soil) DB: ~~Lookup table approach, climatology~~ Database and spectral relations of surface reflectance	DT: Levy et al. (2013, 2018) Gupta et al. (2016) DB: Hsu et al. (2013, 2019) DT&DB: Levy et al. (2013) Sayer et al. (2014)
		MAIAC V6	Simultaneous retrieval of surface and aerosol from time series of observations	Lyapustin et al. (2018)
Ozone Monitoring Instrument (**OMI**) (*UV spectrometer*)	2004– 2016, 2600 km swath, 1° daily, monthly	OMAERUV V.1.8.9.1	Enhanced sensitivity of TOA spectral reflectance to UV aerosol extinction and absorption	Jethva and Torres (2011), Torres et al. (2007, 2013, 201
Polarization and Directionality of the Earth's Reflectances (**POLDER**) 3 (*Multispectral, multiangle polarimeter*)	Dec 2004 – Dec 2013 ~~Global coverage in ~two days, swath~~ 2100 ×1600 km swath, ~~5.3 ×6.2 km at nadir, and~~ 1° daily, monthly, seasonally	GRASP V.1	Simultaneous retrieval of surface and aerosol in frame of multi-pixel approach: statistically optimized fitting of large of pixels groups (aggregated in time and space), the aerosol is assumed as an external mixture of several predefined aerosol components~~groups of pixels~~	Dubovik et al. (2011, 2014, 20
Visible Infrared Imaging Radiometer Suite (**VIIRS**) (*Multispectral, single-view, broad-swath radiometer*)	2012-present, ~~-~~3040 km swath,~~6 km and~~ 1°, daily, and monthly	Deep Blue/SOAR, V.1	Land: surface modeled using data base or spectral relationship. ~~Water~~Ocean: multispectral simultaneous retrieval	Sayer et al (2018a, b, 2019) Hsu et al (2019)

Earth Polychromatic Imaging Camera (**EPIC**) (*Multispectral radiometer orbiting at Lagrange point*)	2015-2016, 1°, daily, and monthly~~10 km~~	MAIAC V1	Simultaneous retrieval of surface and aerosol from time series of observations	Huang et al. (2019)

3 AOD products inter-comparison and evaluation with AERONET

The AOD deviations of the individual products from the median AOD (Figs. S1 and S2) are discussed in detail in the Supplement (Sect. S2). The AOD deviations from the median show regional differences, even for products retrieved from the same instruments with similar algorithm. As both negative and positive deviations are observed in regions with high AOD, the surface type is also likely to influence the AOD retrieval. High AOD might, in turn, be wrongly screened as cloud and thus the resulting lack of high AOD retrieval leads to a low bias in monthly AOD.

To further reveal differences among the AOD products retrieved with different algorithms and applied to different satellites, the diversity of the satellite annual mean AOD for years 2000, 2008 and 2017 was calculated and discussed in Sect. S3 (Figs. S3 and S4). The diversity is lower in 2017, when only MODIS-family and VIIRS AOD products are available.

3 AOD data and regions of interest

3.1 Monthly AOD aggregates

AOD L3 (1°x1° resolution) monthly global products were utilized in the current study. The MISR Standard L3 AOD product (0.5°x0.5° resolution) was aggregated to 1° to match the other datasets by simple averaging. Note that because of differences in instrument capabilities and swath widths, the spatial and temporal coverage of the data for calculating the monthly average are quite different among the satellite products. Some datasets provide measures of internal diversity (e.g., standard deviation), but none currently provides estimates of the monthly aggregate uncertainty against some standard. This is an area currently being investigated by the AeroSat group due to the wide use of the L3 products.For the inter-comparison between AOD products, three "reference" years were chosen:

- 2000, when the AOD products from TOMS, AVHRR, SeaWiFS, ATSR2, MODIS Terra and MISR are available (for the full year, except for MISR and MODIS Terra, which were available from March to December);
- 2008, when the AOD products from AATSR, MODIS Terra and Aqua, MISR, AVHRR, SeaWiFS and POLDER are available;
- 2017, when the AOD products from MODIS Terra and Aqua, MISR, VIIRS and EPIC are available;

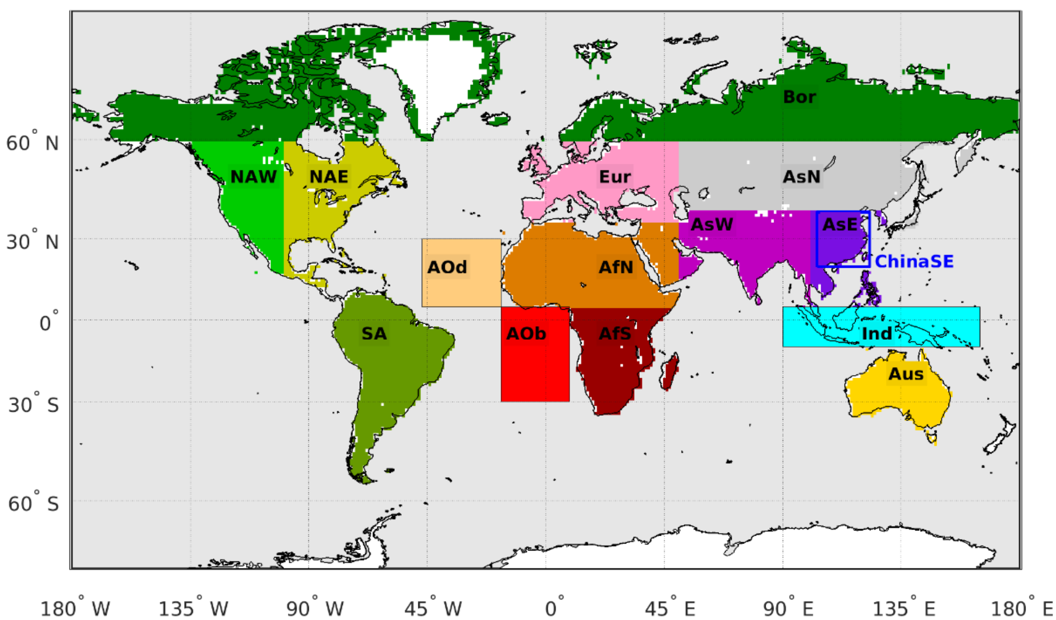
For products with incomplete or no coverage over ocean (TOMS, OMI, and MAIAC-types products (Terra MAIAC, Aqua MAIAC, EPIC)), the AOD over land product was considered only.

3.2 Regions of interest

This study focuses on different regions across the globe, as regional differences in AOD loading, types, seasonality, surface reflectance exist (Holben et al., 2001; Dubovik et al., 2002; Pinty et al., 2011), which can affect the retrieval regional quality considerably. As such, applications drawn from the products will be analysed on a regional level.

~~To study regional differences over the globe, we chose 15 regions that seem likely to represent different aerosol/surface conditions (Fig. 1). There are 11 land regions: Europe (Eur), Boreal (Bor), Northern, Eastern and Western Asia (AsN, AsE and AsW, respectively), Australia (Aus), Northern and Southern Africa (AfN and AfS), Southern America (AmS), east and west of Northern America (NAE and NAW), two regions over ocean: Saharan dust outbreak over the central Atlantic (AOd) and possible biomass burning outbreak over southern Atlantic (AOb), and one region, Indonesia (Ind), that includes both land and ocean. Furthermore, we studied AOD over all land, all ocean and globally, when observed. South-eastern China (ChinaSE), which is part of the AsE region, was also considered separately as an area with considerable AOD changes during the last 25 years (Sogacheva et al., 2018a, 2018b). Altogether, we consider the AOD in 18 regions.~~

~~For six regions, Eur, ChinaSE, AfN, AOd, Ind, AOb that differ considerably in AOD loading/type and surface characteristics, as well as for land and ocean globally, the results are presented and discussed in the main part of the manuscript. The results for the other 10 regions are shown in the Annex and briefly discussed there.~~



~~**Figure 1: 15 Land and ocean regions defined in this study: Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AfN), southern Africa (AfS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outbreak (AOd), Atlantic Ocean biomass burning outbreak (AOb). In addition, Southeast China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.**~~

4 AOD products inter-comparison and evaluation with AERONET

4.1 AOD spatial distribution

To reveal spatial differences among the annual AOD products, annual offsets were estimated from the median AODs of all available products (Fig. 2, upper panel). Year 2008 was chosen for this exercise, as AOD data are available from most instruments (Sect. 3.1). For TOMS, which is not available in 2008, the annual AOD difference from the median AOD was calculated for year 2000; for VIIRS and EPIC 2017 was chosen for the inter-comparison. Median AOD is slightly higher in 2008 over both land (0.184) and ocean (0.135), and thus globally (0.150), compared with 2000 and 2017 over land (0.180/0.174), ocean (0.128/0.132) and globally (0.143/0.145), respectively. However, all these differences remain below 0.010. Although AOD is higher over ChinaSE in 2008 compared to other years, similar AOD spatial patterns are observed for the 3 years chosen for the inter-comparison. Thus, using those 3 years (2000, 2008, 2017) to compare global, annual-averaged AOD deviation from the median AOD, is suitable for revealing differences among the products.

Over land, TOMS AOD is about twice as high (by 0.162) as the median land AOD for all available products (0.180); OMI is higher by 0.06. One explanation for such high AOD differences for TOMS and OMI is possible cloud contamination, related to the lower pixel resolution (~ 50 x 50 km-square) compared to the other instruments, (Sect 2). Global AOD is also higher for POLDER by 0.049, over both land (by 0.096) and ocean (by 0.029), except for the dust area AfN, were AOD is lower than the median.

Products from different instruments, retrieved with similar algorithms (AVHRR, SeaWiFS, VIIRS), often show similar results. AVHRR and SeaWiFS tend to underestimate AOD compared to the median, whereas VIIRS global AOD provides only a small overestimation (by 0.006) of the median value. MODIS DT&DB show higher AOD compared to the median; the deviation of Terra DT&DB from the median (by 0.027 and 0.031 for land and ocean, respectively) is about twice as high as that for Aqua DT&DB (by 0.013 for both land and ocean). This offset between the two MODIS sensors has been reported previously; although time-of-day differences might contribute, it is thought to be dominated by calibration offsets, and has improved compared to previous data versions (Levy et al., 2018; Sayer et al., 2019).

Interestingly, different algorithms applied to the same instrument, often show opposite results, providing an algorithmic spread from the median value. Unlike MODIS DT&DB, MODIS MAIAC AOD are both lower than the median over land (by 0.038 and 0.044, respectively). ATSR ADV is slightly higher over ocean (by 0.026), whereas ATSR SU is lower than the ocean median by a similar amount (0.024). Over land, ATSR ADV is lower, especially over bright surfaces (Sect. 2.6.1), if retrieved, whereas ATSR SU is considerably higher over bright surfaces, compared to the median value. As a result, ATSR ADV AOD underestimates (by 0.011) and ATSR SU overestimates (by 0.026) the median over land.

Although the global land and ocean AOD values are similar (within ± 0.03 of the median AOD for most products, which is inside the GCOS requirement of the greater of 0.03 or 10%; GCOS, 2016), contrasting AOD behavior appears in several regions. In ChinaSE, OMI, AVHRR, SeaWiFS, VIIRS and MISR AOD are considerably lower (up to -0.25) than the median AOD, whereas the MODIS-family (Terra DT&DB, Aqua DT&DB, Terra MAIAC, Aqua MAIAC) retrieve up to 0.2 higher

AOD than the median. Over bright surfaces (e.g., AfN, AsW), the OMI, AVHRR, VIIRS, ATSR SU, ATSR ensemble retrieves up to 0.2-0.25 higher AOD than the median, whereas the MAIAC-family (Terra MAIAC, Aqua MAIAC, EPIC) and POLDER AOD are lower by up to 0.1-0.15 AOD than the median. Over open ocean, the VIIRS and ATSR ensemble show AOD closest to median among the products. As there is no clear deviation from the median for almost all products over open ocean in the continental outflow areas (e.g. AOd and AOb), this confirms that the datasets contain similar AOD in these regions. The phenomenon of high Southern Ocean AOD is found in several satellite data sets, including MODIS Terra and Aqua DT, MISR and POLDER and may be related to unresolved clouds (e.g., Toth et al., 2013; Witek et al., 2018).

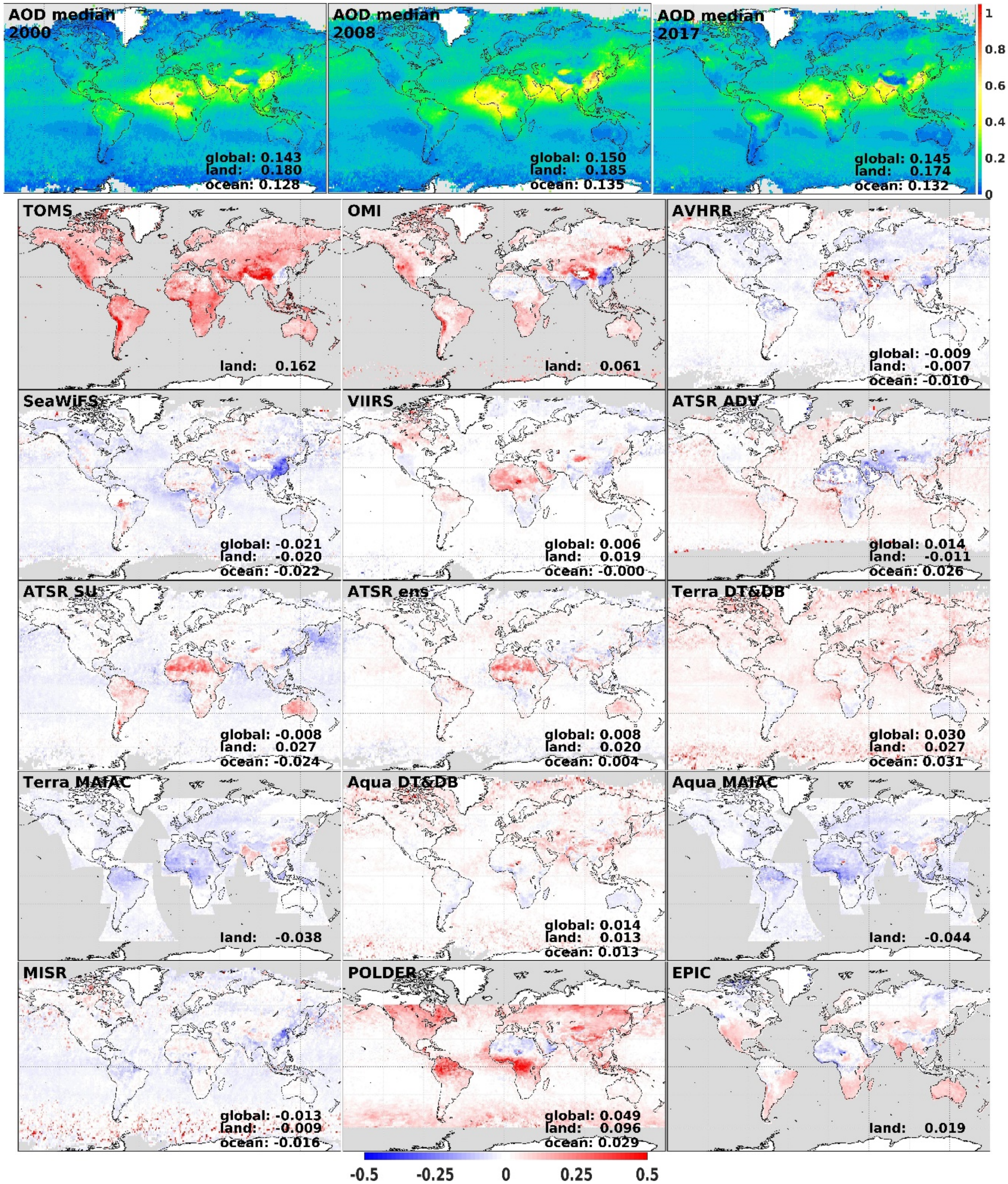


Figure 2. Upper line: annual AOD median for 2000 (*), 2008 and 2017 (X), calculated from the available products. AOD anomalies with respect to the AOD median are shown on the deviation plots. Lines 2-6: AOD deviation of the different products from the annual median AOD for years 2000 (TOMS), 2017 (VIIRS and EPIC) or 2008 (all other products). AOD anomalies with respect to the AOD median are shown on the deviation plots. Global land and ocean AOD mean differences are shown for each product, when available. For summer, see Fig. A1.

Seasonal deviations from the median AOD are similar to the annual patterns throughout the year. However, the spread among the products is slightly more pronounced in summer (Fig. A1 for JJA, summer for the Northern Hemisphere), when the absolute AOD often reaches its maximum in certain regions (e.g., in China, Sogacheva et al., 2018).

Thus, the AOD deviations from the median show regional differences, even for products retrieved from the same instruments with similar algorithms. As both negative and positive deviations are observed in regions with high AOD loading, the surface type is also likely to influence the AOD retrieval. High AOD loading might, in turn, be wrongly screened as cloud and thus bias monthly AOD lower. As with many factors that contribute significantly to AOD retrieval results, surface treatment and cloud screening should be tested with L2 (higher-resolution, swath-based resolution) daily data; but this is beyond the scope of the current study, which considers only L3 monthly AOD products.

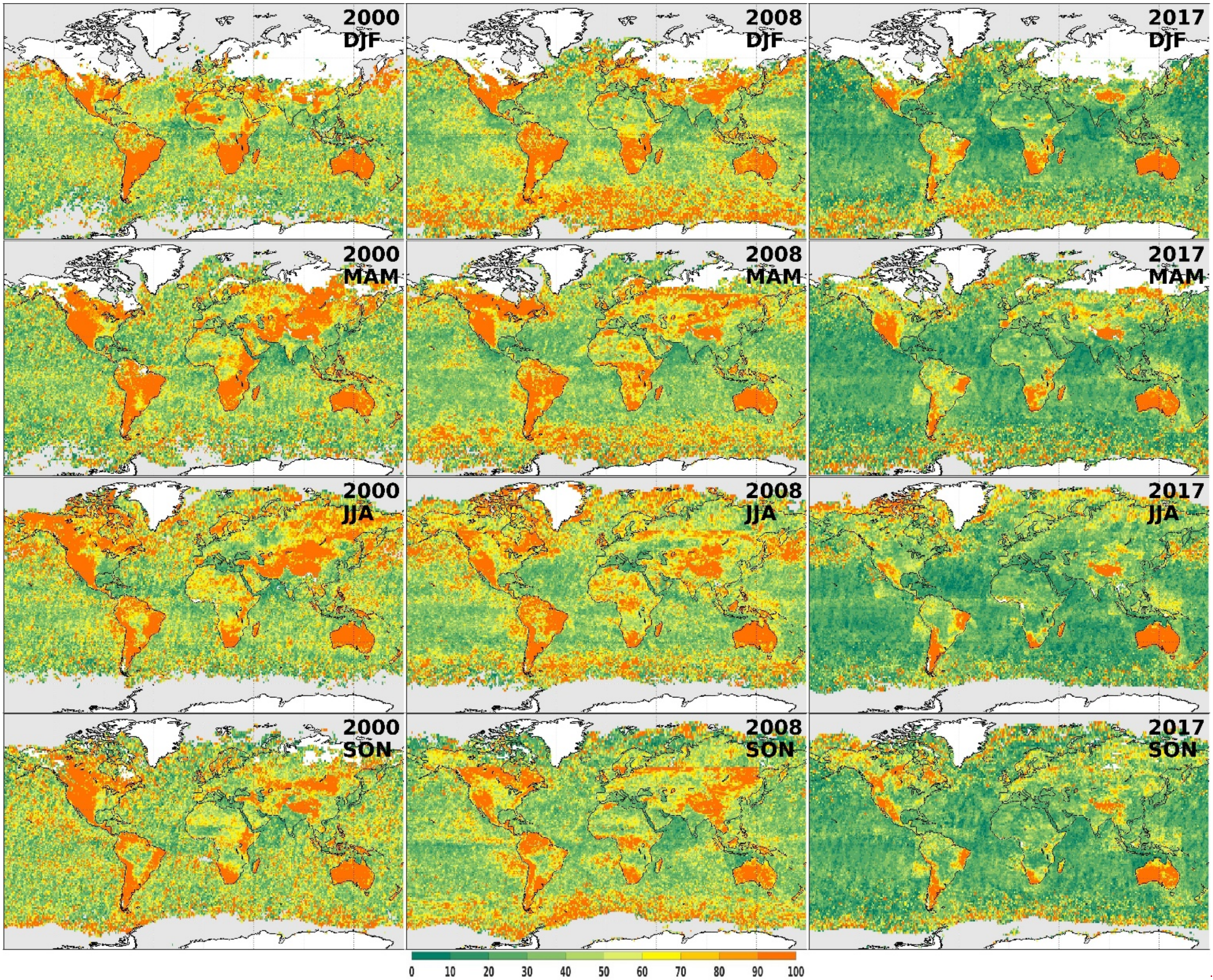
To further reveal differences among the AOD products retrieved with different algorithms applied to different satellites, the diversity of the satellite annual mean AOD ($AODdiv$) was calculated, as in Chin et al., (2014):

$$AODdiv = 0.5 * 100\%$$

The seasonal AOD diversity was calculated for the test years 2000, 2008 and 2017 from all available products (Fig. 3). As expected, satellite data agree best over ocean, and diversity decreases from 2000 (40-60%) towards 2017 (20-30%). Higher over-water diversity is often seen in high latitude oceans, for which retrievals are more challenging due to cloud cover and the high solar zenith angle, which leads to both more limited sampling and more retrieval artefacts (Toth et al., 2013). Note that the results slightly differ from Chin et al., (2014), since in our analysis, besides TOMS, AVHRR and NASA's Earth Observing System (EOS) satellites (SeaWiFS, MISR and MODIS), the European Space Agencies' AATSR radiometer on board the Environmental Satellite (ENVISAT) is also included for years 2000 and 2008. Also, in some cases, newer versions of the satellite products than those available in Chin et al. (2014) were used here.

The diversity is considerably higher over land, reaching more than 90% over certain areas in year 2000 (NAW, SA, Siberia, central Asia, AfS and Aus). AOD diversity decreases considerably in 2017 (to 30-40% on average), when AOD from only NASA Earth Observation System (EOS) satellites (MISR, MODIS, VIIRS) is currently available, and the size of areas with high diversity is lower. Australia stands out as the region where AOD averages disagree most, showing >90% AOD diversity irrespective of the year or number of satellites/products available. The diversity is somewhat higher in summer, which is also clear from the comparison between annual (Fig. 2) and summer (Fig. A1) AOD deviation from the median value.

As discussed early in this section, AOD for TOMS, OMI and POLDER differ most from the median AOD. Those products contribute most (up to the 40-50%) to the AOD spread (figures not included). However, we use all available products in the current exercise to obtain the longest possible data record to be used later for trend assessment.



~~**Figure 3. Seasonal mean AOD diversity for years 2000, 2008 and 2017 for all available products**~~

~~4.2~~3.1 Evaluation of monthly AOD

To evaluate the quality of any AOD product, the verification of the product against more accurate reference measurements, where possible, is obligatory. Ground-based measurements such as from AERONET (cloud screened and quality assured Version 3 Level 2.0, Giles et al., 2019) provide highly accurate measures of AOD that are widely used as ground truth for the validation of satellite AOD data. Extensive L2 AOD validation has been performed for different aerosol products.

However, climate model evaluation is often performed on monthly scales. Thus, climate analysis begs for evaluation~~Verification~~ of ~~the~~ satellite AOD monthly aggregates ~~is of great importance, since model evaluation is often~~

performed on the monthly temporal resolution (Nabat et al., 2013; Michou et al., 2015; Li et al., 2016b). To evaluate the quality of any AOD product, the verification of the product against more accurate reference measurements is obligatory. Ground-based measurements such as from the AERONET (cloud screened and quality assured Version 3 Level 2.0, Giles et al., 2019) provide highly accurate measures of AOD that are widely used as ground truth for the validation of L2 satellite AOD data. Although extensive L2 AOD validation has been performed for different aerosol products, o Only a few attempts have been made to evaluate AOD monthly aggregates retrieved from satellites (e.g., Li et al., 2014b, Wei et al., 2018). This is because V verification of the L3 monthly aggregate satellite AOD is more challenging than L2 validation. The verification exercise is not a true validation (and note the use of "evaluation" and "verification" here instead of "validation")., AERONET provides AOD at a single point and is not necessarily representative of AOD in a 1°x1° grid. While AERONET samples during all cloud-free daylight hours, a given polar-orbiting sensor will only report at once/day and at the same time each day (e.g., 13:30 Local Time for sensors in the A-Train). unlike instantaneous matchups with satellite retrievals which are commonly compared with AERONET. The possible spatial representativity issues associated with this latter point are a topic of current investigation (e.g., Li et al., 2016a; Schutgens, 2019). Nevertheless, AERONET's instantaneous AOD uncertainty (around 0.01 in the mid-visible, Eck et al. 1999) is significantly lower than most satellite products and its temporal sampling is much more complete. As such it remains a useful source for evaluating these L3 products, and for this purpose we compare AOD monthly aggregates of all available data from both AERONET and each satellite product. Deviations between monthly aggregatessatellite and AERONET monthly aggregates are expected, based, e.g., on due to differences in satellite spatial and temporal sampling (Sec. 2.2, Table 1 4.1). This issue is more significant for satellites with lower coverage and can result in missing extreme AOD events. Differences in cloud screening affect mainly high AOD events that can be erroneously removed in some products. Both coverage and cloud-screening issues can decrease monthly aggregated AOD.

By performing the verification of the AOD monthly aggregates, we aim to reveal how well different monthly satellite products meet monthly AOD values observed with AERONET. For that purpose, we compare AOD monthly aggregates from all available data from both AERONET and each satellite product.

Results from this comparison have limitations, . since As mentioned previously, AERONET provides the data over sertain certain locations within a grid cell, while whereas satellites cover a larger fraction of the area of a grid cell (depending on sampling and cloud cover). So, for example, if AERONET is likely to miss extreme high values (localized plumes close to AERONET station), that will result in AERONET showing lower AOD than satellite. which can produce low biases, unless Conversely, if a station happens to be directly under an aerosol plume—, in which case -and the satellite algorithm filters as a cloud, the AERONET value there might be skewed highwould be higher. A related issue for comparing satellite-retrieved monthly AOD aggregates with ground-based AERONET AOD is the spatial representativeness of the AERONET stations for their grid cells, and their chosen regions more generally, which is an ongoing subject of investigation (Shi et al., 2011; Li et al., 2014a, 2016; Virtanen et al., 2018). Note, that both AERONET and satellite monthly AOD aggregates are not "true" monthly AOD values., When we refer to "AOD monthly aggregate" we mean the daytime, cloud-free AOD monthly aggregate from whatever data are available. as AOD is not measured/retrieved under cloudy conditions or across the full diurnal cycle by either technique; here “AOD monthly aggregate” means the daytime, cloud-free AOD monthly aggregate. In addition to which data are aggregated into a monthly mean, how the monthly mean is calculated is also important. AOD distributions on monthly scales are often closer to Lognormal than Normal, as some data sets will be providing poor sampling of skewed distributions, which suggests that the arithmetic monthly mean may not be the most appropriate metric (O’Neill et al., 2000; Sayer and Knobelspiesse, 2019). The discrepancies between different statistics can be exacerbated when a dataset is providing poor sampling of the extreme conditions. Nevertheless, as it is the most widely-used statistic within the community and is the standard output of current L3 products, monthly means are presented in this analysis. The general framework could be applied to other AOD summary statistics (e.g., monthly median or geometric mean, advocated by Sayer and Knobelspiesse 2019) if these L3 outputs become more widely available in the future.

In the evaluation exercise, AERONET monthly mean AOD and Ångström exponent (AE, (which describes how AOD depends on wavelength and is sometimes used, together with AOD, to constrainas a proxy for aerosol type) were calculated from AERONET daily means. AOD Vverification was performed for all available AERONET monthly data, and separately for different aerosol types, which were defined with AOD and AE thresholds. Although these thresholds are subjective, we consider “background aerosol” to be cases where AOD<0.2, “fine-dominated” to be where AOD>0.2 and AE>1, and “coarse-dominated” to be cases where AOD>0.2 and AE<1 : background aerosol (AOD<0.2), fine-dominated (AOD>0.2, AE>1) and coarse-dominated (AOD>0.2, AE<1) aerosol. These thresholds are somewhat subjective, but this simple categorization reflects typical differences between fine-mode dominated and coarse-mode dominated aerosol types (e.g., Eck et al., 1999). This classification has also been used by, e.g., by Sayer et al. (2018b) and Sogacheva et al. (2018a, b). The annual and seasonal maps of prevailing aerosol type for AERONET locations, calculated from the AERONET data available for the period of 1995-2017, are shown in Fig. S5. Such classification differentiates major aerosol scenarios. The biomass burning seasons over Amazon and South Africa are clearly identified by domination of the fine aerosol particles in JJA (June, July, August) and SON (September, October, November), and the Asian dust transport season in MAM (March, April, May) is clearly coarse-dominated.

As the deviation of each satellite product from the median is regionally dependenthas regional components (Figs. 2S1 and S2), we performed the AOD comparison between monthly aggregated products and the AERONET monthly product for each selected study region separately. Even though we tried to choose regions with (somewhat) homogeneous aerosol conditions during a given season, AOD conditions (and thus algorithm performance) might vary somewhat within the region, and AERONET stations, which may represent different aerosol/surface conditions within one study regions, might may have different record lengths. To keep similar weighting for each station in a region, we first calculated statistics for each AERONET station separately, and then calculated the regional median validation statistics from all available stations.

To reveal how retrieval quality depends on AOD loading, offsets between AERONET AOD and satellite product AOD were estimated for binned AERONET AOD, and the number of observations in each AOD bin is reported. Correlation coefficient (R, Pearson), offset (satellite product-AERONET),and root-mean square error (rms), as well as the fraction of falls to the

~~spread envelope (SE) of ± (0.05 + 0.2*AOD)~~ and fraction of points ~~which~~ that fulfil the GCOS (GE) requirements ~~(GE)~~ of 0.03 or 10% of AOD are reported.

Monthly AOD verification results ~~were~~ are used in this ~~manuscript~~ work to estimate weights for each satellite dataset in ~~some~~ one of the merging approaches in Sect. 4.2. Knowledge about how well a satellite AOD record describes global monthly AOD ~~globally~~ is also important for AOD trend estimation, as well as for estimating~~on of~~ the aerosol impact on the global radiation balance and, thus, on climate change.

~~4.2.1~~3.1.1 Binned offset, global evaluation

As an example, AOD-binned evaluation results are shown in Fig. 2. for Terra DT&DB, and in Fig. S6 for all products. A general tendency toward positive satellite-retrieved AOD offsets ~~(Fig. 4)~~ is observed for most products under background conditions. On~~Since, on~~ average, 70-80% of monthly AOD fall into class "background" (AOD<=0.2), so total AOD mean biases are expected to have similar behaviour. TOMS and~~,~~ OMI ~~and POLDER~~ have the highest positive offsets globally, which is in line with the results from the dataset spatial inter-comparison (Sect. ~~4.1~~S2). ~~Notable bias (overestimation) overall land for situations with low AOD and certain underestimations of coarse mode aerosol (desert dust), revealed during the validation of the L2 AOD of the first version of the POLDER global product is being addressed in a new retrieval version.~~ Offsets close to zero for background AOD are observed for the MODIS MAIAC products.

For most products, except MODIS DT&DB, AOD offsets become negative for AOD>0.2 (fine- and coarse-dominated aerosol types) with increasing amplitude (up to 0.2 - 0.5) towards highest AOD values. MODIS DT&DB ~~Terra and Aqua~~ show the lowest offsets for 0.2<AOD<1.~~. Note that the MODIS DT&DB retrieval is not entirely independent of AERONET, as the over-land algorithm uses a region-specific aerosol-type climatology derived from AERONET measurements (Levy et al., 2013).~~ Offsets for VIIRS are close to 0 for AOD<0.5 and reach ca. 30% of AOD at AOD≈1. For the current MISR standard product, AOD is systematically underestimated for AOD >~0.5; this is largely due to treatment of the surface boundary condition at high AOD (Kahn et al., 2010), and is addressed in the research aerosol retrieval algorithm (~~e.g.,~~ Garay et al., 2019; Limbacher and Kahn, 2019). Except for TOMS and Terra MAIAC, offsets are smaller for coarse-dominated AOD.

~~In summary, similar offsets (positive for AOD<0.2 and negative for AOD>0.2) may plausibly indicate systematic overestimation (by 0-0.05 for different products) of AOD<0.2 and underestimation (by more than 0.25) of high AOD (>0.2) for satellite AOD monthly aggregates. AOD products differ in their ability to meet higher AERONET monthly AOD >0.3. AOD products retrieved from satellites with better coverage show better agreement between AOD and AERONET monthly aggregates. Thus, sampling differences (swath and pixel selection) are critical, as expected. However, MODIS DT&DB show slightly better performance than MODIS MAIAC for AOD>0.3, which might result from differences in the retrieval approach and/or cloud screening.~~

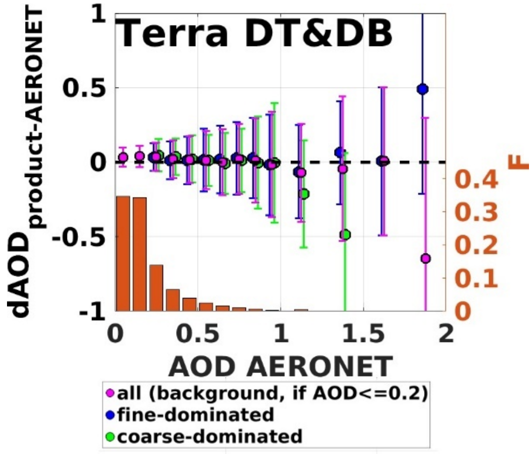


Figure ~~4~~2. Difference between Terra DT&DB and AERONET monthly AOD for selected AOD bins: median bias (circles), bias standard deviation (error bars)~~Global AERONET comparison of the AOD monthly aggregates (circles – median bias, error bar – bias standard deviation)~~ for "all" AOD types (purple), "background" aerosol (purple, AOD<=0.2), "fine-dominated" AOD (blue) and "coarse-dominated" AOD (green); ~~and~~ fraction of points in each bin (bar, orange). For all individual products see Fig. S6.

In summary, positive offsets for AOD<0.2 and negative offsets for AOD>0.2 are observed for most products. AOD products differ more strongly in representing AOD when AERONET monthly AOD is roughly >0.4. AOD products retrieved from satellites having better coverage show better agreement with AERONET monthly aggregates. Thus, sampling differences (swath and pixel selection) are critical in evaluation of monthly products, as expected, but not only factor influencing the evaluation results. MODIS DT&DB shows slightly better performance than MODIS MAIAC for AOD>0.3, which results from differences in the retrieval approach and/or cloud screening.

~~4.2.2~~3.1.2 AOD evaluation over selected regions

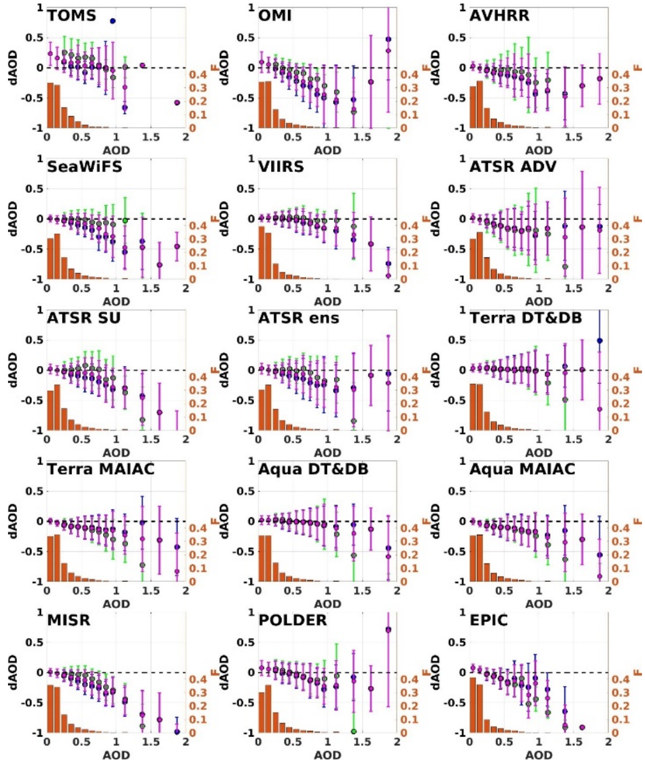
~~Because of~~Due to differences in instrument specifications and retrieval approaches, the performance of retrieval algorithms depends largely on aerosol type, aerosol loading and surface properties at certain locations (e.g., Sayer et al., 2014). In this section we show the evaluation results fore AOD products ~~by comparing statistics between each product and AERONET~~ in ~~four~~ two selected regions: ~~Eur, China SE, AfN, AOd~~Europe and ChinaSE (Fig. 3). Results for ~~the other 14~~all regions are shown in ~~the Appendix~~Fig. S7. For each region, statistics (R, % of points in ~~SE and~~ GE, offset and RMSE) for all ~~15~~ 16 products are combined into one subplot ~~(Fig. 5 and Fig. A2)~~. The merged AOD product M is introduced in Sect. ~~6~~5.2; evaluation results ~~of~~ for that product ~~is discussed~~are summarised in sect. ~~6~~5.2.~~3~~1.

Algorithm performance over Europe is similar for most products, with R of 0.55-0.65, 45-55% of the pixels in the GE ~~and 70-80% of the pixels in the SE~~, offset of 0.05-0.1, and RMSE of ~0.1. For TOMS and~~,~~ OMI, ~~and POLDER~~ the performance is slightly worse than for other products in Europe. In ChinaSE (a region of particular interest having high AOD loading, ~~mainly~~ due mainly to high levels of anthropogenic aerosols, which, however, decreased steadily during the last decade, Sogacheva et al., 2018b), the offset (0.1-0.2) and RMSE (0.2-0.3) are considerably higher than in Europe, and fewer pixels fit within the ~~SE (55-70%) and~~ GE (15-30%). This is likely due to a combination of high AOD loading and accompanying~~ied~~ high uncertainty in the products, high variability in aerosol composition and surface properties. In Indonesia and for the biomass burning outflow over the Atlantic, the ~~EOS~~ MODIS and MISR products show better agreement with AERONET than the ATSR-family products.

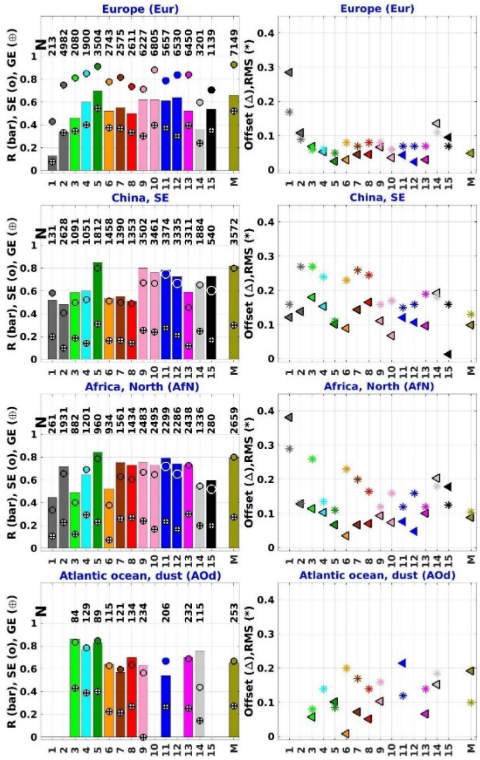
Several ~~P~~products ~~with~~ which use different surface treatment (ATSR SU, MODIS-family, MISR) show similarly higher R over AfN, an area of high surface reflectance. However, a high R does not imply that ~~the~~ performance is better, only that variations in AOD are captured better. Other statistics (number of pixels within GE, offset and RMSE) in AfN are worse compared with those in Europe.

Overall, no single product has the best statistics for all metrics and regions. Retrievals tend to perform well in areas with darker (more vegetated) surfaces and where aerosol type is less variable over time. In these cases, biases are small and retrieval uncertainties are often better than the GE, tracking temporal AOD variability well but with a tendency to underestimate high-AOD events. In more complex tropical environments, data should be used with greater caution, as there is a greater tendency to underestimate AOD. However, correlation often remains high, suggesting good ability to identify monthly AOD variations, despite this underestimation.

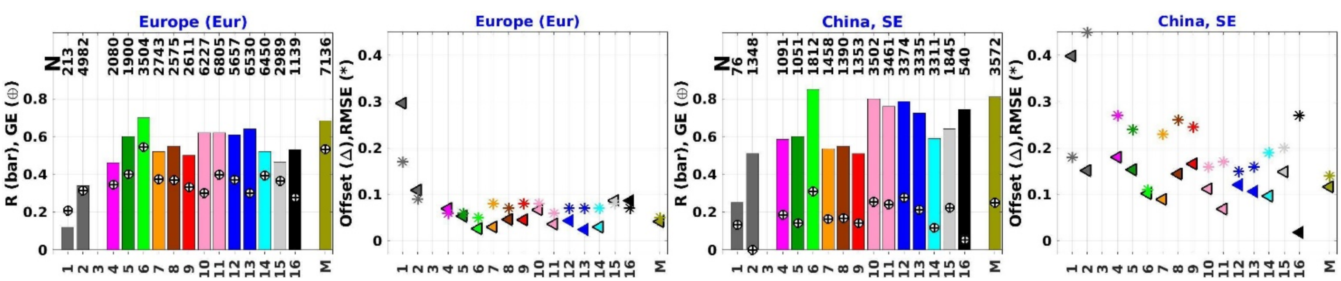
~~Overall, no single product has the best statistics for all metrics and regions. Retrievals tend to perform well in areas with darker (more vegetated) surfaces and where aerosol type is less variable over time. In these cases, biases are small and retrieval uncertainties are often better than defined SE, tracking temporal AOD variability well but with a tendency to underestimate high-AOD events. In more complex tropical environments, data should be used with greater caution, as there is a larger tendency to underestimate AOD. However, correlation often remains high, suggesting good ability to identify monthly AOD, despite this underestimation.~~



~~Figure 4. Global AERONET comparison of the AOD monthly aggregates (circles – median bias, error bar – bias standard deviation) for all AOD types (purple), fine-dominated AOD (blue) and coarse-dominated AOD (green) and fraction of points in each bin (bar, orange)~~



1 TOMS 3 AVHRR 6 ATSR_ADV 9 Terra DT&DB 13 MISR
2 OMI 4 SeaWiFS 7 ATSR_SU 10 Terra MAIAC 14 POLDER
5 VIIRS 8 ATSR_ens 11 Aqua DT&DB 15 EPIC MAIAC
12 Aqua MAIAC M merged product



1 TOMS 3 AVHRR_NOAA 7 ATSR_ADV 10 Terra DT&DB 14 MISR
2 OMI 4 AVHRR_DT/SOAR 8 ATSR_SU 11 Terra MAIAC 15 POLDER
5 SeaWiFS 9 ATSR_ens 12 Aqua DT&DB 16 EPIC
6 VIIRS 13 Aqua MAIAC M Merged product

Figure ~~5~~3. AERONET evaluation~~comparison~~ statistics for Europe and ChinaSE(~~Left~~: correlation coefficient R, bar;~~ fraction of pixels in the Spread Envelope, SE, ○;~~ fraction of pixels satisfying the GCOS requirements, GE, ⊕;~~ Right:~~ Offset (satellite product-AERONET), Δ; root mean square error RMSE, *) for AOD monthly aggregates for each product (1:~~15~~16, legend for products below the plot) and the L3 merged product (M, approach2, RM2 for "all" aerosol types, for details see Sect.~~6~~4.2) with corresponding colours (legend) for the selected regions (as in Fig. 1). N is a number of matches with AERONET. Note, for products which do not provide the global coverage (e.g., no retrieval over oceans), the results are missing. For ~~other~~ all studied regions, see Fig. ~~A2~~S7.

4~~3~~.~~3~~ 2 AOD time series

In order to move towards consistency in regional and global AOD records derived from multiple satellites using different sensors and retrieval techniques, this section examines annual regional AOD time series obtained from the different products. Besides the positive offset for TOMS and~~,~~ OMI~~ and POLDER~~ (~~see Sect. 4.2~~Figs. S1, S2, S6 and S7), consistent temporal patterns are observed, and similar interannual AOD variability is tracked by all datasets (Fig. ~~6~~ 4 and Fig. ~~A2~~S8). AOD peaks in Europe in 2002, in ChinaSE in 2006/2007, 2011 and 2014, (possibly related to changes in anthropogenic emissions, Sogacheva et al., 2018a, 2018b). Relative AOD peaks over the Atlantic dust area in 1998, 2012 and 2015 (Peyridieu et al., 2013), and obvious AOD peaks in Indonesia related to the intensive forest fires in 1997, 2002, 2006 and 2015 (Chang et al., 2015;~~,~~ Shi et al., 2019) are clearly seen.

~~However, significant regional offsets between products exist, which are largest in regions with high aerosol loading. Over ChinaSE, MODIS-family products show higher monthly AOD compared to all others. Over AfN, the ATSR_SU and ATSR_ensemble reach higher monthly aggregated AOD than the MODIS-family products, whereas comparisons with AERONET are similar for ATSR and MODIS (with slightly higher RMSE for ATSR by 0.05); differences are likely tied to the small number of stations in this region. A high offset between MODIS and ATSR is revealed over Australia (Fig. A3).~~

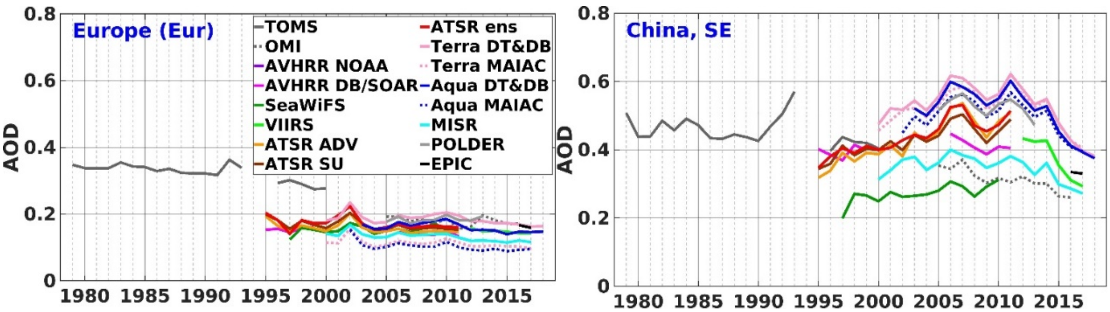


Figure 4. Annual AOD time series from different products (see legend) for Europe and ChinaSE. For all selected regions see Fig. S8.

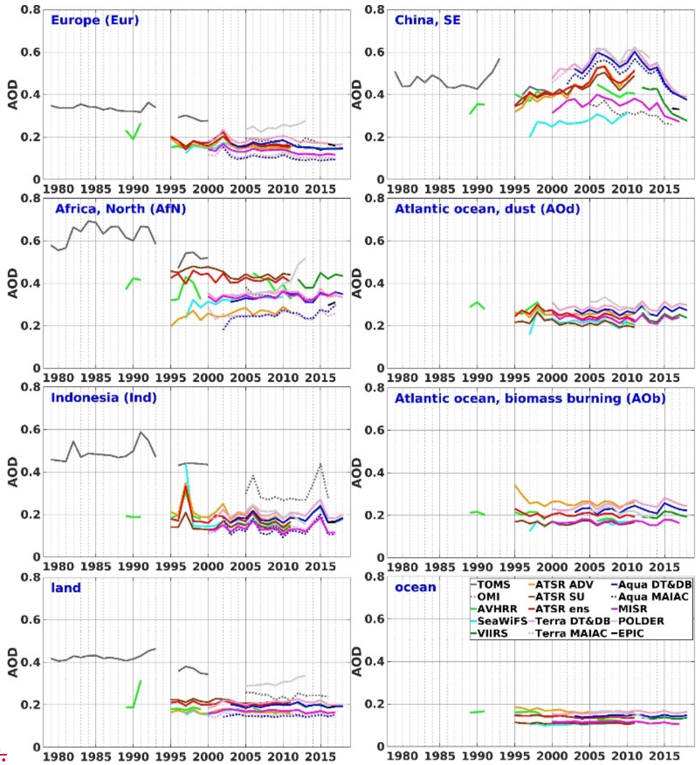
However, significant regional offsets between products exist, which are largest in regions with high aerosol loading. Over ChinaSE, MODIS-family products show higher monthly AOD compared to all others. Over AfN, the ATSR_SU and ATSR_ensemble reach higher monthly aggregated AOD than the MODIS-family products, whereas comparisons with AERONET are similar for ATSR and MODIS (with slightly higher RMSE for ATSR by 0.05); differences are likely tied to the small number of stations in this region. A high offset between MODIS and ATSR is revealed over Australia (Fig. S8). AOD annual cycles for individual products for year 2008 are discussed in Sect. S8. As the annual time series (Fig.4 and S8), the annual AOD cycle between the products (Fig. S9) are similar, with more pronounced deviation in the areas of high AOD.

4.4 AOD annual cycles

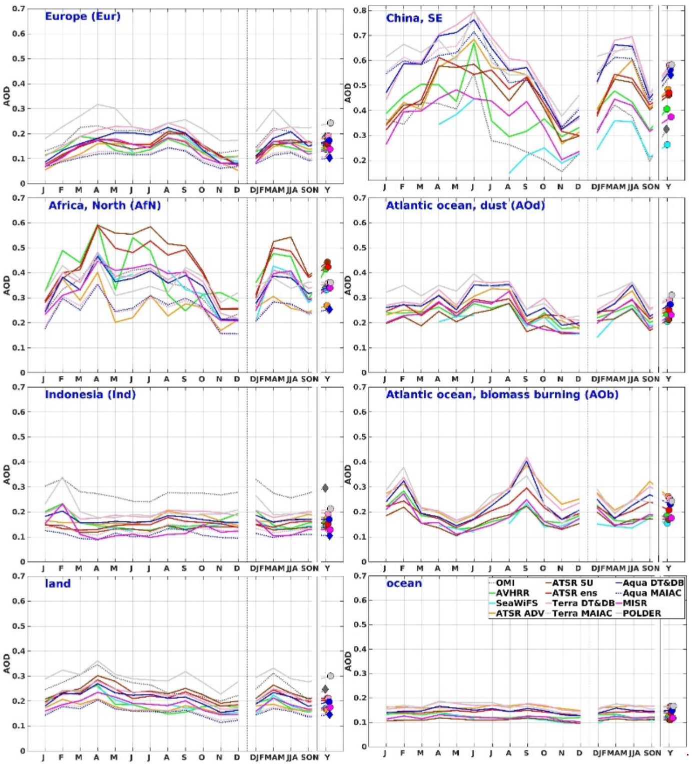
Year 2008 was chosen for annual AOD cycle inter-comparison, when data from all products except TOMS, VIIRS and EPIC are available.

Annual AOD variations among the products (Fig. 7 and Fig. A4) agree better in regions with relatively low AOD (e.g., in Europe). In Europe, the relative increase towards summer and further increase in August-September is captured by all products. In ChinaSE, where AOD loading and annual variation are higher, the AOD increase from winter to summer and another AOD peak in September are observed. Outbreaks of the biomass burning aerosols over Atlantic produce clear peaks in February and September, as shown in all available products.

Thus, besides the similar interannual variation among the satellite products, the AOD annual cycles among the products is also similar.



~~**Figure 6. Annual AOD time series for the selected regions. For 2018, MODIS DT&DB and VIIRS AOD products are shown here but not used in the merging exercise. For other regions, see Fig. A3.**~~



~~**Figure 7. Monthly (J for January, F for February, etc., left plots) and seasonal (DJF, MAM, JJA, SON, middle plots) AOD time series for the selected regions for 2008, and yearly aggregate AOD for 2008 (Y, right) for selected regions. For other regions, see Fig. A4.**~~

~~5~~ 4 AOD merging ~~methods~~approaches

In the current study, 12 AOD products, all available at 0.55 µm, were used to create a merged AOD product for the period of 1995-2017. The temporal availability of the AOD products is shown in Table 2.

Table 2. Availability and coverage of the AOD products for merging for each year in the period 1995-2017. N – annual number of available products

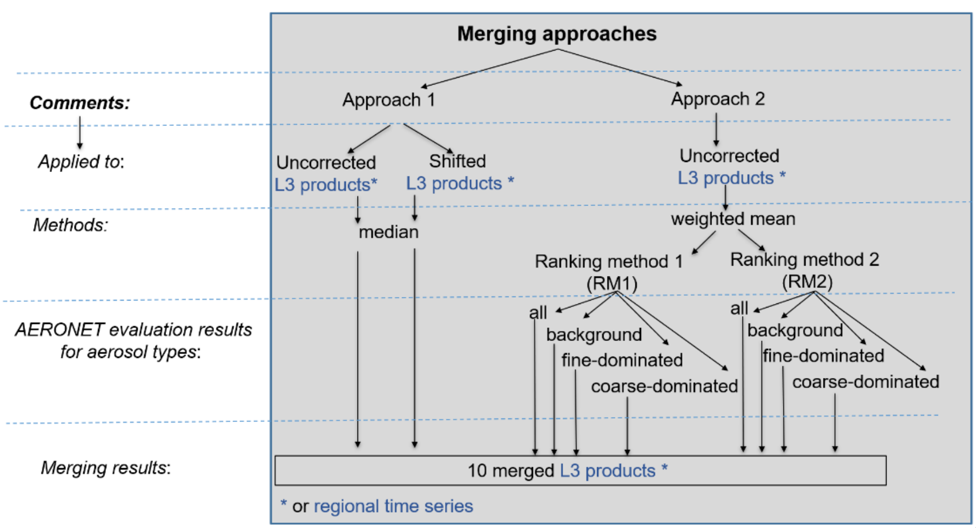
year	AVHRR DT/SOAR global	SeaWiFS global	VIIRS global	ATSR ADV global	ATSR SU global	ATSR ensemble global	Terra DT&DB global	Terra MAIAC land	Aqua DT&DB global	Aqua MAIAC land	MISR global	POLDER global	N
1995	x			x	x	x							4
1996	x			x	x	x							4
1997	x	x		x	x	x							5
1998	x	x		x	x	x							5
1999	x	x		x	x	x							5
2000		x		x	x	x	x	x			x		7
2001		x		x	x	x	x	x			x		7

2002		x		x	x	x	x	x		x	x		8
2003		x		x	x	x	x	x	x	x	x	x	9
2004		x		x	x	x	x	x	x	x	x	x	9
2005		x		x	x	x	x	x	x	x	x	x	10
2006	x	x		x	x	x	x	x	x	x	x	x	11
2007	x	x		x	x	x	x	x	x	x	x	x	11
2008	x	x		x	x	x	x	x	x	x	x	x	11
2009	x	x		x	x	x	x	x	x	x	x	x	11
2010	x	x		x	x	x	x	x	x	x	x	x	11
2011	x			x	x	x	x	x	x	x	x	x	10
2012			x				x	x	x	x	x	x	7
2013			x				x	x	x	x	x	x	7
2014			x				x	x	x	x	x		6
2015			x				x	x	x	x	x		6
2016			x				x	x	x	x	x		6
2017			x				x	x	x	x	x		6

We tested ~~and compared three~~ two ~~different~~ approaches for merging ~~regional and global AOD time~~ , Fig.5. In the first approach, the ~~records:~~AOD ~~Mean and~~ median ~~for~~ of the ~~15~~ available (10 globally and 2 over land) individual uncorrected ~~records~~ and offset-adjusted (shifted to a common value) products were calculated (approach 1, Sect. 4.1 for details). In the second approach, an ~~Mean and median of 15 offset-corrected records (approach 2)~~AOD weighted mean, the weights for individual products derived from the evaluation with ~~results of~~ the AERONET were utilised ~~verification~~ (approach 2~~3~~, Sect. 4.2 for details). The same merging scheme was applied to the L3 uncorrected products (Sect 2.2) and regional time series (Sect. 3.1); as a result, 10 merged AOD products and 10 merged regional time series for the period 1995-2017 were created.

~~AOD annual time series from all available products were merged for the period 1995-2017.~~

~~To achieve best estimates of the regional AOD by merging multi-sensor monthly AOD data, the uncertainties should be considered explicitly. However, this cannot (yet) be done, as currently only some of the L2 products contain pixel-level propagated or estimated uncertainties, and none of the monthly products contain anything beyond simple averages of those~~



~~L2 uncertainties.~~

Figure 5. Scheme for the merging approaches; applied for L3 products or regional time series.

To achieve best estimates of the regional AOD by merging multi-sensor monthly AOD data, the systematic and random components of uncertainties within each product should be considered explicitly. However, this cannot be done (yet), as currently only some of the L2 products used to create the L3 monthly products contain pixel-level propagated or estimated uncertainties. These uncertainties have not yet been aggregated to the L3 products, and sampling components of L3 uncertainty have not yet been quantified robustly. The analysis herein therefore represents and initial effort in the absence of a full uncertainty budget. Standard uncertainties for the chosen merged L3 product (details are discussed in Sect. 5.2.2) were estimated as the root mean squared sum of the deviations between the chosen merged product, the median from the all uncorrected products (approach 1) and each of the other seven merged products (approach 2)

54.1 Approach 1: AOD mean, median for uncorrected and offset-adjusted (shifted) AOD products

4.1.1 Method

The mean (arithmetic average) value, while althougha commonly statistic used in climate studies, is not generally equal to the most commonly frequently occurring value (the mode) and may not be representative ofreflect the central tendency (the median) of strongly asymmetrical distributions such as can be found for AOD (O’Neill et al., 2000; Sayer and Knobelspiesse 2019). Although the central limit theorem implies that this should be less of an effect when making an estimate of the mean AOD from a cluster of AOD data sets (i.e. a merged time series), in practice this is unlikely to be fully the case because the different data sets are not independent estimates of the underlying AOD field This is because they are made with sensors and techniques which are not independent (i.e. typically similar spectral/spatial bands and sampling limitations), and they may have different bias characteristics. Further, bBy itself, the mean does not provide any information about how the observations are scattered, whether they are tightly grouped or broadly spread out. Thus, to describe the asymmetry, in addition to the mean, we also reportstudy the median (which is more robust to outliers which might be caused by a poorly-performing algorithm in a certain region) and standard deviations (as a metric of diversity) between the products chosen for merging.

5.2 Approach 2: AOD mean, median for shifted AOD

As shown in Sect. 43, the AOD time series of different products are display highly consistent, showing similar temporal patterns. However, offsets between the products exist, which vary globally and seasonally (Figs. 64, S8, Fig. 7 and S9). We use the Terra DT&DB product as a reference tTo estimate the average offsets between the products, because its time period (early 2000 onwards) , we chose as a reference the ATSR_ensemble product, which is closest to the median value ofoverlaps with each AOD product considered in the current study. The mean biases between each product and the Terra DT&DB product were calculated for the overlapping periods of the annual, seasonal and monthly products separately. Shifted AOD products were obtained by adjusting the AOD products with the corresponding biases.all products AOD (with some exceptions

~~discussed below). As in approach 1,~~ AOD ~~mean,~~ median and standard deviation were calculated from all available AOD shifted products. ~~datasets, shifted to the ATSR_ensemble products.~~

4.1.2 Regional offset adjustment

Regional differences in the offsets between Terra DT&DB and other products exist. Means and standard deviations of the offsets for all individual products from the Terra DT&DB AOD are shown in Fig. 6 for Europe and ChinaSE and in Fig. S10 for all selected regions. Offset magnitudes and their variations depend on AOD loading: offsets are typically higher for high AOD. Over land, ocean and thus globally, the offset is negative relative to Terra DT&BD for most of the products. This includes Europe and ChinaSE. However, over the bright surface area in Northern Africa, AVHRR DT/SOAR, VIIRS, ATSR SU and ATSR ensemble show high (0.05-0.1) positive bias. Also, all ATSR products are biased high in Australia and South America. Thus, the median for the offset-adjusted product is expected to be positive-biased. For details, see Sect.5.1, where evaluation results for the AOD products merged with different approaches are discussed.

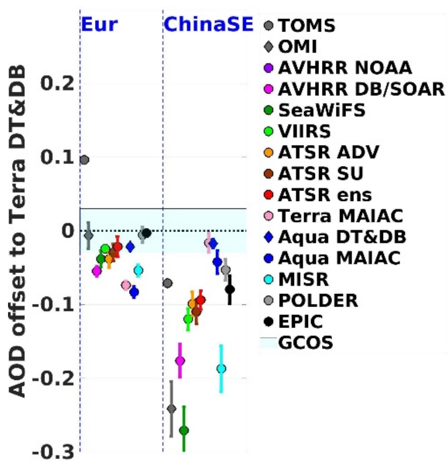


Figure 6. Regional annual average AOD offset between each dataset and the Terra DT&DB dataset. GCOS requirement of ±0.03 is shown as a background colour. For all selected regions, see Fig. S10.

With the shifted median merging approach, each AOD product was shifted based on its regional offset with respect to the Terra DT&DB (Sect. 5.2). Median and standard deviation of AOD time series were then derived from these 10 shifted and Terra DT&DB data records.

~~5~~4.~~3~~ 2 Approach ~~3~~2: Weighted AOD

4.2.1. Method

As shown in Sect. ~~4~~3.~~2~~1, the products differ in the degree to which each represents the AERONET values on the monthly scale. Our second approach is a weighted mean AOD, where the weights are assigned based on each data set's agreement with monthly AERONET averages. This thus represents an initial attempt to adjust the level of confident assigned to each product on a regional basis; the better-comparing products are given more weight in the calculation of a combined product.~~A weighted AOD mean is calculated as our third approach, by assigning weights based on their agreement with AERONET. The better-compared products are given more weight in the calculation of a combined product.~~

An AOD-weighted mean was calculated, with a ranking approach based on the statistics from the AERONET comparison for AOD: R, bias, RMSE, GE (Fig. ~~5~~4 and S8) and median bias of the binned AOD in the range [0.45, 1] (Fig. ~~4~~3 and S7). The last criterion was added to specifically consider algorithm performance for higher AOD.

Two ranking methods were tested. For the first ranking method (RM1) based on best statistics, the ~~15~~ 12 products were ranked from 1 (worst) to ~~15~~ 12 (best) for each statistic (R, GE, RMSE, bias, binned bias) separately. The 5 separate ranks were then summed, so the maximum possible rank is ~~15~~12*5=~~75~~60. Possible errors in RM1 occur when several products have similar statistics, so small variations in statistics can produce large changes in ranking. Note, that no product received a perfect (~~75~~60) rating.

In the second ranking method based on binned statistics (RM2), a rank from 1 to 10 was assigned to each metric separately according to the bin number. For each statistic, the following windows: [0.5, 1] for R, [0, 0.5] for GE, [0, 0.2] for bias, [0, 0.15] for RMSE, and [-0.5, 0] for the binned bias were divided into 10 bins. In that exercise, several algorithms could be ranked ~~similarly~~ equally for certain statistics, if their statistics fell within the same bin. For example, if R for three products is between 0.8 and 0.85, all three were ranked 8 for that statistic. If for all ~~15~~ 12 the R was between 0.6 and 0.65, they all would receive rank 4 with RM2, whereas with the RM1 approach, they were ranked from 1 to ~~15~~12, which caused an artificial bias in the ranking. In RM2 ~~is more logical, as~~ possible errors in statistics are considered when operating with the bins. For example, R might be slightly biased by a few outliers. If R is 0.82 or 0.81, the same rank of 8 is given in RM2. In RM1, the lower rank is given for the product with R= 0.815 than for the product with R= 0.82. ~~Ranking results for RM1and RM2 are slightly different, but the resulting merged products are similar (not shown here). Thus, RM2 results (discussed below in Sect. 6.3) were used to determine the weight of each AOD product.~~

The sum of the five ranks (R, GE, RMSE, bias, binned bias) , w, for each product i was calculated and transformed to a weight of each product (as a fraction of total sum for the product from the total sum of ranking for all products) to calculate the AOD weighted mean, $\overline{AOD}$:~~.~~

$$\overline{AOD} = \frac{\sum_{i=1}^{n}(w_i * AOD_i)}{\sum_{i=1}^{n}(w_i)}$$

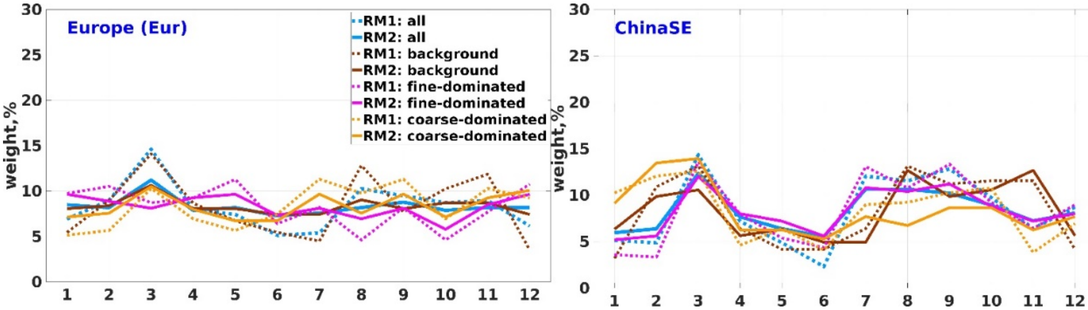
As shown in Sect. 4.2.1, the performance of the retrieval algorithms often depends on the aerosol conditions (aerosol type and loading, Fig. 42) and surface properties. Accordingly, weights for the different AOD products were calculated separately for each region for different aerosol types (“background”, “” or “coarse-dominated”) separately and “all” aerosol types together considering the corresponding regional statistics from the AERONET comparison. However, aerosol types often change in time and space within the same region (Fig. S5) and thus can’t be defined with a high level of confidence (with the methods applied in the current study). Thus, those weights for each aerosol type were applied globally to merge both L3 monthly products and time series. As a result, eight merged AOD products were obtained, calculated with RM1 and RM2 for three different aerosol types and “all” aerosol types.

4.2.2 Ranking results (weights) for individual products

Results for RM1 and RM2 ranking, the contribution of each product to the merged data product expressed as a weight, based on evaluation of the single products with AERONET, are shown in Fig. 7 (Europe and ChinaSE) and Fig. S11 (all selected regions) for three aerosol types (“background”, ”fine-“ and “coarse-dominated”) and all aerosol types together (“all”). With some exceptions (e.g., in AOb, where the RM2 weight of Aqua DT&DB is ca 15% higher for coarse-dominated type, and in Australia, where the RM2 weight of SeaWiFS and Aqua MAIAC is 10-15% higher for coarse-dominated type, Fig. S11), the difference in weights obtained with RM1 and RM2, if they exist, do not exceed 5-10%. Thus, the ranking methods RM1 and RM2 introduced in the current study produce similar results. Some products show better performance for certain aerosol types (Fig. 4 and S4). Thus, the weight of the product depends on which aerosol type is assumed for merging. E.g., in Europe, VIIRS has lower weight for fine-dominated aerosols, whereas the corresponding weight for ATSR SU is higher for that aerosol type. In ChinaSE, Terra DT&DB performs worse than Terra MAIAC for background aerosols, thus for that aerosol type the weight for Terra MAIAC is higher.

As with the results discussed in Sect. 3, none of the algorithms consistently outperforms the others in all regions. There is no clear leader over Europe, a region with low AOD, indicating similar performance of all algorithms under background conditions. Over land globally, also a region with low AOD, the ranks are similar for EOS and ATSR, with somewhat higher number for VIIRS. Over ocean globally, the ranks are similar for all existing products. One likely reason that the VIIRS and MODIS ranks are often higher is their better coverage, which enables them to better represent AERONET monthly means over land. However, MODIS is ranked lower over the Atlantic dust region. The lowest ranks are obtained consistently for TOMS, OMI and POLDER, due to their high biases.

Ranks for the different aerosol classes (all, background, fine-dominated and coarse-dominated) are different, which raises another aspect of using multiple products. Over land, MODIS MAIAC often has a higher rank for background AOD, whereas MODIS DT&DB are better for other aerosol types.



1 AVHRR DB/SOAR	4 ATSR_ADV	7 Terra DT&DB	10 Aqua MAIAC
2 SeaWiFS	5 ATSR_SU	8 Terra MAIAC	11 MISR
3 VIIRS	6 ATSR_ens	9 Aqua DT&DB	12 POLDER

Figure 7. Weights of each product obtained with RM1 and RM2 for Europe (left) and China (right) for different aerosol types (all, background, fine- and coarse-dominated). For all regions, see Fig. S11

6 AOD merging results

6.1 Approach 1 and 2 : AOD mean, median for uncorrected and shifted AOD time series

6.1.1 Mean offset between AOD products and ATSR_ensemble AOD

Means and standard deviations of the offsets for all datasets from the ATSR_ensemble AOD are shown in Fig. 8 and Fig. A5 for selected regions. For VIIRS and EPIC, which do not overlap (in time) with ATSR, the offset to the ATRS ensemble was estimated by adding the offset to Terra DT&DB plus the offset of Terra DT&DB to the ATSR ensemble.

Offset magnitudes and their variations depend on AOD loading: offsets are typically higher for high AOD. In Europe, the offset for most products falls within ±0.04 AOD. OMI, VIIRS, Terra, Aqua and EPIC are offset positive, whereas SeaWiFS, AVHRR, ATSR_ADV, ATSR_SU and MISR show negative offsets from the ATSR ensemble. TOMS and POLDER show larger positive offsets (0.12 and 0.08, respectively) whereas MODIS MAIAC show high (ca. 0.06) negative offsets.

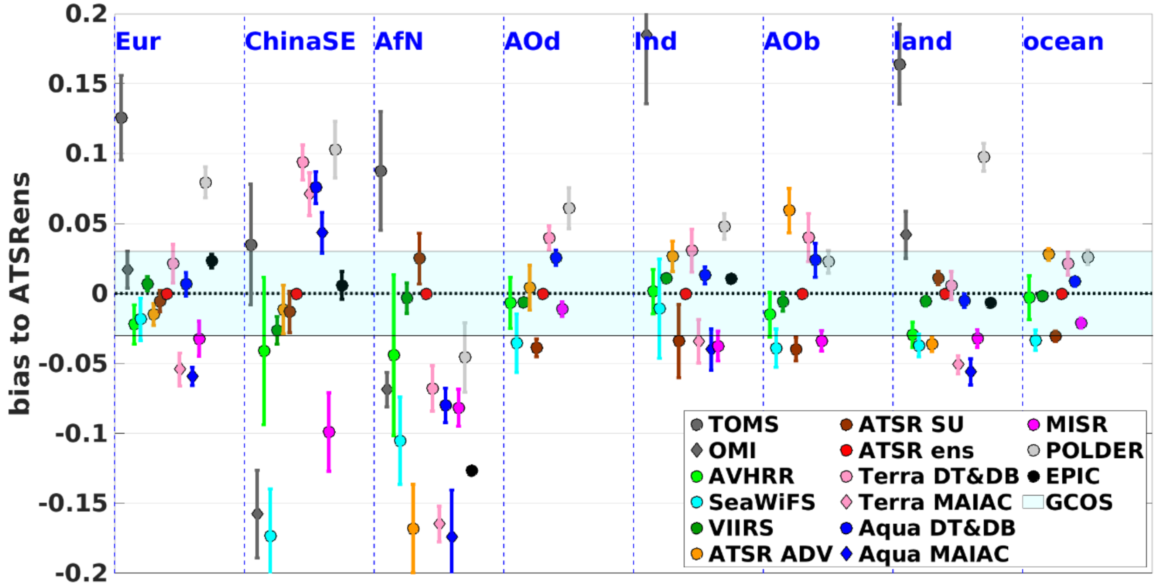


Figure 8. Regional annual average AOD offset between each dataset and the ATSR_ensemble dataset. GCOS requirement for ±0.03 AOD is shown as a background color. For other regions, see Fig. A5.

For most products and regions, Fig. 8 shows similar tendencies in the signs of offsets. Offset magnitudes and their variations depend on AOD loading: offsets are typically higher for high AOD. However, there are exceptions. In ChinaSE, MODIS MAIAC shows positive offsets. In AfN, most products, except TOMS and ATSR_SU, have negative offsets.

Over land, VIIRS, MODIS DT&DB have the lowest offsets (±0.01) relative to the ATSR ensemble product, whereas over ocean AVHRR, VIIRS and Terra DT&DB offsets are close to zero.

With the second merging approach, each AOD product was shifted based on its regional offset with respect to the ATSR ensemble (Sect. 5.2). Mean, median and standard deviation AOD time series were then derived from these 15 shifted data records.

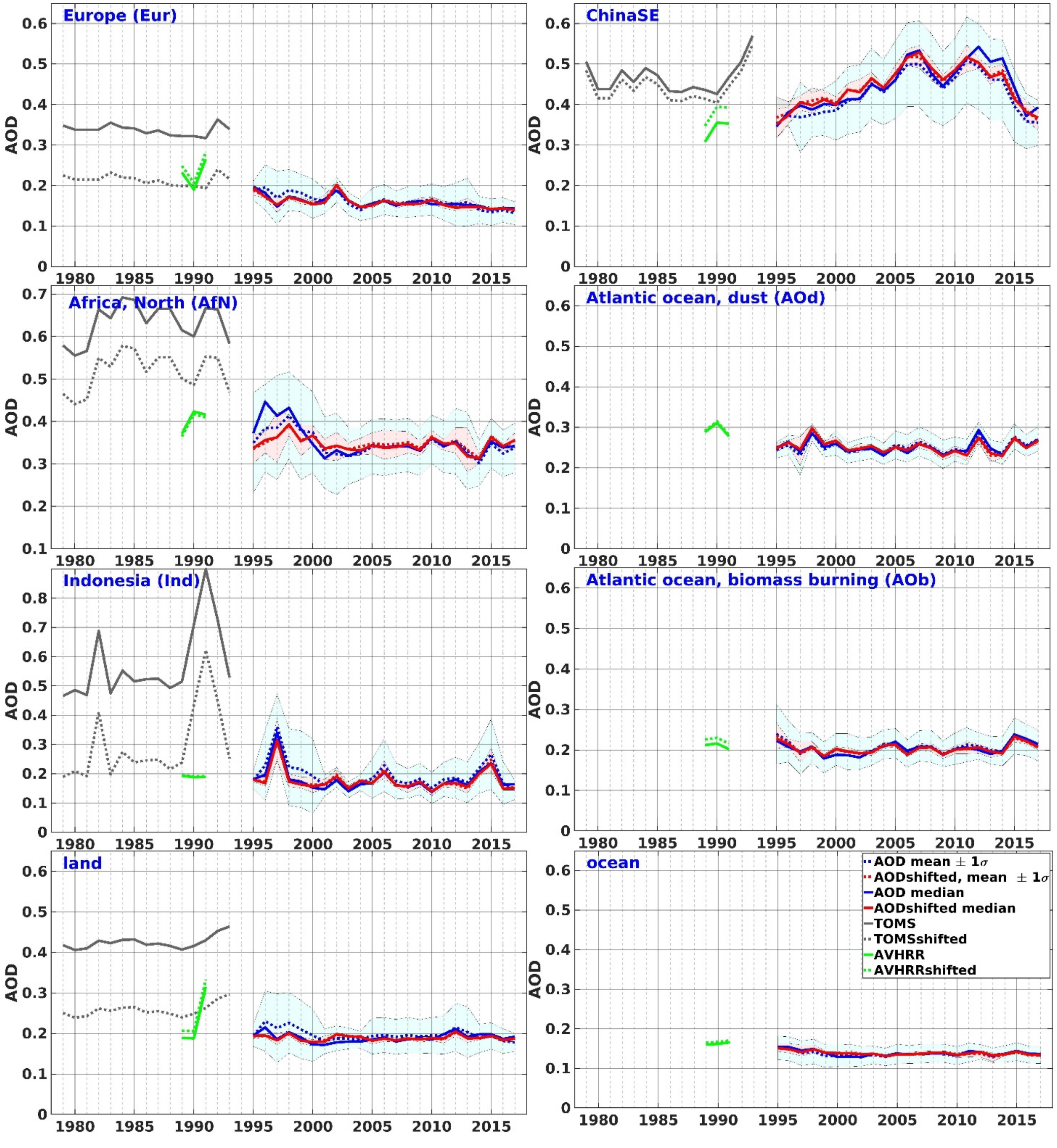
6.1.2 Mean and median AOD time series

Regional mean and median AOD time series calculated from all available uncorrected data records (approach 1) are shown in Fig. 9 and Fig. A6 for the period 1995-2017. Mean and median AOD time series from products shifted to the ATSR_ensemble (approach 2) are also plotted. Products and shifted products from TOMS (1979-1993) and AVHRR (1989-1991) are shown, when available.

The two merging approaches tested here agree well except in NAf, where the ATSR_ensemble (used as reference) is biased high, as shown from comparison with AERONET, which leads to a positive bias in the shifted median compared to the simple median of uncorrected products. The agreement of the two approaches is encouraging, as we can conclude that for the big-picture analysis of overall trends, details of the methodology do not matter so much.

There are some smaller differences between the two approaches. Time series from uncorrected records are biased high compared to the records from the shifted products, due to the high bias of TOMS, OMI and POLDER, which also leads to a larger standard deviation. By adding both MODIS products beginning in 2002, the uncorrected AOD standard deviation becomes lower over land. For the period before 1995, where only one or two products (TOMS and AVHRR) exist, no meaningful merging can be applied; only in some regions the shifted records for this period come close to the average level of the second part of the merged records.

Using the temporal records to monitor regional AOD trends seems plausible. As demonstrated in previous studies (e.g. Zhang and Reid, 2010; Hsu et al., 2012; Chin et al., 2014), the merged records show a small but gradual decrease of AOD over Europe (with one small peak in 2002). Spatial consistency is indicated by high correlation (similar positions of peaks) in NAfr and its Atlantic dust outflow region. Interannual variation as well as the standard deviations are highest for regions with the largest AOD, e.g., over ChinaSE (anthropogenic emissions) and Indonesia (biomass burning). The time series of ChinaSE follows the known patterns caused by step-wise regional emission reductions in the last 25 years (Sogacheva et al., 2018b).



~~Figure 9. Merged time series from uncorrected AODs (approach 1, blue) and from shifted AODs (approach 2, red) for the 8 selected regions (for both: mean, dotted line, median, solid line, and 1σ shadowed area). For other regions see Fig. A6~~

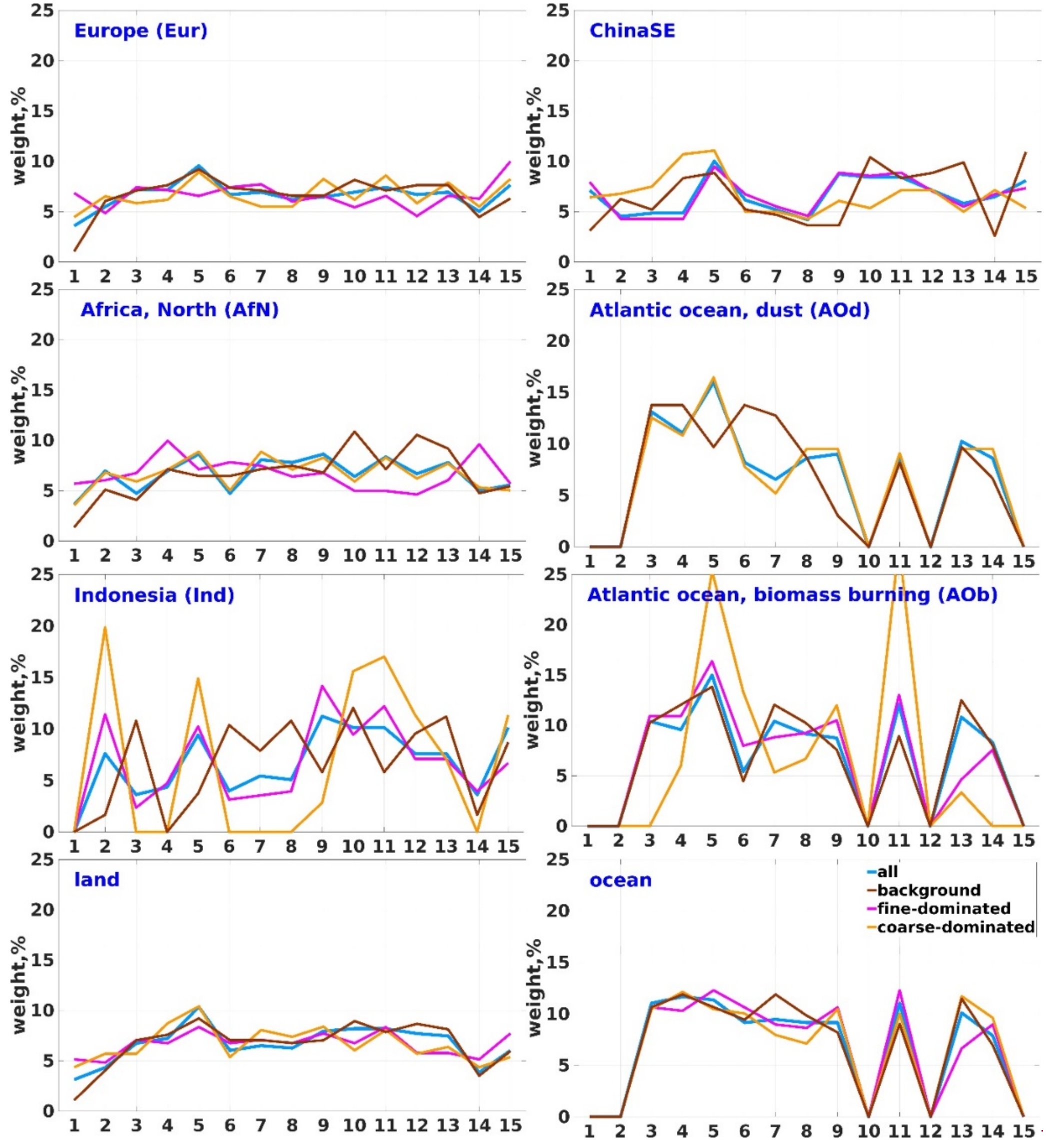
6.2 Approach 3: weighted AOD

6.2.1 Ranking results

Results for RM2 ranking are shown in Fig. 10 and Fig. A6 for three aerosol types (background, fine- and coarse-dominated) and all aerosol types. As with the results discussed in Sect. 4, none of the algorithms consistently outperforms the others in all regions. There is no clear leader over Europe, a region with low AOD, indicating good performance of all algorithms under background conditions. MODIS is ranked lower over the Atlantic dust region. Over land globally, the ranks are similar for EOS and ATSR, with somewhat higher number for VIIRS.

Over ocean globally, the ranks are similar for all existing products (again a region with low AOD). One reason that the VIIRS and MODIS ranks are often higher is likely their better coverage, which enables them to better represent AERONET monthly means over land. The lowest ranks are obtained consistently for TOMS, OMI and POLDER, due to their high biases.

Ranks for the different aerosol classes (all, background, fine-dominated and coarse-dominated) are different, which raises another aspect of using multiple products. Over land, MODIS MAIAC often has a higher rank for background AOD, whereas MODIS DT&DB are better for other aerosol types.



~~1 TOMS 3 AVHRR 6 ATSR_ADV 9 Terra DT&DB 13 MISR~~
~~2 OMI 4 SeaWiFS 7 ATSR_SU 10 Terra MAIAC 14 POLDER~~
~~5 VIIRS 8 ATSR_ens 11 Aqua DT&DB 15 EPIC MAIAC~~
~~12 Aqua MAIAC~~

~~**Figure 10. Weight of each product for the selected regions for different aerosol types. For other regions, see Fig. A7**~~

~~6.2.2~~5 ~~AOD monthly aggregates merged based on ranking results~~Merged L3 AOD products

Using the~~-~~ median and median for shifted products (approach 1)~~regional weights (Fig. 10 and Fig. A7) for three different aerosol types and all types, the Approach 3 (weighted AOD),~~ ~~was applied to the~~ two merged L3 monthly 0.55 µm AOD ~~aggregates~~ products were created ~~of all~~from all available AOD products (at 0.55 µm) ~~available~~ since 1995 (Table 1). Regional ranking results for four aerosol types (“all”, “background”, “fine-” and “coarse-dominated”) were considered in the weighted AOD (approach 2, ranking methods RM1 and RM2). ~~As the ranking was slightly different for different aerosol types, merging of the L3 AOD was done for three aerosol types separately, and for all aerosol particles together when the aerosol type was not specified.~~ Altogether, 10 merged L3 AOD products were created and evaluated against AERONET.

5.1 Evaluation of the all merged L3 AOD products with AERONET

To estimate the quality of the AOD ~~merged~~ L3 monthly products merged with different approaches, we performed an exercise to evaluate the merged AOD ~~(approach 3, RM2, all aerosol types)~~ against AERONET AOD, similar to the one used for ~~merging~~ evaluation of the individual~~all~~ products (Sect. ~~4.2~~3.1). Evaluation results reveal similarities in the accuracy of products merged with different approaches. The AOD binned bias of the merged products (Fig. ~~11~~S12~~, left~~) shows a similarly small deviation from AERONET (±0.03) for AOD<0.5 (positive for AOD <0.3 and negative for 0.3<AOD<0.5), where the fraction of values falling within the bins is about 0.95. The offset is slightly higher for the median of the shifted AOD product (approach 1), because as discussed earlier, Terra DT&DB has a positive bias relative to most of the other individual products; this results in slightly elevated AOD merged with approach 2. For AOD>0.5, where the number of cases is very low, the underestimation increases as AOD increases. As for individual products, the coarse-dominated merged products have offset with AERONET.~~For AOD<1.25, the deviation is similar for all aerosol types.~~

Correlation coefficient, number of the pixels in the GE, offset and RMSE for the AOD merged product are shown in Fig. ~~5~~ 8 for ~~selected regions~~Europe and ChinaSE and in Fig. S13 for all regions. For the AOD merged products ~~, which~~ has the best temporal coverage and~~,~~ the number of points used for validation (N) is higher than for any individual product. ~~Comparison with these metrics for the individual products in Fig. 5 shows that the~~ The correlation coefficients and the number of the pixels in the ~~SE~~ GE are as high as for the ~~1-2~~one or two best -ranked products in the corresponding regions, except for the product merged with approach 2. The offset is close to the average~~d~~ offset, and the RMSE tends to be lowest. Thus, the quality of the merged products, except for the shifted AOD product, is as good as that of the most highly ranked individual AOD products in each region.

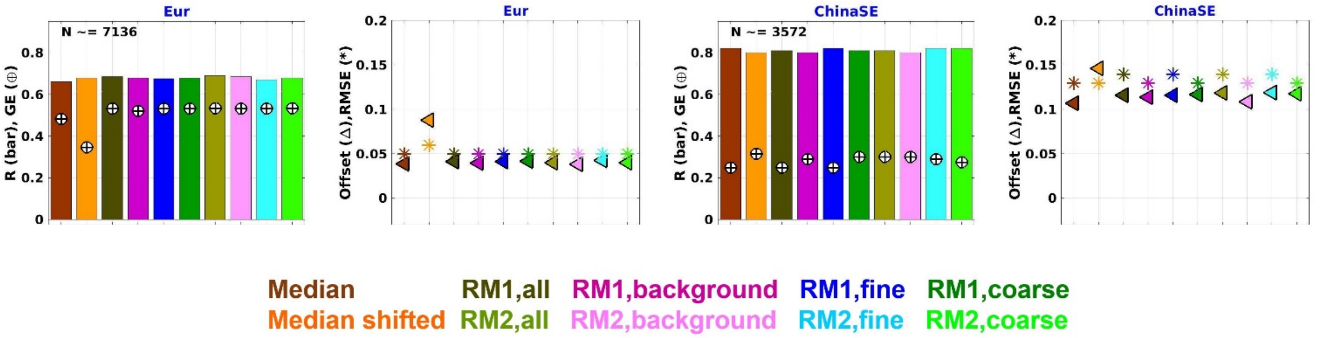


Figure 8. AERONET comparison statistics: correlation coefficient R, bar; fraction of pixels satisfying the GCOS requirements, GE, ⊕; b, d: Offset, Δ; root mean square error RMSE, *; for AOD products merged with different approaches, median, shifted median, RM1, RM2 for different aerosol types for Europe and ChinaSE. For all regions, see Fig. S13.

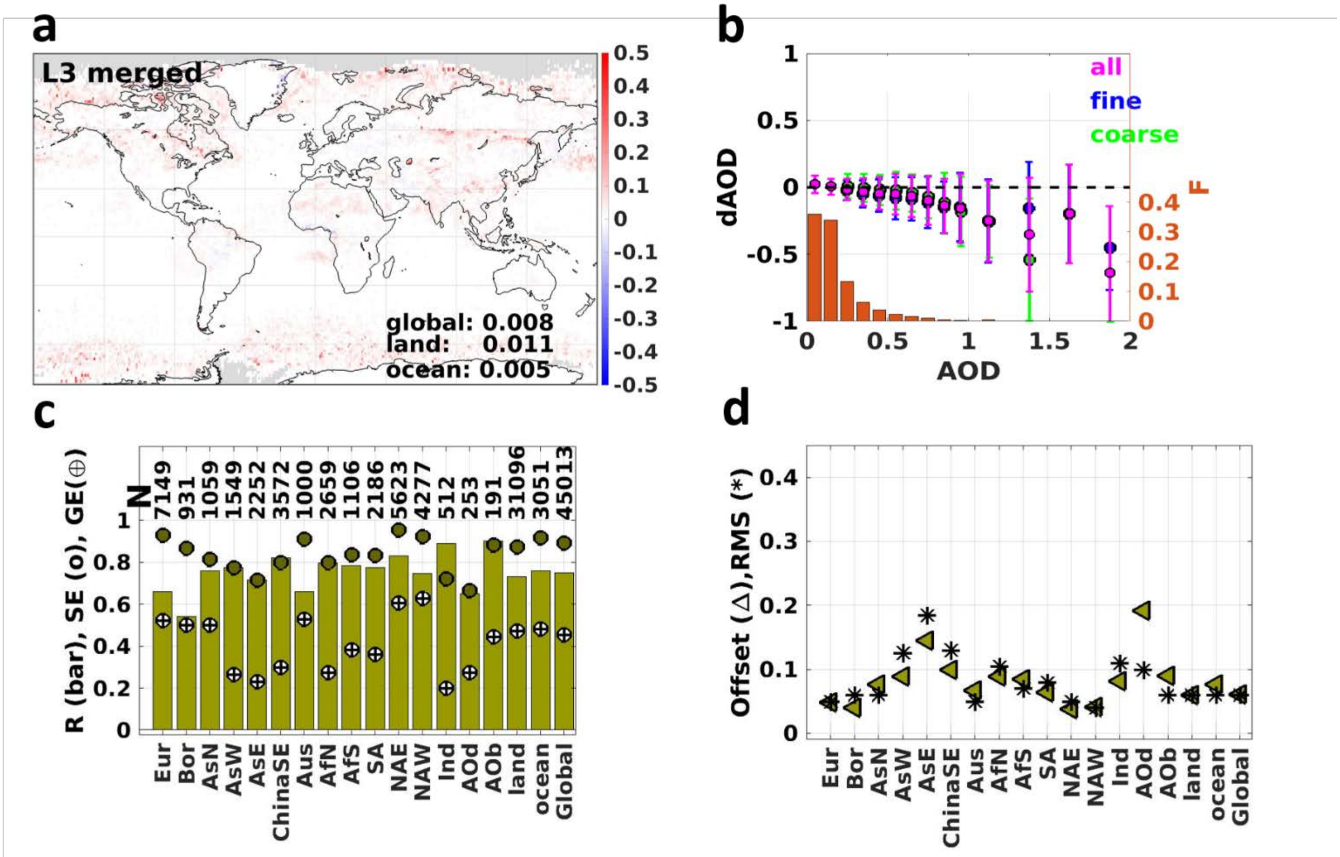
5.2 Final merged product: evaluation and inter-comparison with individual products

The agreement of the RM1 and RM2 approaches is encouraging, as we can conclude from the big-picture analysis (Sect. 5.1); details of the methodology do not matter much. As there is no significant difference in the evaluation results for products merged with approaches 1 and 2, we choose the RM2 approach for “all” aerosol types as the main merged product. We use RM2 for further inter-comparison with individual products to reveal the regional and seasonal differences between the products. If not specifically stated, the merged product mentioned below is the one obtained with RM2 for “all” aerosol types (RM2, “all”).

5.2.1 Summarised evaluation results

The difference between the L3 merged product and the median of all individual products used for merging (Table 2)The L3 merged AOD difference from the median in 2008 was calculated for year 2008 (Fig. 119a, upper, as Fig. S1 for individual products). The isdifference is close to 0 over ocean (except for near the poles regions) and within GCOS requirements over both land and ocean (0.009 and 0.007, respectively) and globally (0.011 and 0.008, respectively). High latitudes contribute most to the positive bias over oceans, whereas a positive bias is observed over land mostly over bright surfaces.

The statistics used to evaluation statistics fore the L3 merged product against AERONET (Fig 11, lower, as in Figs. 5 and A2) extracted from Figs. S12 and S13 are combined in Fig. 9 (b, c, d) for all 15 regions, as well as for land, ocean and globally., show that for For most regions, R is between 0.75 and 0.85, 80%-90% fall within the SE, 20%-60% fall within the GE, and the RMSE and offset are between 0.05 and 0.1, though somewhat higher for the regions with potentially high AOD loading (Indonesia and AOd, up to 0.15-0.2 for AsW and AsE). Statistics for the merged product (M) are also shown in Figs 3 and S7, for comparisons with individual products.


a
L3 merged
global: 0.008
land: 0.011
ocean: 0.005
b
all
fine
coarse
dAOD
AOD
F
c
N
R (bar), SE (o), GE(⊕)
7149
931
1059
1549
2252
3572
1000
2659
1106
2186
5623
4277
512
253
191
31096
3051
45013
Eur
Bor
AsN
AsW
AsE
ChinaSE
Aus
AfN
AfS
SA
NAE
NAW
Ind
AOd
AOb
land
ocean
Global
d
Offset (△),RMS (*)

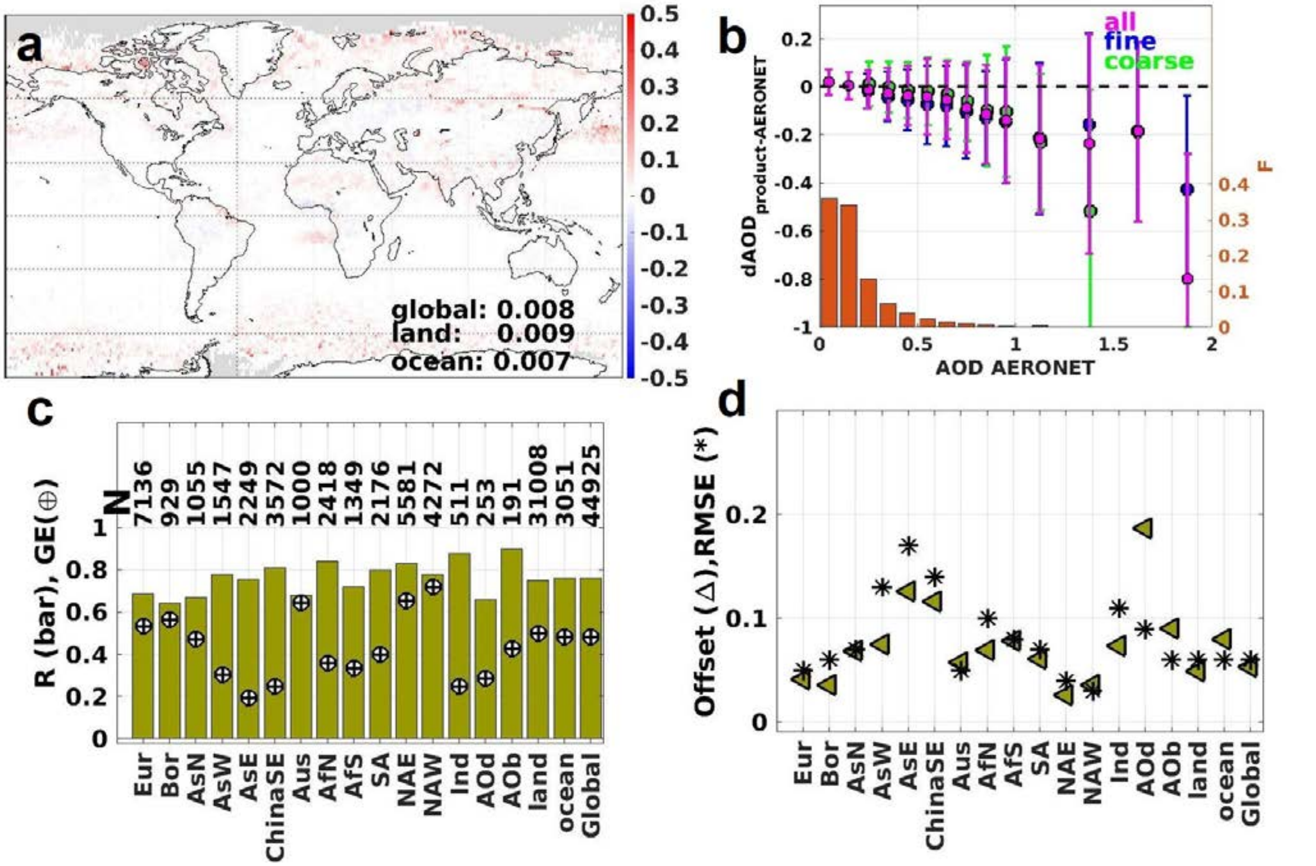


Figure ~~11~~9. L3 merged (approach ~~3~~2, RM2, ~~all aerosol types~~"all") AOD product deviation from the annual median AOD calculated from individual products used for merging (Table 2) for year 2008 (as Fig. ~~2~~ S1 for ~~other~~ individual products), a; ~~. b:~~ L3 monthly merged AOD product evaluation with AERONET~~-~~: binned AOD bias for "all" (purple), "background" (AOD<0.2, purple), "fine-dominated" (blue) and "coarse-dominated" (green) aerosol types, b; regional statistics (c: correlation coefficient R, bar; fraction of pixels ~~in the Error Envelope~~that fulfil the GCOS requirements, ~~EE~~GE, circle; d: Offset, Δ; ~~root mean square~~RMSE, ~~star~~*).

5.2.2 Uncertainties

Standard uncertainties (*unc*, meaning 1-σ of the uncertainty distribution) for the merged L3 products (monthly, seasonal, and annual) were estimated as the root mean squared sum of the deviations between the chosen merged product M (RM2,"all"), the median from the all uncorrected products (approach 1) and each of the other seven merged products (approach 2, RM1 for "all", RM1 and RM2 for "background", "fine-" and "coarse-dominated" particles).

$$unc = \sqrt{\frac{1}{N}\sum_{1}^{N}(m_i - M)^2}$$

where m_i is AOD from alternative merged product i, M is AOD from the chosen merged product (RM2, "all"), and N is the number of the alternative merged products.

Seasonal and annual uncertainties for year 2008 are shown in Fig.10. These uncertainties show artefacts at regional boundaries because the merging was done according to regional statistics.

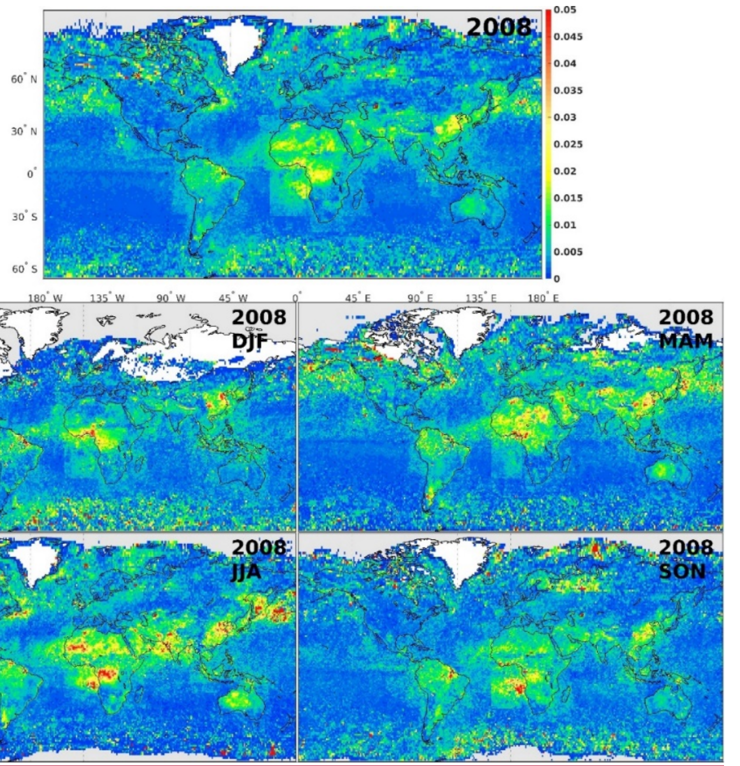


Figure 10. Seasonal and annual uncertainties between the L3 merged product (M, approach 2: RM2, "all") and other L3 merged products calculated with the approaches 1 and 2 for year 2008.

The estimated annual and seasonal uncertainties are low, 0.005-0.006 globally. They show seasonal dependence, reaching 0.008 and 0.009 on average over land in MAM and JJA, respectively. The uncertainties are larger (0.01-0.03, on average, up to 0.05) in regions with high AOD (e.g., ChinaSE, India (in JJA), AfN (in MAM, JJA), AfS (in JJA, SON)). This means that the merged dataset uncertainties fulfil the requirements calculated by Chylek et al. (2003) for AOD uncertainty of 0.015 over land and 0.010 over ocean, in order to estimate the direct aerosol radiative effect to within 0.5 W m^{-2}.The fact that this merging uncertainty estimate is smaller than the previously-discussed GCOS goal uncertainties implies that reasonable merging method decisions may be of secondary importance in terms of meeting those goals. It is cautioned, though, that since many of the algorithms are susceptible to the same error sources and subject to similar sampling limitations, the uncertainty estimates calculated here are likely to be a lower bound on the true uncertainty of the merged data sets. And it should be remembered that these uncertainties cover only the aspect of choosing the merging method, but not the entire uncertainties of the merged datasets versus AERONET.

5.2.3. Spatial and temporal inter-comparison with other products

The deviation between individual products and the merged product for year 2008 is shown in Fig. 11. Among the products used for merging, POLDER has highest positive offset (0.026) and SeaWiFS the highest negative offset (-0.026), from the merged AOD global average. Over land, POLDER has the highest positive offset (0.031); the offsets for ATSR SU and Terra DT&DB are also high (0.024 and 0.023, respectively). Highest negative offsets relative to the merged product are for MAIAC (-0.046 and -0.041 for Terra and Aqua, respectively). Over ocean, POLDER, Terra DT&DB and ATSR ADV are offset high by 0.022-0.024, whereas ATSR SU and SeaWiFS are offset low (-0.030 and -0.027, respectively) compared to the merged AOD product. Most of the observed global, land and ocean AOD offsets (except for Aqua MAIAC over land) are within the GCOS requirement of ±0.03. VIIRS agrees best with the merged product globally (0.003) and over ocean (-0.003); AVHRR DT/SOAR and AQUA DR&DB agree best with the merged product over land, showing opposite in sign offsets of -0.011 and 0.009, respectively. Regional biases between the individual products and the merged product exist. These are similar to regional biases shown in Fig. 2, where the individual products were compared with median AOD calculated from all individual products available at 0.55 µm.

Regional annual offsets between individual AOD products and the merged AOD product are shown in Fig. S14 (similar regional annual offsets between individual AOD products and the median AOD product, as shown in Fig. 6 and S10). For AsE, which includes ChinaSE, and AfN, the AOD offset is higher than 0.03 (GCOS requirements) for some products. However, those areas are characterised with high AOD loading (annual AOD is between 0.4 and 0.8) related to e.g., anthropogenic pollutions and/or dust events. If the GCOS requirement of 10% of AOD is also applied here, then most of the offsets are within the GCOS requirements. The highest regional offsets relative to the merged AOD dataset are associated with products which provide AOD at other than 0.55 µm wavelength – TOMS (0.50 µm), OMI (0.50 µm) and EPIC (0.44 µm) – and thus are not used for merging.

In some regions, AOD offsets between individual products and the merged product show seasonal behaviour (Fig. S15). In ChinaSE, the negative offsets for AVHRR NOAA, SeaWiFS and VIIRS are most pronounced in JJA. In AsW, the ATSR ADV positive offset is higher for that season. In AfN, most products have their highest negative offsets in JJA, whereas ATSR SU and ATSR ensemble (which includes the ATSR SU product) have their highest positive biases. In SA, offsets are lower in JJA for all products. In AOb offsets are lower in MAM, and in AOd offsets are lower in SON for all products.

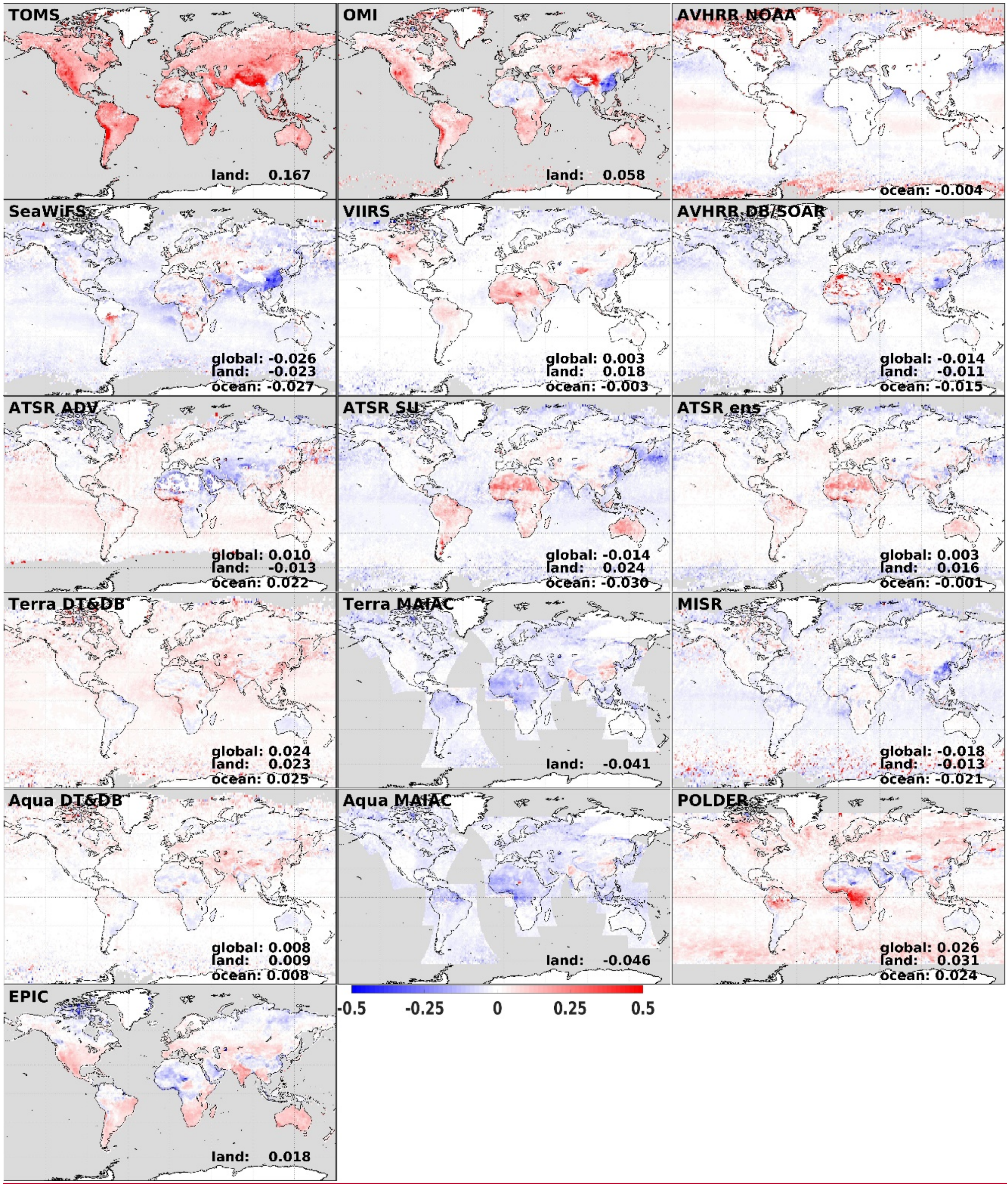


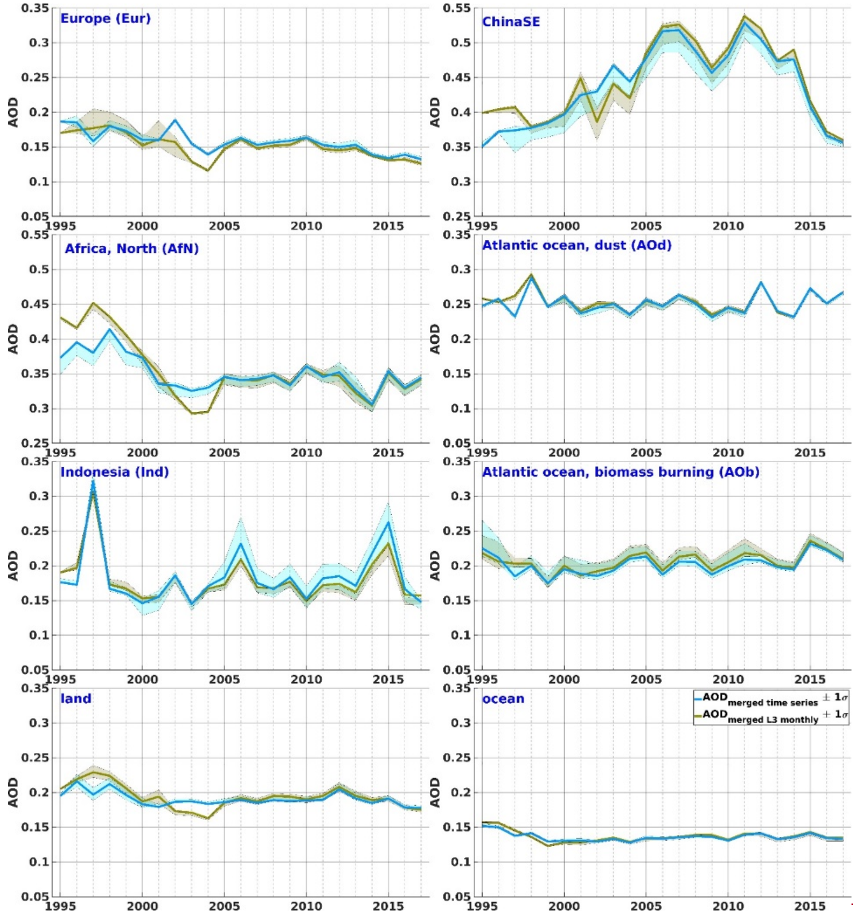
Figure 11. AOD deviation of the individual products relative to the merged AOD product for year 2008. Global, land and ocean AOD mean differences are shown for each product, when available.

6.2.3 AOD time series merged based on ranking results

The AOD time series produced by merging all available products, assuming the ranking for all types of aerosol particles, are shown in Fig. 12 and Fig. A7 together with the time series calculated from the L3 merged aerosol product (Sect. 6.2.2). Overall low (1-3% of AOD) 1-σ AOD standard deviation shows agreement with the time series merged for all four classes, which means that aerosol type is not critical for merging. The deviation between time series for different aerosol types increases slightly in the regions with the potentially high AOD loading. For some regions (e.g., Europe, AfN, Asia), the deviation is higher for the period when MODIS-family products are not yet available, i.e. before 2002, which agrees well with the AOD diversity for years 2000, 2008 and 2017 for all available satellite products (Fig. 3).

An attempt was performed to explain the offset in time series (AOD from the merged time series minus merged L3 AOD time series) obtained with different methods (Fig. 12) by looking at the AOD standard deviation between the available products for particular years and at the number of products. However, no clear dependence was found, except that the negative offset above GCOS requirements (between -0.03 and -0.08) was observed before 2002 and positive offset (between 0.03 and 0.04) was observed mostly in 2002-2004 in some of the areas, as indicated in Fig. A9.

Thus, there is general consistency and similar temporal patterns between the two approaches (the merged time series and the timeseries from merged L3 AOD product), despite small differences, which are more pronounced at the beginning of the period, when less products are available.



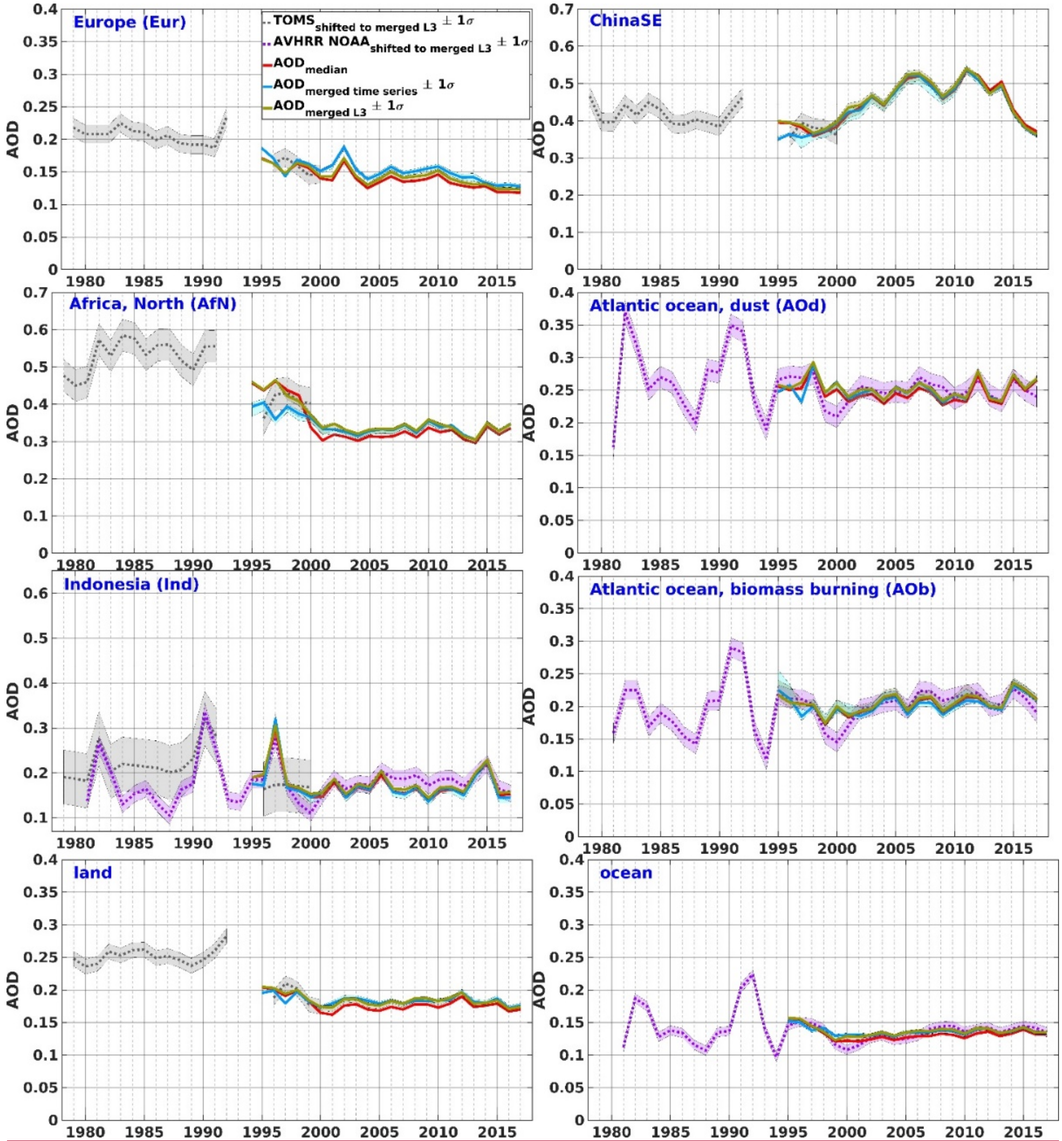
~~**Figure 12. Time series of merged (approach 3, all aerosol types) annual time series AOD (blue) and annual AOD time series calculated from L3 merged (with approach 3, for all aerosol types) AOD product (olive) for the selected regions. Deviations between time series calculated with different statistics for all, background, fine-dominated and coarse-dominated aerosol are shown as ±1σ. For other regions, see Fig. A8.**~~

~~7~~ 6 ~~Regional and global annual, seasonal and monthly~~Merged AOD ~~merged~~ time series

As the L3 AOD merged products (Sect.5), the AOD time series from the individual products (Figs. 4 and S8) were merged, using approach 1 (median for uncorrected AOD) and approach 2 (RM1 and RM2 for different aerosol types).~~We also show the ±1σ of AOD from all available products.~~ The shifted AOD median (approach 1 for shifted products) has clear limitations when the product chosen as a reference (Terra DT&DB, in our case) deviates considerably from other products over most of the regions (except for Aus, AfN and Sa, Fig. S8). Thus, the median for shifted products is not discussed here. However, the median shifted AOD approach allows extending the time series back to 1978-1994, where only the TOMS AOD (over land) and AVHRR NOAA (over ocean) long-term products currently exist and thus the merging approaches introduced in the current study are not applicable.

The two merging approaches (approach 1 for uncorrected products and approach 2 for weighted AOD) tested here agree well (Fig.12). The offsets between time series calculated with different approaches are low (0.004-0.011), which shows that different merging approaches provide similar results. Spatial consistency is indicated by high correlation (similar positions of peaks) in AfN and its Atlantic dust outflow region. Interannual variation as well as the standard deviations are highest for regions with the largest AOD, e.g., over ChinaSE (anthropogenic emissions) and Indonesia (biomass burning). The time series of ChinaSE follows the known patterns caused by step-wise regional emission reductions in the last 25 years (Sogacheva et al., 2018b). AOD time series merged with different approached show a good agreement for all time scales: annual (Fig.12), seasonal and monthly (Fig.13a and 13b, respectively, for Europe and China and Figs. S16 and S17 for all studied regions).

The offsets between the merged time series and time series calculated from the merged L3 product has regional component and, as, discussed above, depends on the availability of the products (Table 2). The offsets between the time series merged with different approaches (Table 3) are slightly higher for all regions for the periods 1995-1999 and 2012-2017, when fewer products are available for merging (Table 2). The deviation up to 0.05 ($AOD_{approach1}$ >$AOD_{approach2}$) is observed over Indonesia and North America before 2002, when both MODIS satellites become operational. For other regions, the deviation is considerably lower (below 0.03). By adding MISR and both MODIS products in 2000/2002, the offset between the time series is reduced. ATSR products are not available starting from 2012, when the VIIRS product becomes available. In 1995-1999, the mean offset is similar for all three time series. The offsets are higher for regions with high AOD loading (e.g., Asia and Northern Africa, Fig. S18). In 2000-2011 and 2012-2017, the offset is lowest (0.004) between the merged and the median time series, as well as between the merged time series and the time series calculated from the merged L3 product. The agreement in the time series obtained with different approaches supports the conclusion made based on the evaluation results, that for the big-picture analysis of overall trends, details of the methodology do not matter very much.


Europe (Eur)
TOMS shifted to merged L3 ± 1σ
AVHRR NOAA shifted to merged L3 ± 1σ
AOD median
AOD merged time series ± 1σ
AOD merged L3 ± 1σ
ChinaSE
Africa, North (AfN)
Atlantic ocean, dust (AOd)
Indonesia (Ind)
Atlantic ocean, biomass burning (AOb)
land
ocean
AOD

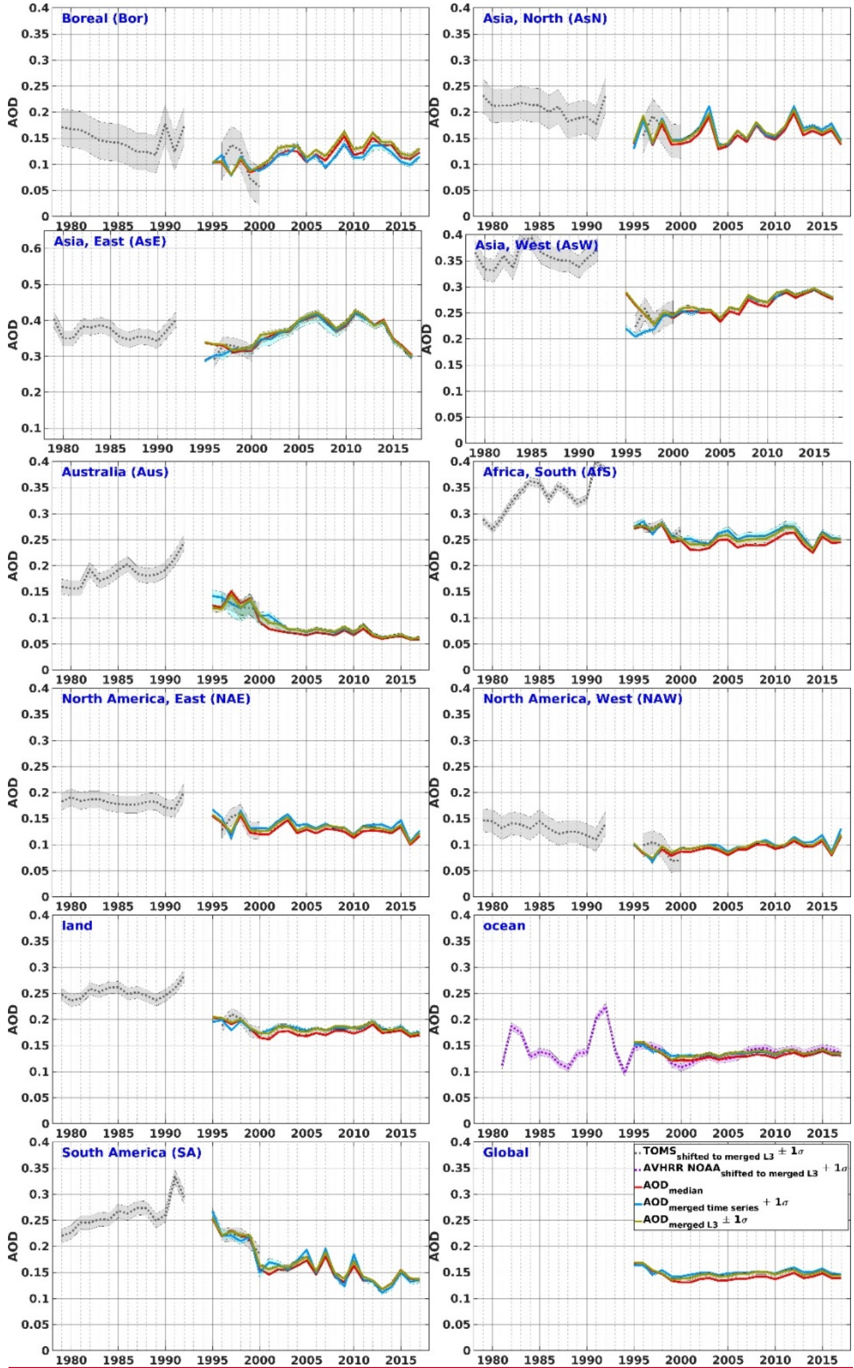


Figure 12 Annual AOD time series merged with two different approaches (red, light blue for approaches 1 and 2, respectively) and AOD time series from the L3 merged data (approach 2, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow (often small, thus not visible). TOMS over land and AVHRR NOAA over ocean products shifted to the merged time series are also shown (grey and purple dashed lines, respectively), when available.

Table 3. Mean offset and standard deviation (in parentheses) between time series obtained with different approaches for three time periods, determined based on products availability.

	1995-1999	2000-2011	2012-2017
time series from merged L3 to median time series	0.009 (0.009)	0.007 (0.005)	0.011 (0.006)
merged time series to median time series	0.011 (0.010)	0.004 (0.002)	0.009 (0.006)
time series from merged L3 to merged time series	0.010 (0.014)	0.004 (0.004)	0.004 (0.004)

Annual, seasonal and monthly time series from the merged L3 monthly AOD show slightly higher deviation of both signs, compared to the merged time series discussed above. Interestingly, the seasonality is observed in the deviation. In AfN, the AOD from the monthly merged L3 is higher in autumn for the period of 1995-1999. In Bor and AsN (Figs. S16 and S17), the deviation is higher in spring for the period of 1997-1999. The possible explanation might be the sparser coverage in those areas (due to restrictions in retrieval algorithms to retrieve bright surfaces, e.g., desert or snow). Regional offsets between the annual, seasonal and monthly AOD merged time series and the time series from the merged L3 monthly product are summarised for three time scales in Fig. S19. The offset is lower for annual data and generally increases with the time resolution. As the previous analysis showed, the offset is bigger in high AOD regions (e.g., Asia, AfN and SA).

Overall, good agreement exists in current study among the time series calculated using different merging approaches and different orders of the processing steps. There is a general consistency and similar temporal patterns are observed between the time series merged with two approaches and the timeseries from merged L3 AOD product, despite small differences, which are more pronounced at the beginning of the period, when less products are available. With only few exceptions, the offsets between the AOD time series calculated with different approaches are within the GCOS requirement of ±0.03 or 10% of AOD. A separate paper is planned, where the merged AOD L3 product will be analysed in order to reveal regional and global AOD trends for the period from 1995 to 2017.

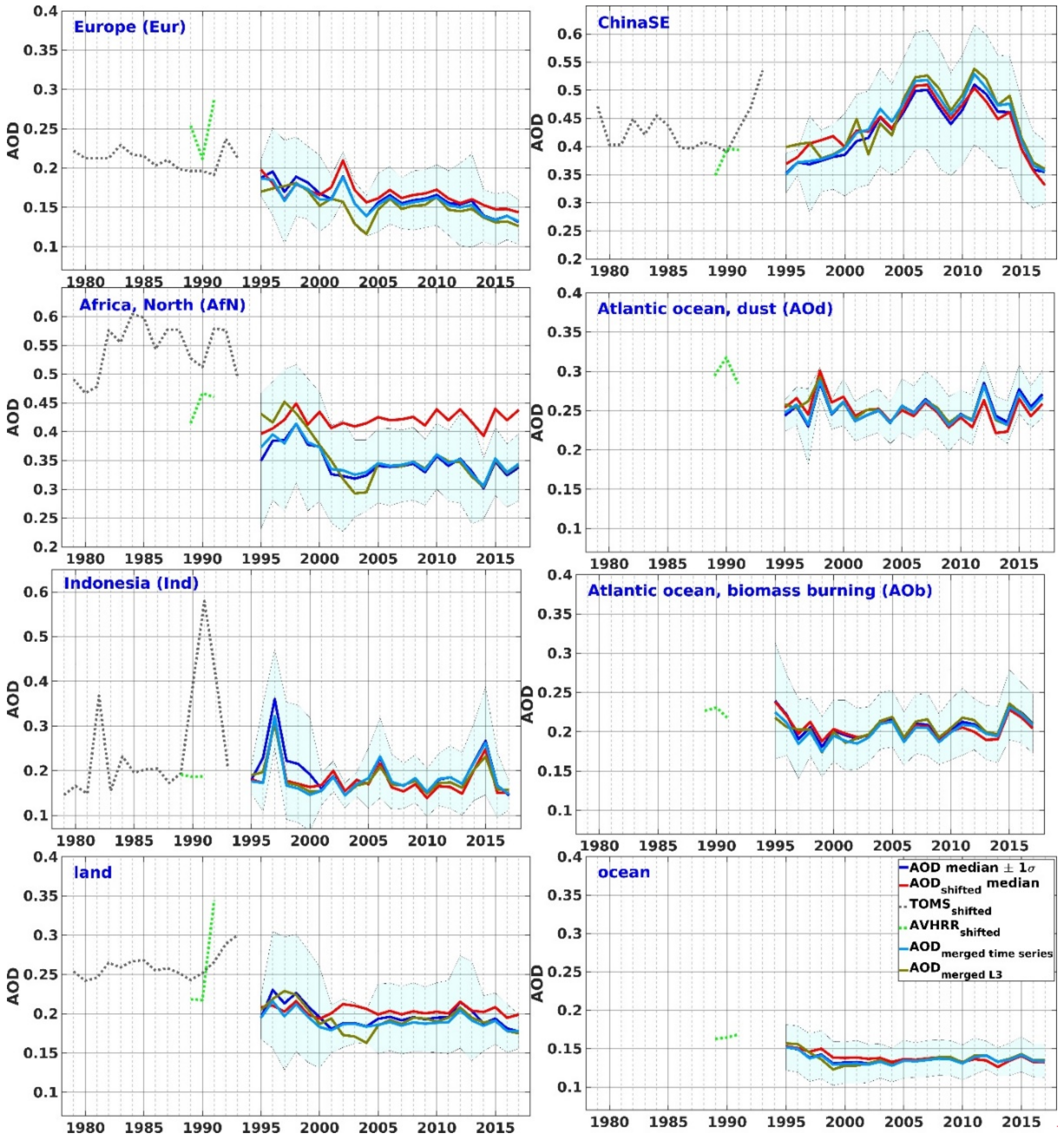
In Figs. 13-15 and Figs. A10-A12, we show the AOD time series obtained with different approaches – median for uncorrected AOD from 15 available products (approach 1), median for shifted to the ATSR ensemble AOD (approach 2), weighted time series (approach 3, all aerosol types) and time series from weighted AOD for annual, seasonal and monthly products, respectively. We also show the ±1σ of AOD from all available products.

Median of the uncorrected AOD (approach 1) and merged (approach 3) AOD time series show the best agreement for all time scales (annual, seasonal, monthly). The deviation up to 0.05 ($AOD_{approach1}$->$AOD_{approach3}$) exists for Indonesia and North America before 2002, when both MODIS satellites become operational. For other regions, the deviation is considerably lower (below 0.03). Median of shifted AOD (approach 2) has its clear limitations when the product chosen as a reference (ATSR ensemble, in our case) deviated considerably from other products in certain regions (e.g., AfN). In that case, offset exists between the time series merged with other approaches (1 and 3). However, this approach allows extending time series back to 1978-1994, where only the TOMS AOD long term product currently exists and thus other merging approached are not applicable for that period.

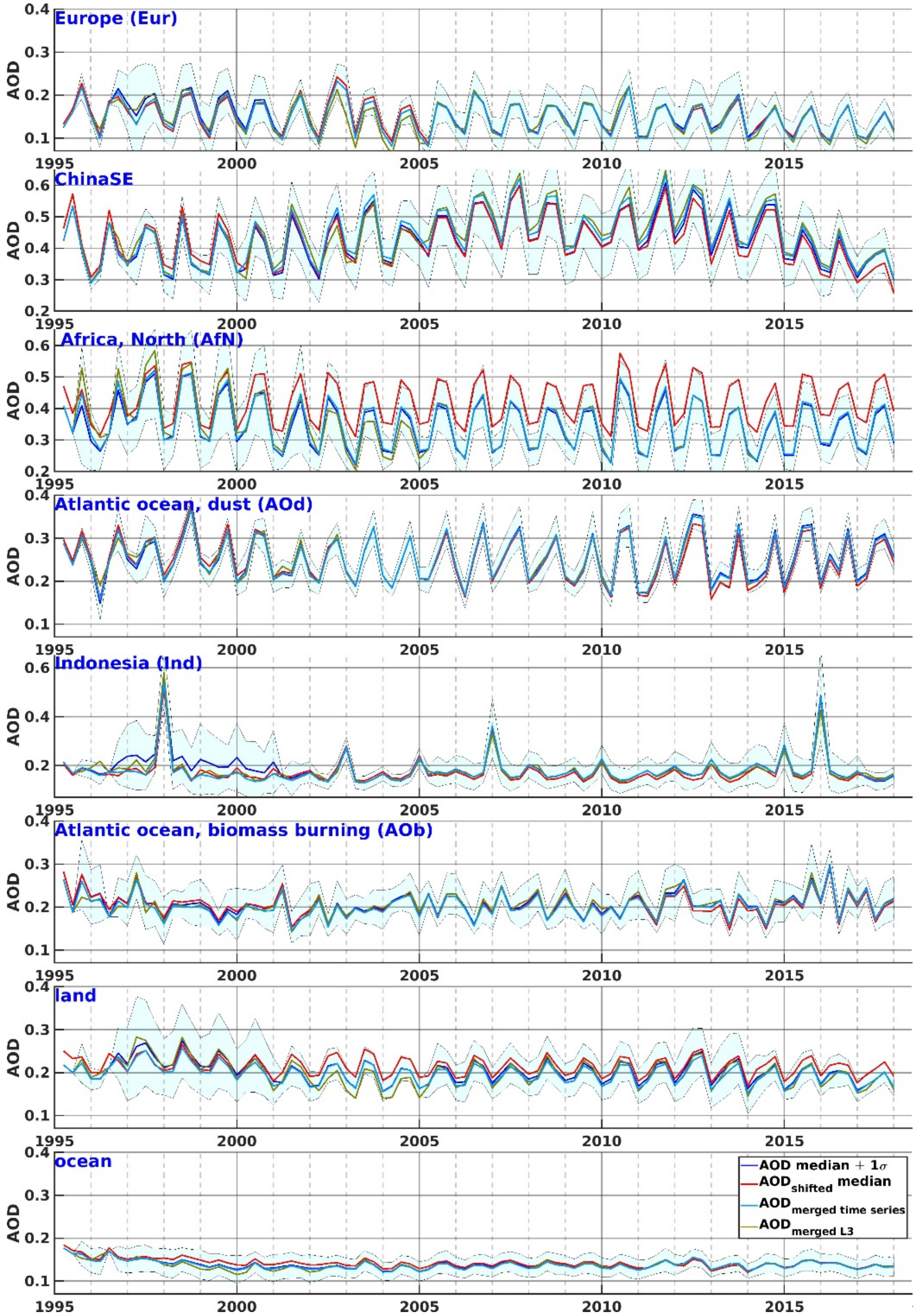
Annual, seasonal and monthly time series from the merged (approach 3) L3 monthly AOD show a bit higher both signs deviation from time series discussed above. Interestingly, the seasonality is observed in the deviation. In AfN, the AOD from the monthly merged L3 is higher in autumn for the period of 1995-1999. In Bor and AsN (Figs. A10-A12), the deviation is higher in spring for the period of 1997-1999. The possible explanation might be due to the sparser coverage in those areas (due to restrictions in retrieval algorithms to retrieve bright surfaces, e.g., desert or snow).

Overall, good agreement exists among the time series calculated using different approaches. With only few percent exception, similar temporal patterns are reproduced and the offset between the AOD time series calculated with different approaches is within the GCOS requirement of 0.03 AOD.

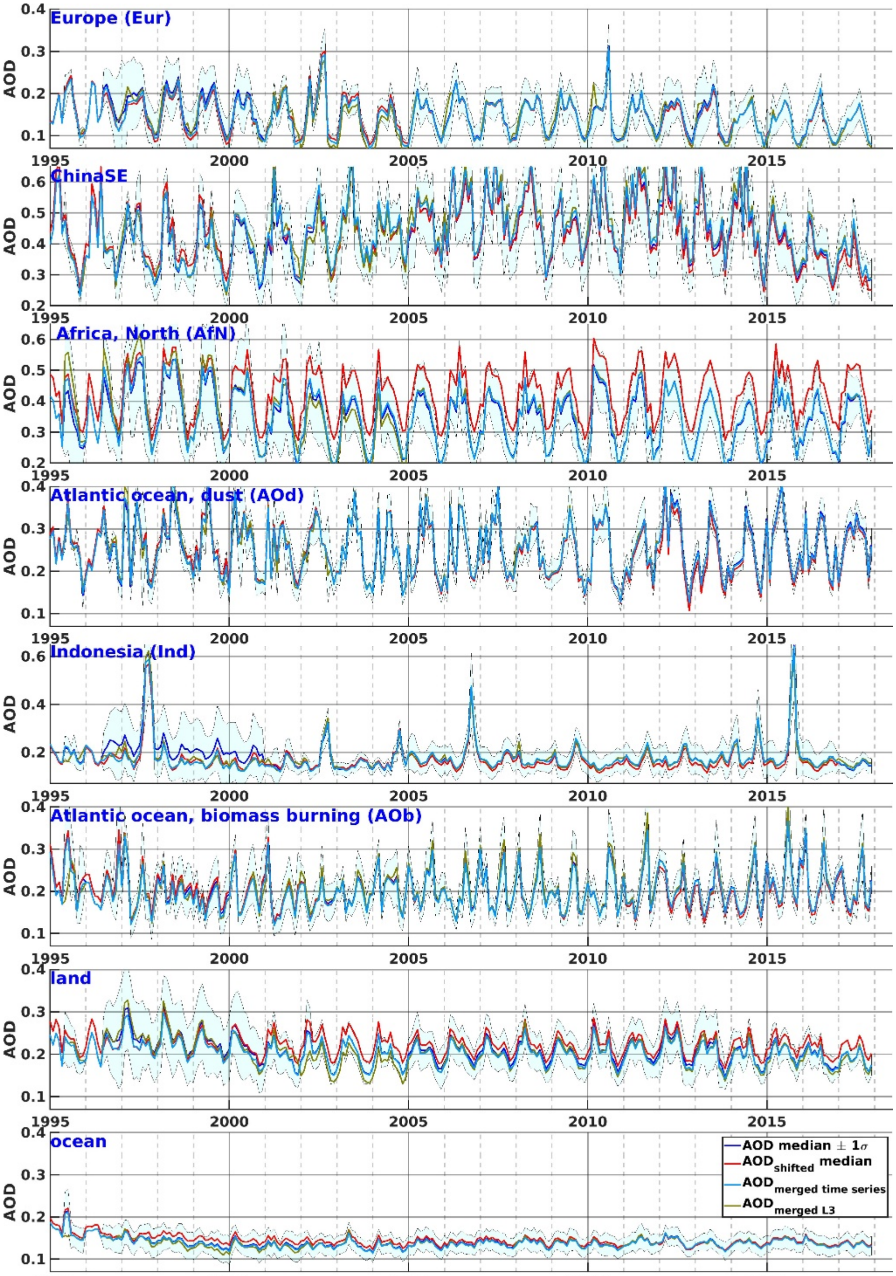
There are of course caveats to these rather simple and straightforward merging approaches, which do not consider in much detail the differences in sampling and sensitivity to different conditions (e.g. surface brightness, number of independent observables) of the different instruments and algorithms. It is well known that monthly, seasonal or annual gridded mean values carry large uncertainties, whether inferred from a few ground-based stations meant to represent a full grid cell, or from satellite images containing large gaps due to limited swath, clouds or failed retrievals. Pixel-level uncertainties are becoming available for a growing number of satellite products, and it would be highly beneficial if these estimated errors could be propagated consistently to those gridded monthly products. However, this requires deeper insight and new methods to take into account correlation patterns among parts of the uncertainties, and to estimate practically the sampling-based uncertainties in light of approximated AOD variability. Altogether, as frequently requested from a user point of view, the stability and consistency of regional, merged AOD time series can be seen as strengthening our confidence in the reliability of satellite-based data records.



~~**Figure 13. Annual AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. For other regions, see Fig. A10.**~~



~~**Figure 14. Seasonal AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. For other regions, see Fig. A11.**~~


Europe (Eur)
ChinaSE
Africa, North (AfN)
Atlantic ocean, dust (AOd)
Indonesia (Ind)
Atlantic ocean, biomass burning (AOb)
land
ocean
AOD
1995
2000
2005
2010
2015
AOD median ± 1σ
AODshifted median
AODmerged time series
AODmerged L3

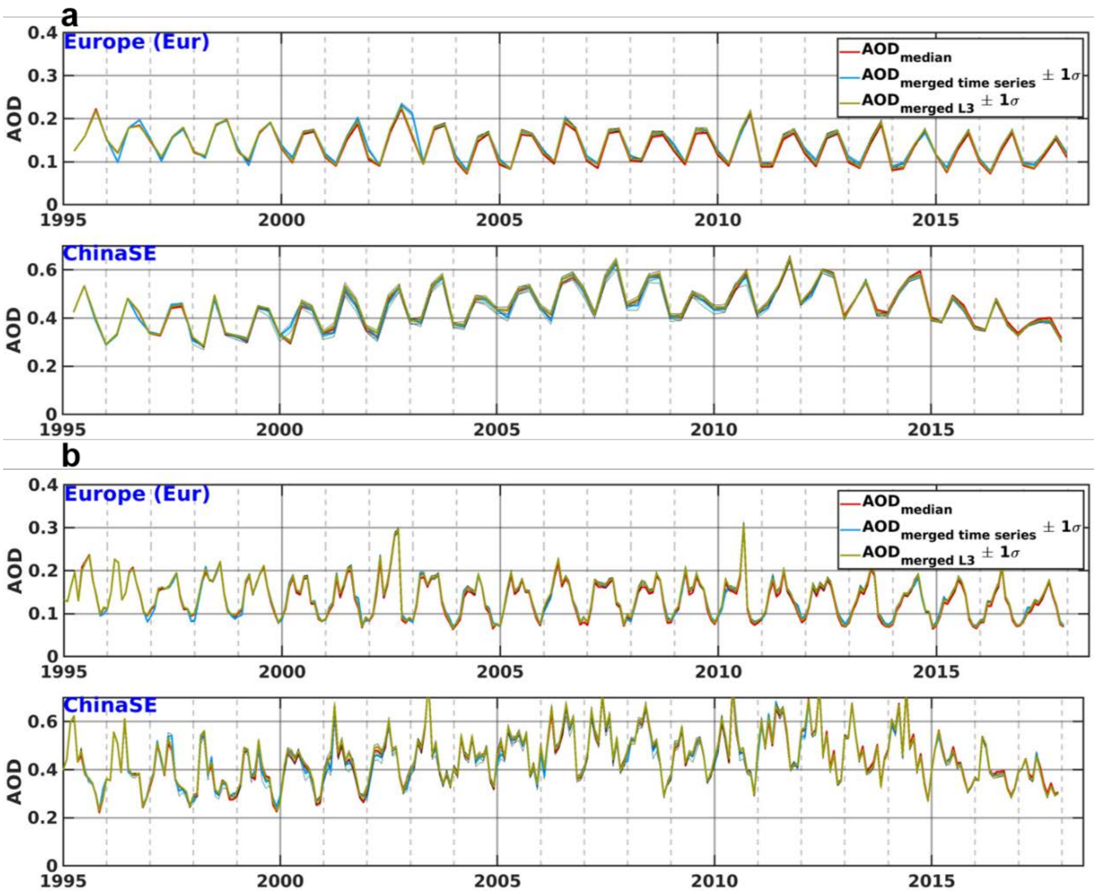


Figure ~~15~~13. Seasonal (a) and monthly (b) AOD median time series (red), merged time series (blue) and time series from the merged L3 product (olive) for Europe and ChinaSE. AOD ±1σ for the merged time series and for the time series from the merged L3 products are shown as light blue and light olive shadows, respectively. Note the different scale. For all selected regions, see Figs. S16 and S17.~~Monthly AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. Note the different scale for ChinaSE, AfN and Ind. For other regions, see Fig. A12.~~

~~8~~ 7 Conclusions

This study has analysed the consistency of regional time records of monthly AOD from ~~15~~ 16 different satellite products. These were obtained from a wide range of different instruments – TOMS, AVHRR, SeaWiFS, ATSR-2, AATSR, MODIS, MISR, POLDER, VIIRS and EPIC - with largely varying information content and sampling, and with different algorithms based on different remote sensing approaches, quality filtering, cloud masking and averaging.

Differences between those ~~15~~ 16 data records in a set of regions with different characteristics across the globe were demonstrated and verified against a ground-based AERONET monthly mean ~~gridded~~ dataset in order to answer the question how well a satellite dataset can reproduce monthly gridded mean AERONET values in a region.

Regional AOD time series (monthly, seasonal, annual) from ~~15~~ 16 different products (with different algorithms, measurement principles, number of independent observables, sampling) show good consistency of temporal patterns but significant biases due to all those differences. In many cases the more pronounced differences were between different algorithms applied to the same sensor, rather than between similar algorithms applied to different sensors. This is encouraging in that it implies that algorithmic uncertainties (either retrieval assumptions or pixel selection criteria) can be similar to or larger than sensor ones (e.g., calibration quality and sampling limitations), and as such refining individual algorithms can still make meaningful steps towards providing better L3 products.

To build ~~the~~ an AOD ~~data set~~product merged from ~~15~~ 12 ~~different~~ individual satellite products, ~~three~~ two different ~~merging~~ approaches were introduced and tested. ~~First~~In approach 1, a simple median of the ~~15~~ 12 uncorrected and shifted to Terra DT&DR product time records was conducted. ~~Second~~In approach 2, ~~each record was shifted with a constant offset to one chosen record before the median was calculated. Third,~~ the AOD evaluation results (for different aerosol types) against AERONET ~~analysis~~ w~~as~~ere used to infer a ranking which was then used to calculate a weighted AOD mean. Two different ranking methods, RM1, simple ranking based on better statistics, and RM2, ranking based on binned statistics, were tested in approach 2. In addition,~~The third approach~~ the order of the processing steps in approach 2 was ~~applied to the AOD time series and to L3 monthly AOD product, later used to calculate regional time series~~interchanged (L3 dataset merging or regional merging) to test the stability of the results.

Ten merged L3 AOD monthly products were created and evaluated with AERONET. The evaluation ~~The AOD L3 monthly merged product (1995-2017) was developed and evaluated against the AERONET in the current paper. Evaluation results~~ show~~ed~~s that the quality of~~-~~ the merged products (except for one created with the approach 1 for shifted AOD) is as good as that of the most highly ranked individual AOD products in each region. One of the merged products (approach 2, RM2, "all"), was chosen as a final merged product, based on slightly better evaluation results. Uncertainties for the final merged product were estimated.

All merged regional AOD time series show ~~the~~ a very high consistency of temporal patterns and ~~(where expected)~~ between regions and the time records with their uncertainties (standard deviations shaded around the median values) ~~are~~ clearly ~~able to~~ illustrate the evolution of regional AOD. With few exceptions all merging ~~the three~~ methods lead to very similar results ~~(only the spread is much smaller and likely underestimated for the third approach)~~, which is reassuring for the usefulness and stability of the merged products.

There are of course caveats to these rather simple and straightforward merging approaches, which do not consider in much detail the differences in sampling and sensitivity to different conditions (e.g., surface brightness, number of independent observables) of the different instruments and algorithms. It is well known that monthly, seasonal or annual gridded mean values carry large uncertainties, whether inferred from a few ground-based stations meant to represent a full grid cell, or from satellite images containing large gaps due to limited swath, clouds or failed retrievals. Pixel-level uncertainties are becoming available for a growing number of satellite products, and it would be highly beneficial if these estimated errors could be propagated consistently to those gridded monthly products. However, this requires deeper insight and new methods to take into account correlation patterns among parts of the uncertainties, and to estimate practically the sampling-based uncertainties in light of approximated AOD variability. Altogether, as frequently requested from a user point of view, the stability and consistency of regional, merged AOD time series ~~can~~ should be seen as strengthening our confidence in the reliability of satellite-based data records. Recent, ongoing, and future work to improve the level 3 uncertainty budget of the satellite products – as well as assessment of spatiotemporal uncertainties in time-aggregated AERONET data – will benefit the creation and assessment of merged time series.

~~The AOD L3 monthly merged product (1995-2017) was developed and evaluated against the AERONET in the current paper. Evaluation results showed that the quality of the merged product is as good as that of the most highly ranked individual AOD products in each region.~~ The corresponding time series are ~~planned to be used~~for use in regional and global AOD trend analyses, and for comparison with (climate and reanalysis) model AOD fields. The records can also be extended as satellite missions continue and new data versions are released. Aside from the merged data set itself, a main outcome of this research has been a quantification of the diversity between monthly satellite AOD products and their comparability with monthly averages from AERONET, as well as the utility of a merged product for applications requiring data beyond a single instrument's lifetime, and the sensitivity of the merged time series to some sensible decisions which must be made in creating it. ~~the AOD regional and global trend analysis, as well as in the inter-comparison with the modelled AOD product from the re-analysis.~~

~~9~~ 8 Data availability

URL and doi (if available) of the products used in the current study are summarised in Table ~~2~~4.

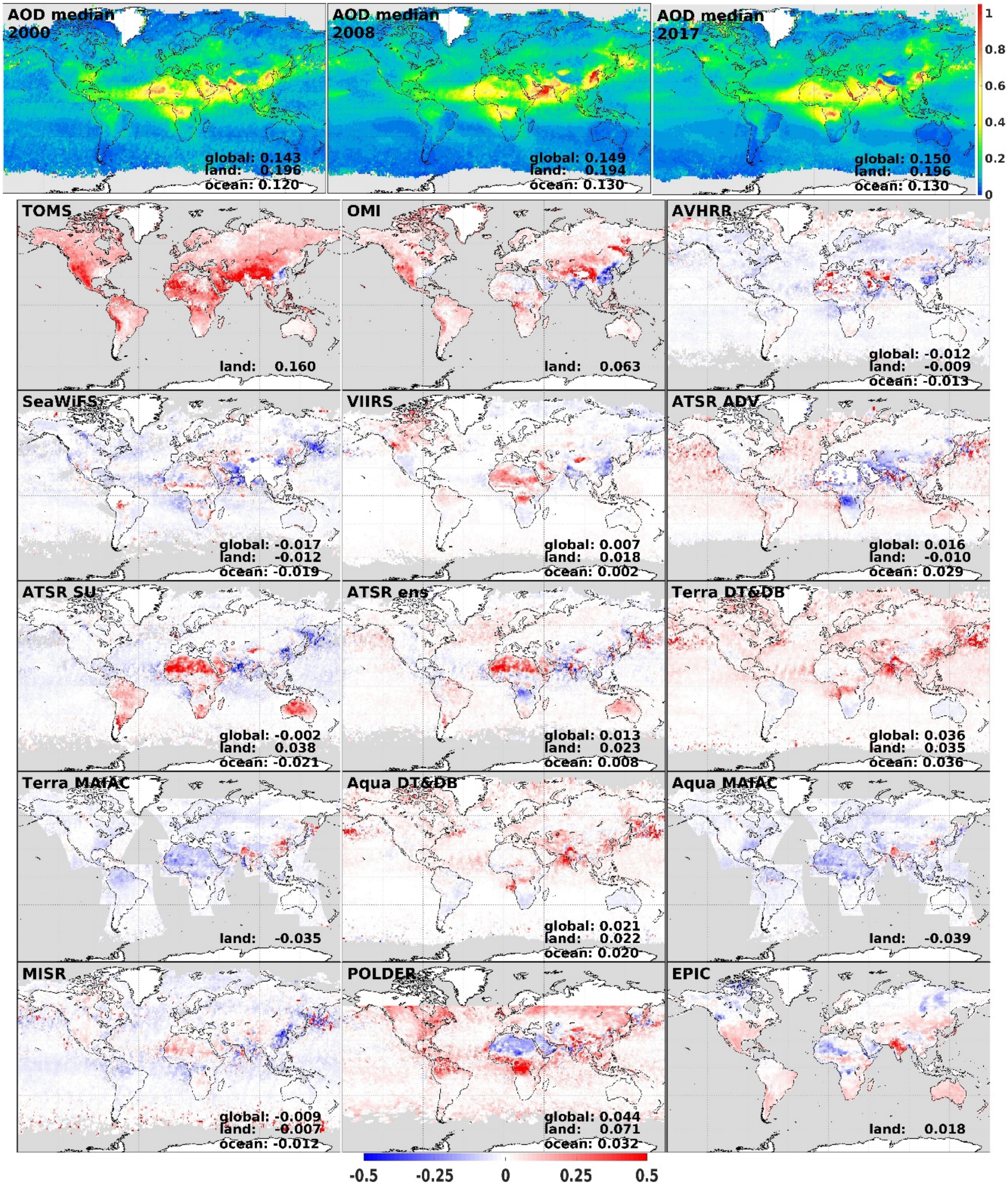
Table ~~2~~4. URL and doi (if available) of the products used in the current study.

Product	url/doi	archive	
TOMS	url	NASA's GES-DISC	https://disc.gsfc.nasa.gov/datasets?page=1&subject=Aerosols&measurement=Aerosol%20Optical%20Depth%2FThickness
OMI	url	NASA's GES-DISC	https://aura.gesdisc.eosdis.nasa.gov/data/Aura_OMI_Level2/OMAERUV.003/
	doi		10.5067/Aura/OMI/DATA2004
AVHRR	url	NOAA	https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00977
	doi		http://doi.org/10.7289/V5BZ642P

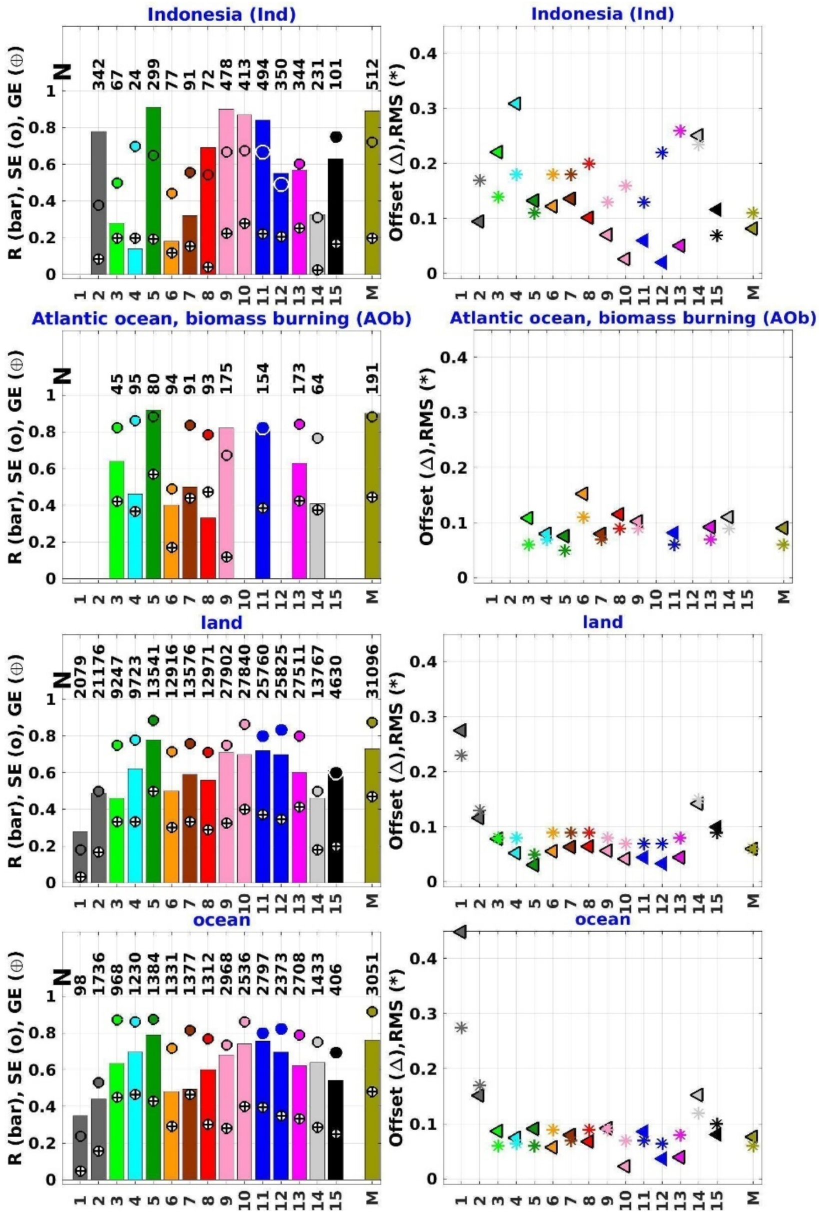
~~AVHRR~~	url	NASA NCCS	https://portal.nccs.nasa.gov/datashare/AVHRRDeepBlue
SeaWiFS	url	NASA GES DISC (via EarthData)	https://earthdata.nasa.gov/
	doi		10.5067/MEASURES/SWDB/DATA304
VIIRS	url	NASA LAADS (via EarthData)	https://earthdata.nasa.gov/
	doi		10.5067/VIIRS/AERSDB_M3_VIIRS_SNPP.001
ATSR ADV	url	ICARE	http://www.icare.univ-lille1.fr/archive
ATSR SU	url	ICARE	http://www.icare.univ-lille1.fr/archive
ATSR ensemble	url	ICARE	http://www.icare.univ-lille1.fr/archive
MODIS DT&DB *	url	NASA LAADS	https://ladsweb.modaps.eosdis.nasa.gov/
	doi		Terra: 10.5067/MODIS/MOD08_M3.061 Aqua: 10.5067/MODIS/MYD08_M3.061
MODIS MAIAC	url		https://search.earthdata.nasa.gov/search?q=MCD19&ok=MCD19
MISR	url		http://eosweb.larc.nasa.gov/project/misr/misr_table
	doi		10.5067/Terra/MISR/MIL3MAE_L3.004
POLDER	url	ICARE	https://www.grasp-open.com http://www.icare.univ-lille1.fr
EPIC	url		https://search.earthdata.nasa.gov/search?q=MCD19&ok=MCD19
AOD merged	url		will be available after the manuscript is accepted to ACP
AERONET	url		https://aeronet.gsfc.nasa.gov/

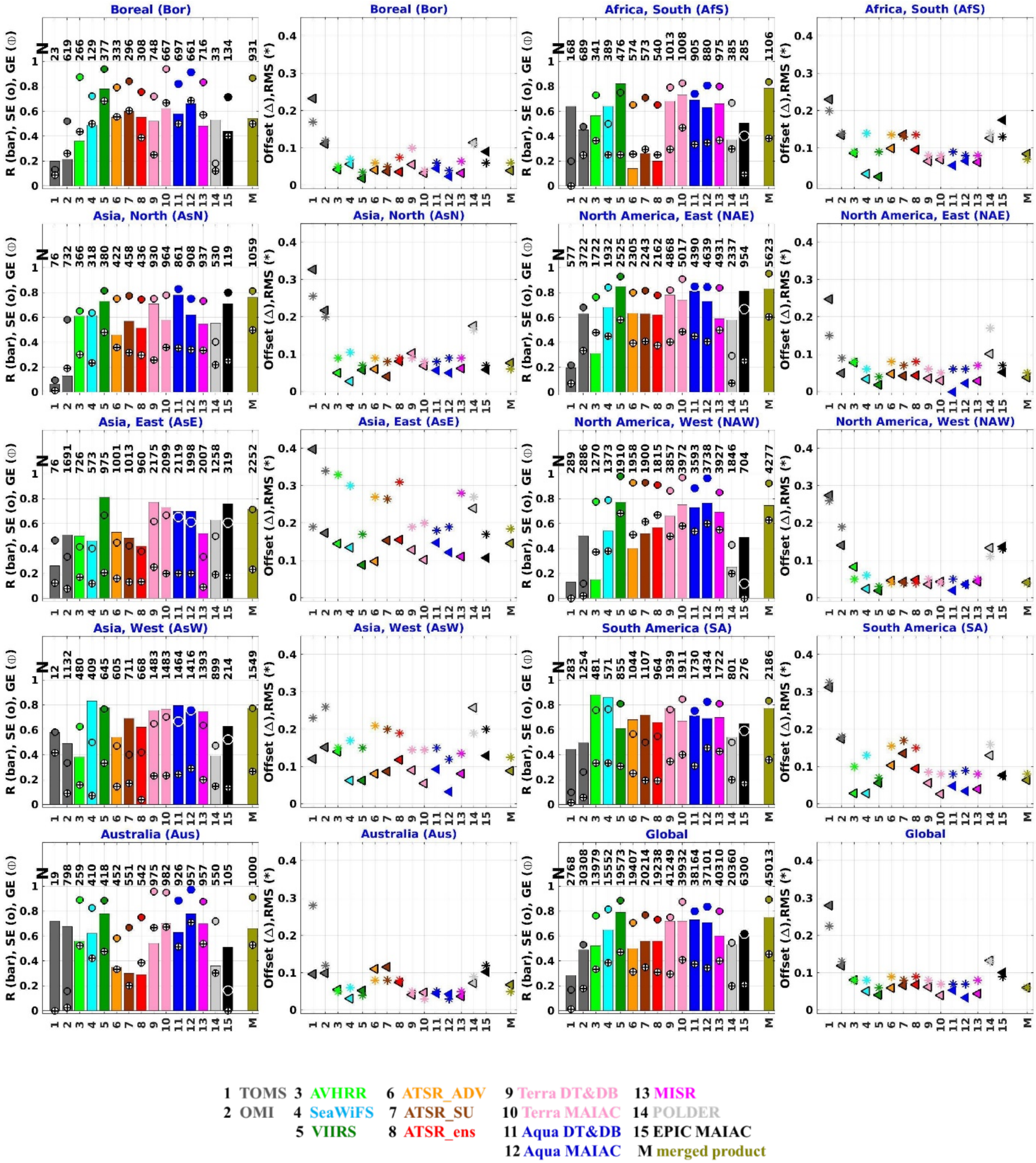
* Additional online documentation at: https://modis-atmos.gsfc.nasa.gov/, https://darktarget.gsfc.nasa.gov/, https://deepblue.gsfc.nasa.gov/

~~10 Appendix~~



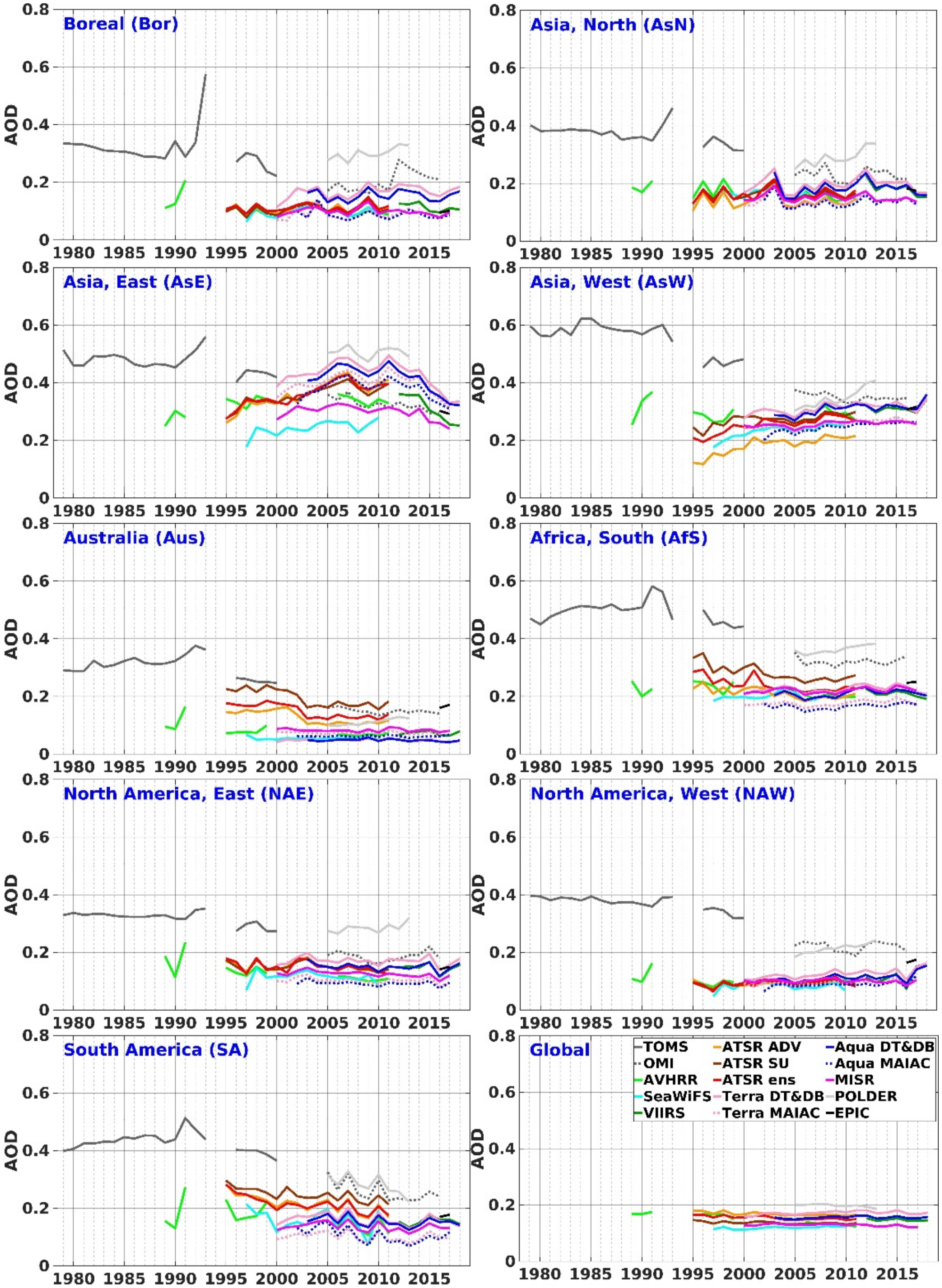
~~**Figure A1. Upper line: summer AOD median for 2000 (*), 2008 and 2017 (X), calculated from the available products. Lines 2-6: AOD deviation of the different products from the annual median AOD for years 2000 (TOMS), 2017 (VIIRS and EPIC) or 2008 (all other products). AOD anomalies with respect to the AOD median are shown on the deviation plots. Global land and ocean AOD mean differences are shown for each product, when available. For annual offset, see Fig. 2.**~~


Indonesia (Ind)
R (bar), SE (o), GE (⊕)
N
342 67 24 299 77 91 72 478 413 494 350 344 231 101 512
Offset (Δ),RMS (*)
Atlantic ocean, biomass burning (AOb)
45 95 80 94 91 93 175 154 173 64 191
land
2079 21176 9247 9723 13541 12916 13576 12971 27902 27840 25760 25825 27511 13767 4630 31096
ocean
98 1736 968 1230 1384 1331 1377 1312 2968 2536 2797 2373 2708 1433 406 3051
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 M

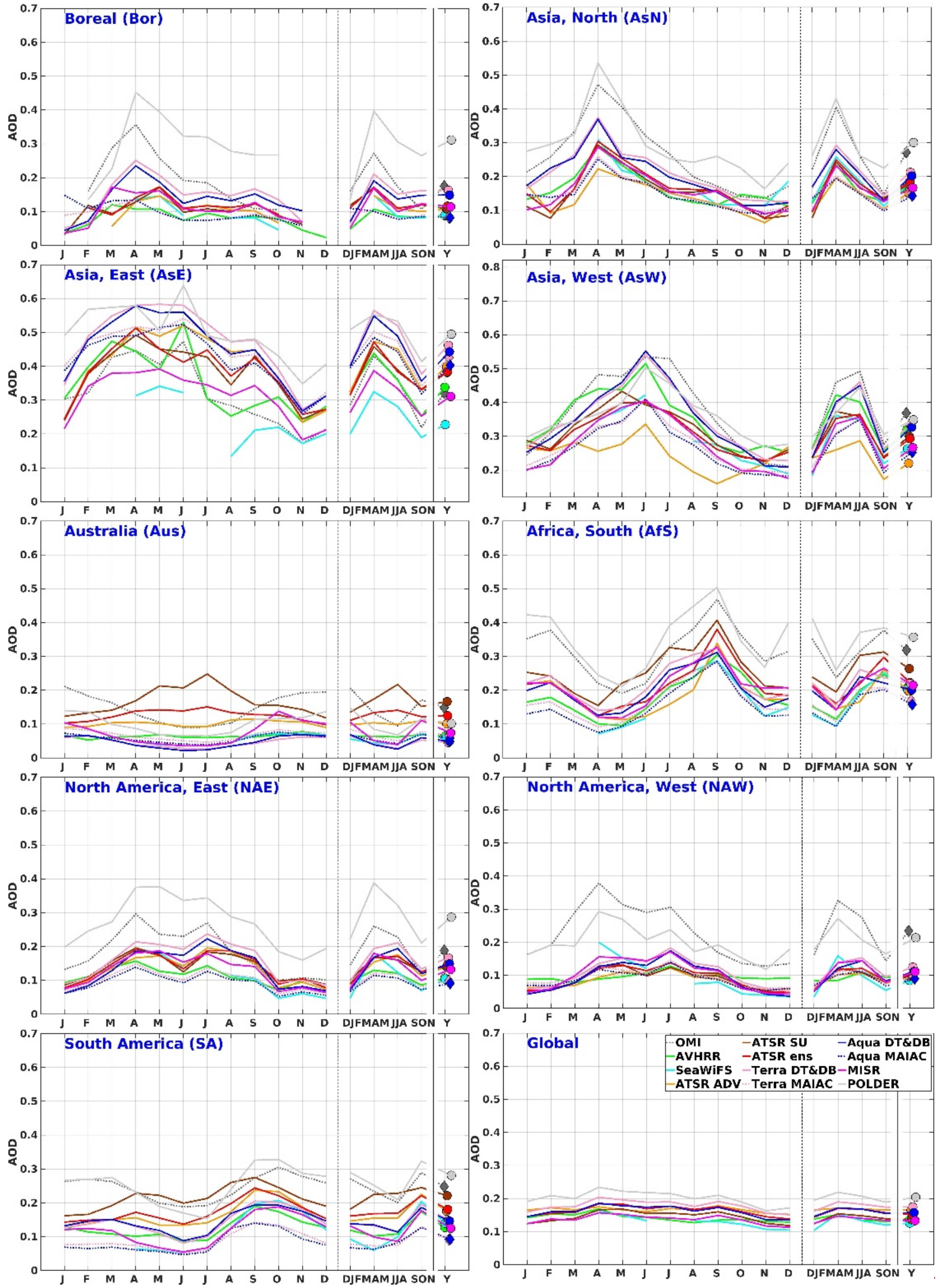


~~**Figure A2. AERONET comparison statistics (correlation coefficient R, bar; fraction of pixels in the Spread Envelope, SE, ○; fraction of pixels satisfying the GCOS requirements, GE, ⊕; Offset, Δ; root mean square error RMS, *) for AOD monthly aggregates for each product (1:15, legend for products below the plot) and the L3 merged product (M, for details see Sect.6.2) with corresponding colors (legend) for the selected regions (as in Fig. 1). N is a number of matches with AERONET. Note, for products which do not**~~

~~provide global coverage (e.g., no retrieval over oceans), the results are missing. For Ind, AOb, land, ocean (upper panel), Bor, AsN, AsE, AsW, Aus (lower left panel), AfS, NAE, NAW, SA, Global (lower right panel). For other regions, see Fig. 5.~~



~~Figure A3. Annual AOD time series for the selected regions. For 2018, MODIS DT&DB and VIIRS AOD products are shown here but not used in the merging exercise. For other regions, see Fig. 6.~~



~~**Figure A4. Monthly (J for January, F for February, etc., left plots) and seasonal (DJF, MAM, JJA, SON, middle plots) AOD time series for the selected regions for 2008, and yearly aggregate AOD for 2008 (Y, right) for selected regions. For other regions, see Fig. 7.**~~

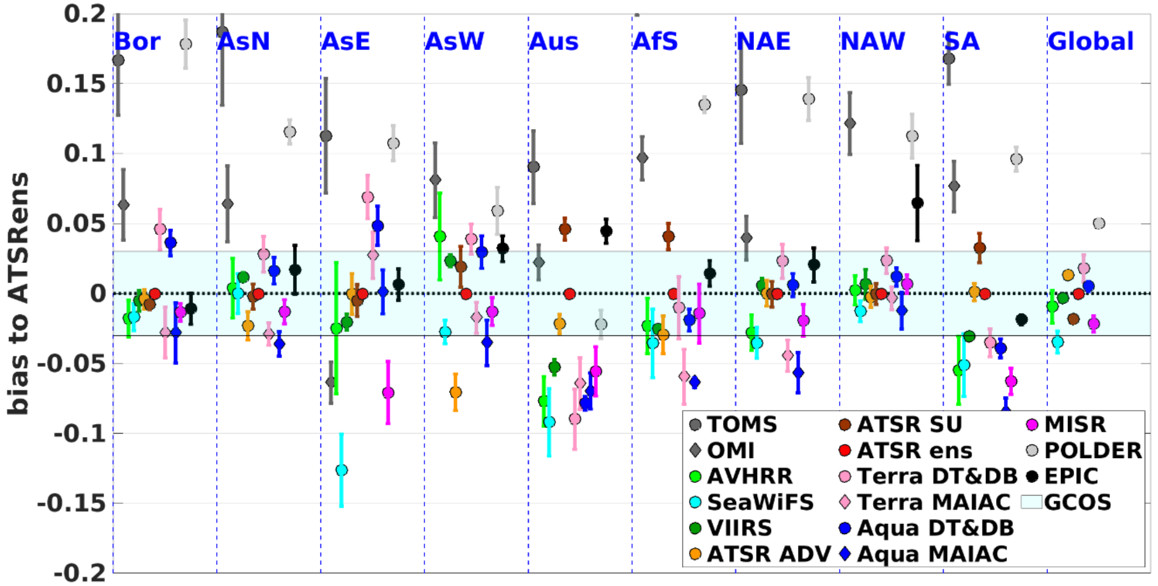
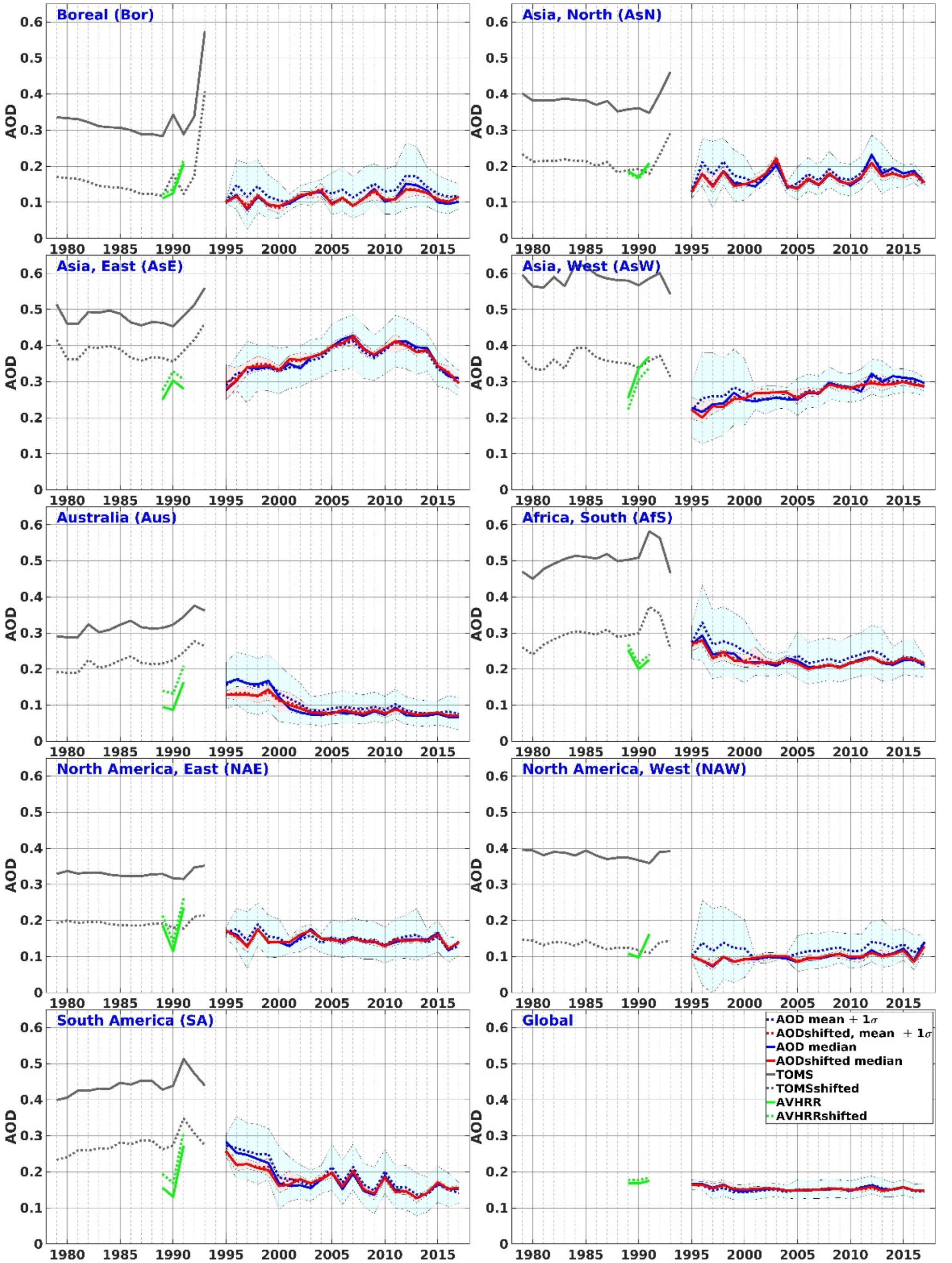
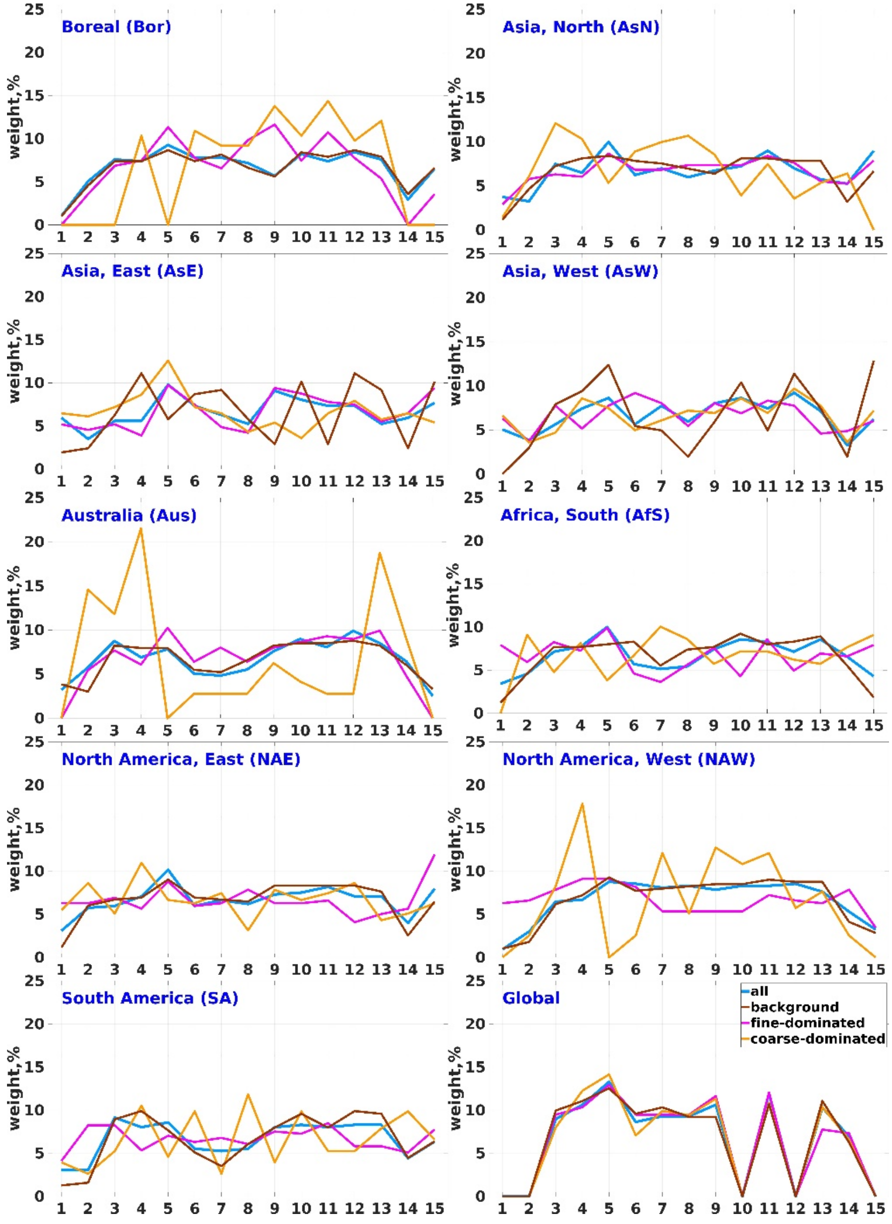


Figure A5. Regional annual average AOD offset between each dataset and the ATSR_ensemble dataset. GCOS requirement for ±0.03 AOD is shown as a background color. For other regions, see Fig. 8.

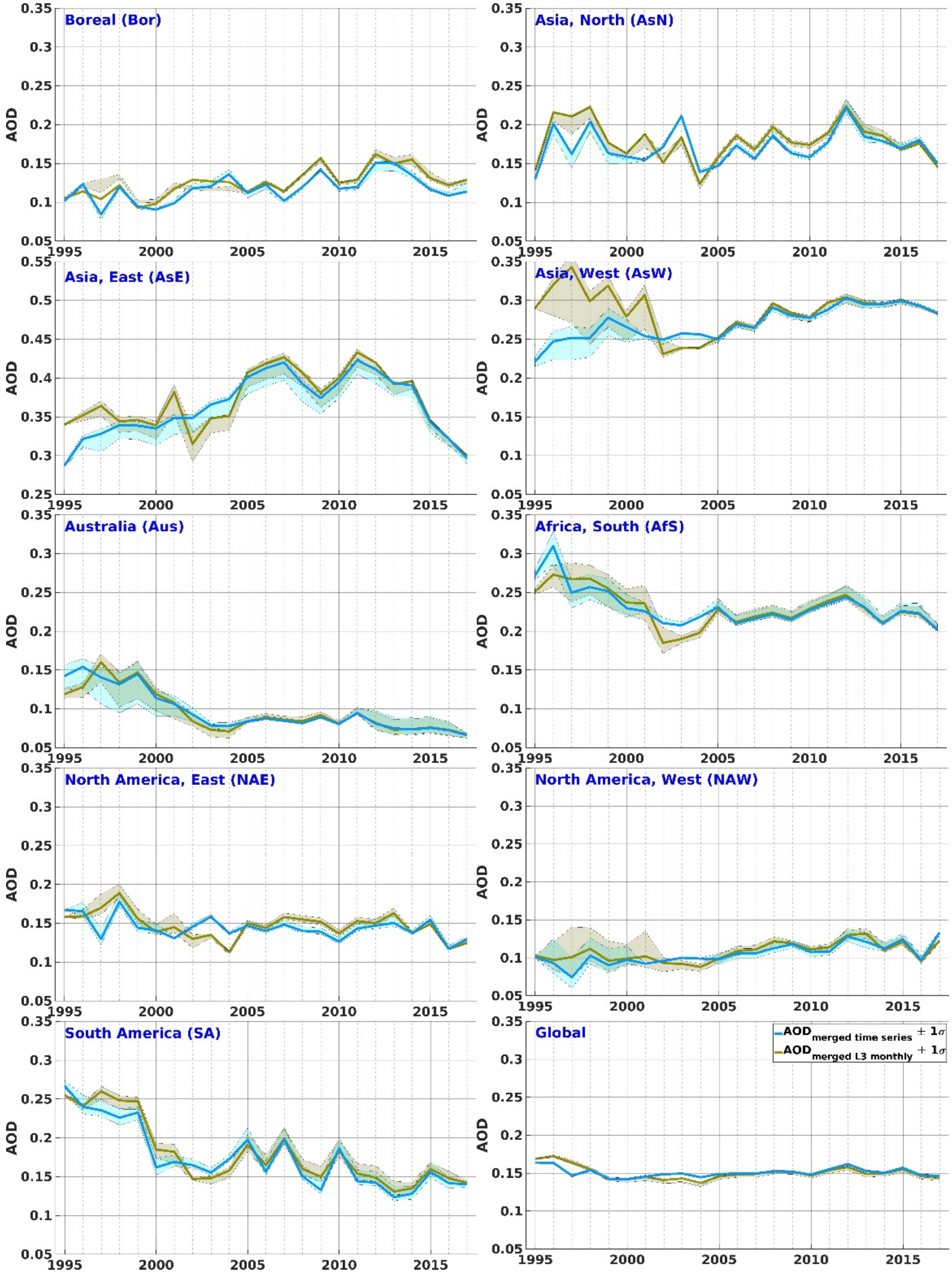


~~**Figure A6. Merged time series from uncorrected AODs (approach 1, blue) and from shifted AODs (approach 2, red) for the 8 selected regions (for both: mean, dotted line, median, solid line, and 1σ shadowed area). For other regions see Fig. 9.**~~

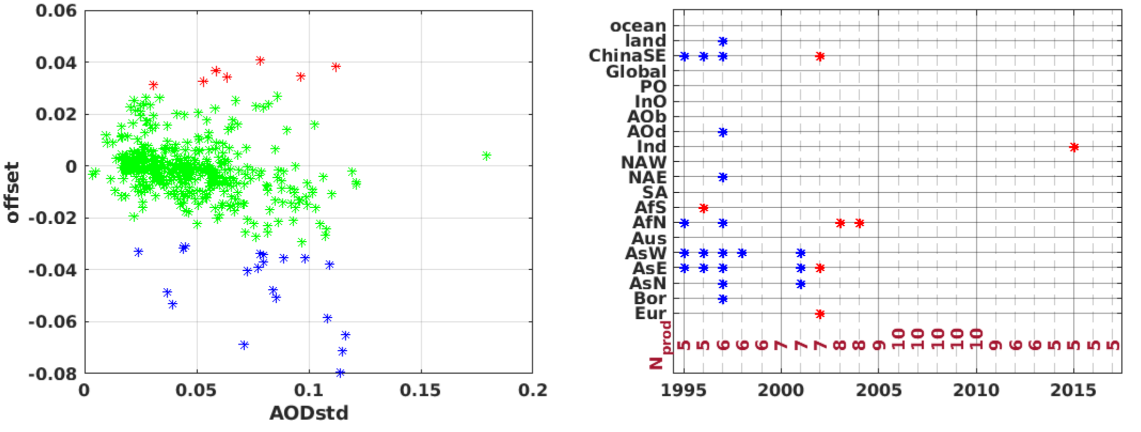


~~1 TOMS 3 AVHRR 6 ATSR_ADV 9 Terra 13 MISR~~
~~2 OMI 4 SeaWiFS 7 ATSR_SU 10 Terra MAIAC 14 PARASOL POLDER~~
~~5 VIIRS 8 ATSR_ens 11 Aqua 15 EPIC MAIAC~~
~~12 Aqua MAIAC~~

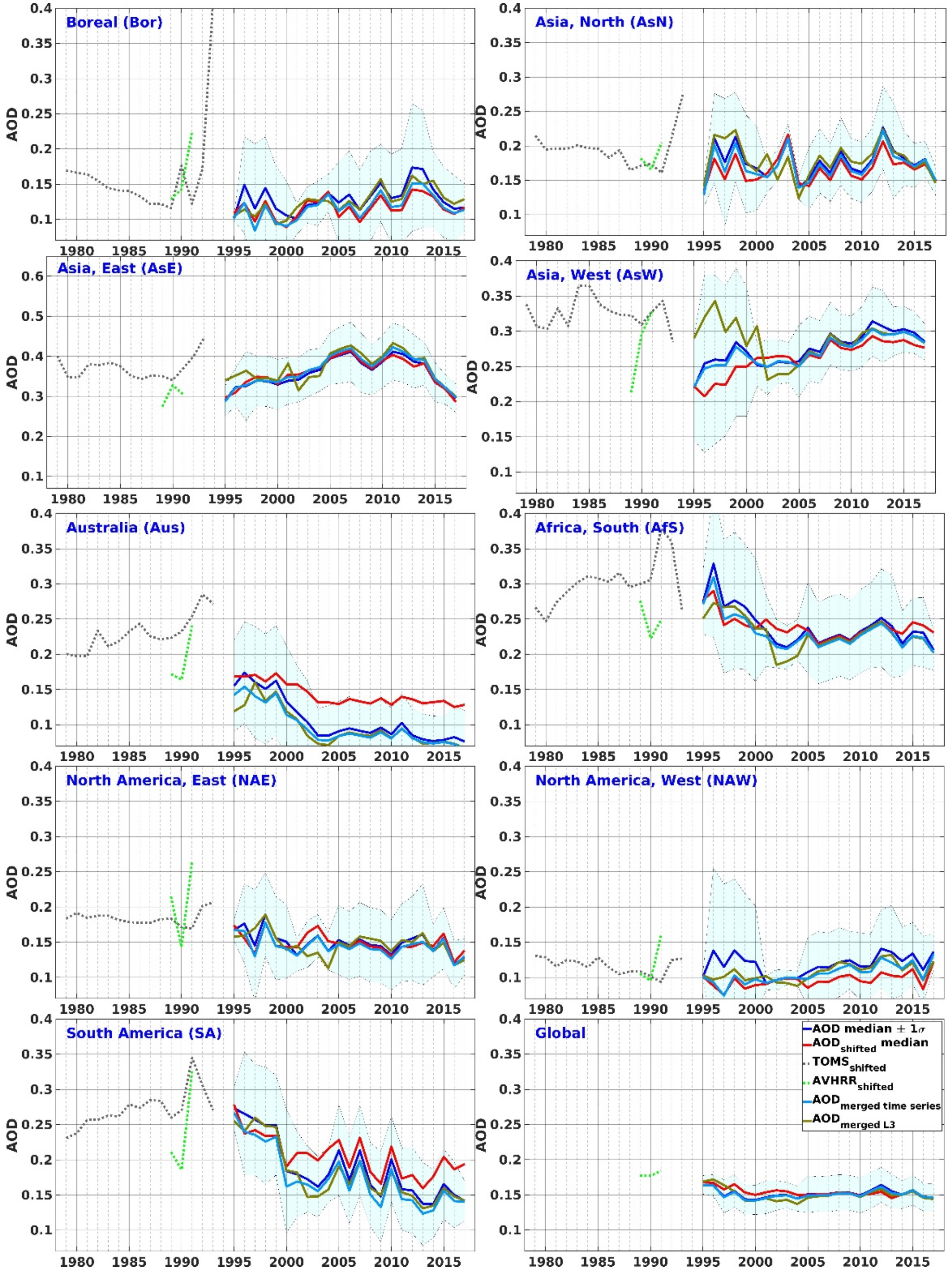
~~**Figure A7. Weight of each product for the selected regions for different aerosol types. For other regions, see Fig. 10.**~~



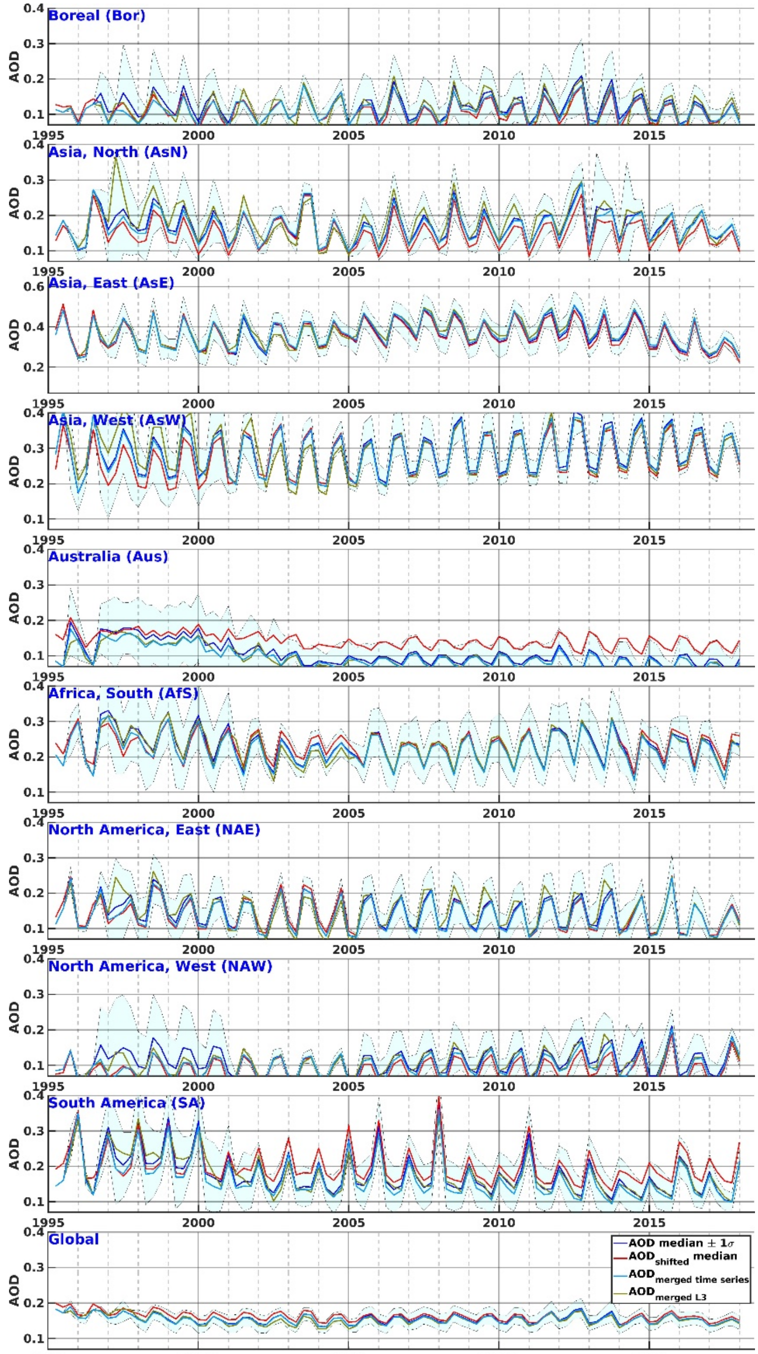
~~**Figure A8. Time series of merged (approach 3, all aerosol types) annual time series AOD (blue) and annual AOD time series calculated from L3 merged (with approach 3, for all aerosol types) AOD product (olive) for the selected regions. Deviations between time series calculated with different statistics for all, background, fine-dominated and coarse-dominated aerosol are shown as ±1δ. For other regions, see Fig. 12.**~~



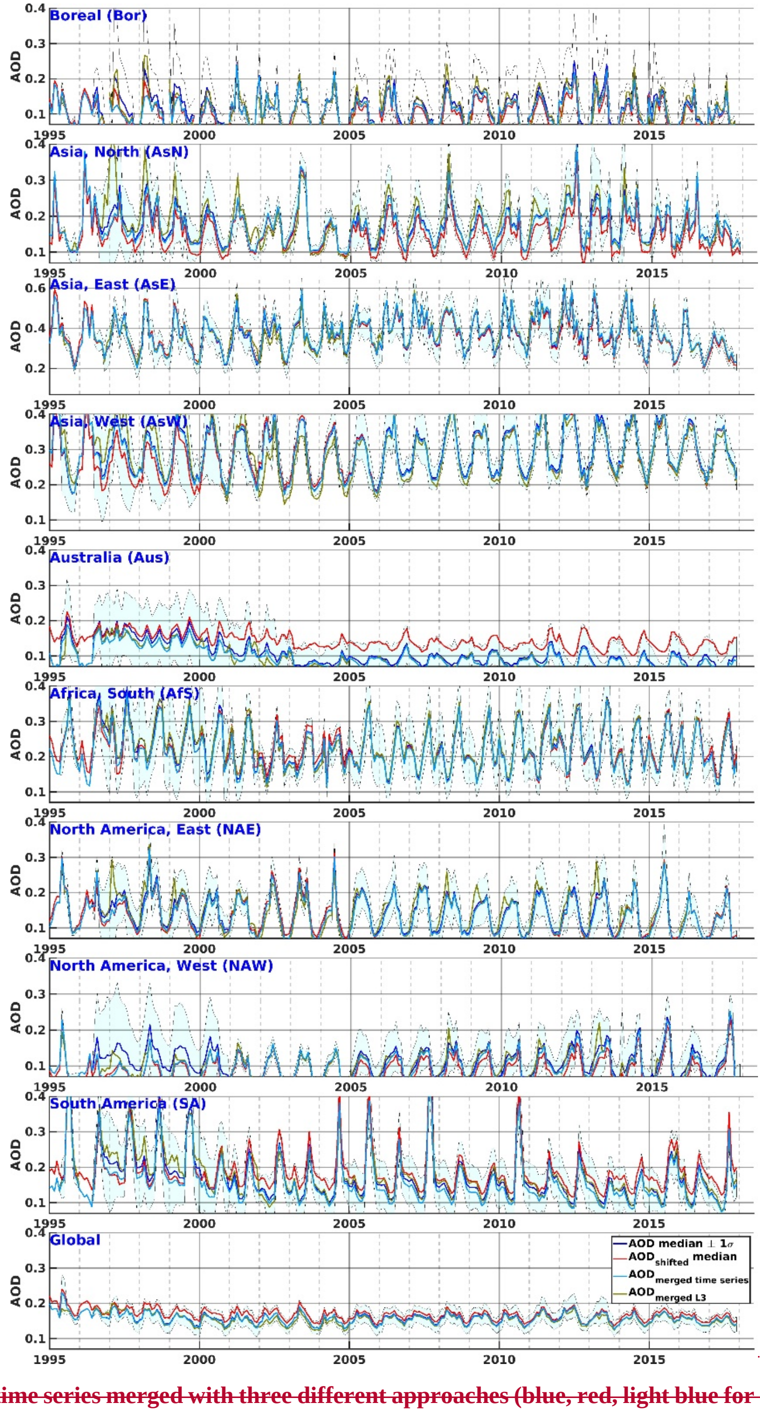
~~**Figure A9. Left: scatterplot for AOD standard deviations (AODstd) between the available products and offset between time series. Green stars - offset within the GCOS requirements of 0.03, red and blue starts – positive and negative (respectively) offset outside the GCOS requirements. Right – annual distribution of the offsets outside the GCOS requirements and number of the products (Nprod) used for merging for each year.**~~



~~**Figure A10. Annual AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. For other regions, see Fig. 13.**~~



~~**Figure A11. Seasonal AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. For other regions, see Fig. 14.**~~



~~**Figure A12. Monthly AOD time series merged with three different approaches (blue, red, light blue for approaches 1-3, respectively) and AOD time series from the L3 merged data (approach 3, olive) for the selected regions. ±1σ of the AOD from all uncorrected AOD products is shown as light blue shadow. Note the different scale for ChinaSE, AfN and Ind. For other regions, see Fig. 15.**~~

~~11~~ 9 Author contribution

The exercise on AOD merging has been initiated and widely discussed by the AeroCom/AeroSat community. The work has been performed by L. Sogacheva, who collected data, performed the analysis and wrote the extended draft of the manuscript. The evaluation results were widely discussed with the AOD data providers, who co-author the paper. Thomas Popp, ~~and~~ Andrew M. Sayer and Ralph Kahn considerably contributed to writing.

~~12~~ 10 Acknowledgments

The authors thank attendees of AeroCom/AeroSat workshops over the past several years for lively and informative discussions, which helped provide the impetus for and shape this analysis. AeroCom and AeroSat are unfunded community networks which participants contribute to within the remit and constraints of their other aerosol research.

The work presented is partly supported by the Copernicus Climate Change Service (contracts C3S_312a_lot5 and C3S_312b_Lot2) which are funded by the European Union, with support from ESA as part of the Climate Change Initiative (CCI) project Aerosol_cci (ESA-ESRIN projects AO/1-6207/09/I-LG and ESRIN/400010987 4/14/1-NB) and the AirQast 776361 H2020-EO-2017 project.

~~13~~ 11 References

Ban-Weiss, G. A., Jin, L., Bauer, S. E., Bennartz, R., Liu, X., Zhang, K., Ming, Y., Guo, H., and Jiang, J. H.: Evaluating clouds, aerosols, and their interactions in three global climate models using satellite simulators and observations, J. Geophys. Res. Atmos., 119, 10876–10901, doi:10.1002/2014JD021722, 2014.

Bellouin, N., Boucher, O., Haywood, J., and Shekar, Reddy M.: Global estimate of aerosol direct radiative forcing from satellite measurements, Nature, 438, 1138-1141, doi:10.1038/nature04348, 2005.

Bevan, S. North, P. Los, S. and Grey, W.: A global dataset of atmospheric aerosol optical depth and surface reflectance from AATSR. Remote Sensing of Environment 116, 199-210, 2012.

Boys, B. L., Martin, R. V., van Donkelaar, A., MacDonell, R. J., Hsu, N. C., Cooper, M. J., Yantosca, R. M., Lu, Z., Streets, D. G., Zhang, Q., and Wang, S. W.: Fifteen-year global time series of satellite-derived fine particulate matter. Environ Sci Technol., 48, 11109–11118, 2014.

Chang, C.-H., Hsiao, Y.-L., and Hwang, C.: Evaluating Spatial and Temporal Variations of Aerosol Optical Depth and Biomass Burning over Southeast Asia Based on Satellite Data Products, Aerosol and Air Quality Research, 15: 2625–2640, doi: 10.4209/aaqr.2015.10.0589, 2015.

Chatterjee, A., Michalak, A. M., Kahn, R. A., Paradise, S. R., Braverman, A. J., and Miller, C. E.: A geostatistical data fusion technique for merging remote sensing and ground-based observations of aerosol optical thickness, J. Geophys. Res., 115, D20207, doi:10.1029/2009JD013765, 2010.

Chin, M., Diehl, T., Tan, Q., Prospero, J. M., Kahn, R. A., Remer, L. A., Yu, H., Sayer, A. M., Bian, H., Geogdzhayev, I. V., Holben, B. N., Howell, S. G., Huebert, B. J., Hsu, N. C., Kim, D., Kucsera, T. L., Levy, R. C., Mishchenko, M. I., Pan, X., Quinn, P. K., Schuster, G. L., Streets, D. G., Strode, S. A., Torres, O., and Zhao, X.-P.: Multi-decadal aerosol variations from 1980 to 2009: a perspective from observations and a global model, Atmos. Chem. Phys., 14, 3657-3690, https://doi.org/10.5194/acp-14-3657-2014, 2014.

Chylek, P., Henderson, B., and Mishchenko, M.: Aerosol radiative forcing and the accuracy of satellite aerosol optical depth retrieval. J. Geophys. Res., 108, no. D24, 4764, doi:10.1029/2003JD004044, 2003.

de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W., Descloitres, J., Grainger, R. G., Griesfeller, J., Heckel, A., Kinne, S., Klüser, L., Kolmonen, P., Litvinov, P., Martynenko, D., North, P. J. R., Ovigneur, B., Pascal, N., Poulsen, C., Ramon, D., Schulz, M., Siddans, R., Sogacheva, L., Tanré, D., Thomas, G. E., Virtanen, T. H., von Hoyningen Huene, W., Vountas, M., and Pinnock, S.: Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis. Remote Sensing of Environment, 162, 295-315. 10.1016/j.rse.2013.04.023, 2015.

de Leeuw, G., Sogacheva, L., Rodriguez, E., Kourtidis, K., Georgoulias, A. K., Alexandri, G., Amiridis, V., Proestakis, E., Marinou, E., Xue, Y., and van der A, R.: Two decades of satellite observations of AOD over mainland China using ATSR-2, AATSR and MODIS/Terra: data set evaluation and large-scale patterns, Atmos. Chem. Phys., 18, 1573-1592, https://doi.org/10.5194/acp-18-1573-2018, 2018.

Dubovik, O., Holben, B. N., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanre, D., and Slutsker, I.: Variability of ab-sorption and optical properties of key aerosol types observed in worldwide locations, J. Atmos. Sci., 59, 590–608, 2002.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.: Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations, Atmos. Meas. Tech., 4, 975-1018, 2011.

Dubovik, O., Lapyonok, T., Litvinov, P., Herman, M., Fuertes, D., Ducos, F., Lopatin, A., Chaikovsky, A., Torres, B., Derimian, Y., Huang, X., Aspetsberger, M., and Federspiel, C.: GRASP: a versatile algorithm for characterizing the atmosphere, SPIE: Newsroom, DOI:10.1117/2.1201408.005558, http://spie.org/x109993.xml, 2014.

Dubovik, O., Li, Z., Mishchenko, M. I., Tanre, D., Karol, Y., Bojkov, B., Cairns, B., Diner, D. J., Espinosa, R., Goloub, P., Gu, X., Hasekamp, O., Hong, J., Hou, W., Knobelspiesse, K. D., Landgraf, J., Li, L., Litvinov, P., Liu, Y., Lopatin, A., Marbach, T., Maring, H., Martins, V., Meijer, Y., Milinevsky, G., Mukai, S., Parol, F., Qiao, Y., Remer, L., Rietjens, J., Sano, I., Stammes, P., Stamnes, S., Sun, X., Tabary, P. Travis, L. D., Waquet, F., Xu, F., Yan, C., and Yin, D.: Polarimetric remote sensing of atmospheric aerosols: instruments, methodologies, results, and perspectives, J. Quant. Spectrosc. Radiait. Transfer, doi.org/10.1016/j.jqsrt.2018.11.024, 474–511, 2019.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosol, J. Geophys. Res, 104, 31333–31350, 1999.

Escribano, J., Boucher, O., Chevallier, F., and Huneeus, N.: Impact of the choice of the satellite aerosol optical depth product in a sub-regional dust emission inversion, Atmos. Chem. Phys., 17, 7111-7126, https://doi.org/10.5194/acp-17-7111-2017, 2017.

Flowerdew, R. J., and Haigh, J. D.: An approximation to improve accuracy in the derivation of surface reflectances from multi-look satellite radiometers, Geophys. Res. Let., 22, 1693–1696, 1995.

Garay, M. J., Kalashnikova, O. V., and Bull, M. A.: Development and assessment of a higher-spatial-resolution (4.4km) MISR aerosol optical depth product using AERONET- DRAGON data, Atmos. Chem. Phys., 17, 5095–5106, https://doi.org/10.5194/acp-17-5095-2017, 2017.

Garay, M. J., Witek, M. L., Kahn, R. A., Seidel, F. C., Limbacher, J. A., Bull, M. A., Diner, D. J., Hansen, E. G., Kalashnikova, O. V., Lee, H., Nastan, A. M., and Yu, Y.: Introducing the 4.4 km Spatial Resolution MISR Aerosol Product, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-340, in review, 2019. ~~Garay, M. J., Witek, M. L., Kahn, R. A., Seidel, F. C., Limbacher, J. A., Bull, M. A., Diner, D. J., Hansen, E. G., Kalashnikova, O. V., Lee, H., Nastan, A. M., and Yu, Y.: The MISR Version 23 aerosol products over land and ocean, Atmos. Meas. Tech., *in preparation*.~~

GCOS, 2016, https://ane4bf-datap1.s3.eu-west-1.amazonaws.com/wmod8_gcos/s3fs-public/aerosols_ecv_factsheet_201905.pdf?Sv_8X3rsnl_rqNQVLEIg5gzig53zTHox, last accessed 19.06.2019.

~~Geogdzhayev, I. V., Mishchenko, M. I., Li, J., Rossow, W. B., Liu, L., Cairns, B.: Extension and statistical analysis of the GACP aerosol optical thickness record, Atmospheric Research, Volume 164, p. 268-277, doi: 10.1016/j.atmosres.2015.05.013, 2015.~~

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmos. Meas. Tech., 12, 169-209, https://doi.org/10.5194/amt-12-169-2019, 2019.

Gupta, P., Levy, R. C., Mattoo, S., Remer, L. A., and Munchak, L. A.: A surface reflectance scheme for retrieving aerosol optical depth over urban surfaces in MODIS Dark Target retrieval algorithm, Atm. Meas. Tech., 9 (7): 3293-3308, ~~[~~10.5194/amt-9-3293-2016~~]~~, 2016.

Han, B., Ding, H., Ma, Y., and Gong, W.: ~~(2017)~~ Improving retrieval accuracy for aerosol optical depth by fusion of MODIS and Caliop data. Tehni ki Vjesnik, 24(3), 791-800, 2017.

Heidinger, A. K., Cao, C., and Sullivan, J.: Using Moderate Resolution Imaging Spectrometer (MODIS) to calibrate Advanced Very High Resolution Radiometer (AVHRR) reflectance channels, J. Geophys. Res., 107(D23), 4702, doi:10.1029/2001JD002035, 2002.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y., Nakajima, T., Lavenu, F. , Jankowiak, I., and Smirnov, A.: AERONET — A federated instrument network and data archive for aerosol characterization, Remote Sens. of Env., 66, 1–16, 1998.

Holben, B.N., Tanre, D., Smirnow, A., Eck, T.F., Slutsker, I., Abuhassan, N., Newcomb, W.W., Schafer, J.S., Chatenet, B., Lavenu, F., Kaufman, Y.J., Vande Castle, J., Setzer, A., Markham, B., Clark, D., Halthore, R., Karneli, A., O'Neili, N.T., Pietras, C., Pinker, R.T., Vos. K., and Zibord, G.: An Emerging Ground-Based Aerosol Climatology." Journal of Geophysical Research 106(D11):12 067–12 097, 2001.

Hoskins, B. J., and Hodges, K. I.: A new perspective on Southern Hemisphere storm tracks, Journal of Climate, vol. 18, no. 20, pp. 4108–4129, 2005.

Hsu, N. C., Tsay, S. C., King, M. D., and Herman, J. R.: Aerosol properties over bright-reflecting source regions, IEEE Trans. Geosci. Remote Sens., 42, 557-569, 2004.

~~Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S. C., and Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, Atmos. Chem. Phys., 12, 8037-8053, https://doi.org/10.5194/acp-12-8037-2012, 2012.~~

Hsu, N. C., Jeong, M.-J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., Huang, J., and Tsay, S.-C.: Enhanced Deep Blue aerosol retrieval algorithm: The second generation, J. Geophys. Res. Atmos., 118, 9296– 9315, doi:10.1002/jgrd.50712, 2013.

Hsu, N. C., Lee, J., Sayer, A. M., Carletta, N., Chen, S.-H., Tucker, C. J., and Tsay, S.-C: Retrieving near-global aerosol loading over land and ocean from AVHRR, J. Geophys. Res., 122, doi:10.1002/2017JD026932, 2017.

Hsu, N. C., Lee, J., Sayer, A. M., Kim, W., Bettenhausen, C., and Tsay, S.-C.: VIIRS Deep Blue aerosol products over land: Extending the EOS long-term aerosol data records. Journal of Geophysical Research: Atmospheres, 124. https://doi.org/10.1029/2018JD029688, 2019.

Huang, D., Lyapustin, A., Korkin, S., Wang, Y., Blank, K., and Marshak, A.: The Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm for DSCOVR EPIC and initial analysis of data products, Rem. Sens. Environ., 2019 (*in review*).

Ignatov, A. and Stowe, L. L.: Aerosol Retrievals from Individual AVHRR Channels. Part I: Retrieval Algorithm and Transition from Dave to 6S Radiative Transfer Model, J. of Atm. Sciences, vol. 59, Issue 3, pp.313-334, 2002.

IPCC, 2018, https://www.ipcc.ch/site/assets/uploads/2018/11/pr_181008_P48_spm_en.pdf, last accessed 30.04.2019

Jethva, H., and Torres, O.: Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from Ozone Monitoring Instrument, Atmos. Chem. Phys., 11, 10,541–10,551, doi:10.5194/acp-11-10541-2011, 2011.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., and Holben, B. N.: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, J. Geophys. Res., 115, D23209, https://doi.org/10.1029/2010JD014601, 2010.

Kinne, S., Lohmann, U., Feichter, J., Schulz, M., Timmreck, C., Ghan, S., Easter, R., Chin, M., Ginouz, P. , Takemura, T., Tegen, I., Koch, D., Herzog, M., Penner, J., Pitari, G., Holben, B., Eck, T., Smirnov, A., Dubovik, O., Slutsker, I., Tanre, D., Torres, O., Mishchenko, M. Geogdzhayev, I., Chu, D.A., and Kaufman, Y.: Monthly averages of aerosol properties: A global comparison among models, satellite data, and AERONET ground data. J. Geophys. Res., 108, no. D20, 4634, doi:10.1029/2001JD001253, 2003.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment – optical properties in aerosol component modules of global models, Atmos. Chem. Phys., 6, 1815-1834, https://doi.org/10.5194/acp-6-1815-2006, 2006.

Kinne, S.: Remote sensing data combinations: superior global maps for aerosol optical depth. In book: Satellite Aerosol Remote Sensing over Land, Springer, 361-381, doi: 10.1007/978-3-540-69397-0_12, 2009.

Kokhanovsky, A. A., and de Leeuw, G. (Eds.): Satellite Aerosol Remote Sensing Over Land, Springer-Praxis (Berlin), 388p, 2009.

Kolmonen, P., Sogacheva, L., Virtanen, T. H., de Leeuw, G. , and Kulmala, M.: The ADV/ASV AATSR aerosol retrieval algorithm: current status and presentation of a full-mission AOD data set, International Journal of Digital Earth, 9:6, 545-561, doi: 10.1080/17538947.2015.1111450, 2016.

Kosmale, M., et al: *in preparation*

Lee, H., Garay, M. J., Kalashnikova, O. V., Yu, Y., and Gibson, P. B.: How Long should the MISR Record Be when Evaluating Aerosol Optical Depth Climatology in Climate Models? Remote Sens., 10, 1326, 2018.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Levy, R. C., Mattoo, S., Sawyer, V., Shi, Y., Colarco, P. R., Lyapustin, A. I., Wang, Y., and Remer, L. A.: Exploring systematic offsets between aerosol products from the two MODIS sensors, Atmos. Meas. Tech., 11, 4073-4092, https://doi.org/10.5194/amt-11-4073-2018, 2018.

Li, J., Carlson, B. E., and Lacis, A. A.: Application of spectral analysis techniques in the inter-comparison of aerosol data. Part I: An EOF approach to analyze the spatial-temporal variability of aerosol optical depth using multiple remote sensing data sets, J. Geophys. Res. Atmos., 118, 8640–8648, doi:10.1002/jgrd.50686, 2013.

Li, J., Carlson, B. E., and Lacis, A. A.: Application of spectral analysis techniques in the inter-comparison of aerosol data. Part II: Using maximum covariance analysis to effectively compare spatio-temporal variability of satellite and AERONET measured aerosol optical depth, J. Geophys. Res. Atmos., 119, 153–166, doi:10.1029/2013JD020537, 2014a.

Li, J., Carlson, B. E., and Lacis, A. A.: Application of spectral analysis techniques to the intercomparison of aerosol data – Part 4: Synthesized analysis of multisensor satellite and ground-based AOD measurements using combined maximum covariance analysis, Atmos. Meas. Tech., 7, 2531-2549, https://doi.org/10.5194/amt-7-2531-2014, 2014b.

Li, J., Li, X., Carlson, B. E., Kahn, R. A., Lacis, A. A., Dubovik, O., and Nakajima, T.: Reducing multisensor satellite monthly mean aerosol optical depth uncertainty: 1. Objective assessment of current AERONET locations, J. Geophys. Res. Atmos., 121, 13,609–13,627, doi:10.1002/2016JD025469, 2016a.

Li, S., Yu, C., Chen, L., Tao, J., Letu, H., Ge, W., Si, Y., and Liu, Y.: Inter-comparison of model-simulated and satellite-retrieved componential aerosol optical depths in China, Atmos. Environ., 141, 320–332, 10.1016/j.atmosenv.2016.06.075, 2016b.

Li, Z., Zhao, X., Kahn, R. A., Mishchenko, M., Remer, L., Lee, K.-H., Wang, M., Laszlo, I., Nakajima, T., and Maring, H.: Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective. Ann. Geophys. 27, 2755–2770, 2009.

Limbacher, J. A. and Kahn, R. A.: Updated MISR dark water research aerosol retrieval algorithm – Part 2: Aerosol and surface-reflectance retrievals over shallow, turbid, and eutrophic water. Atmos. Meas. Tech. 12, 675–689, doi:10.5194/amt-12-675-2019, 2019.

Liu, L., Lacis, A. A., Carlson, B. E., Mishchenko, M. I., and Cairns, B.: Assessing Goddard Institute for Space Studies ModelE aerosol climatology using satellite and ground-based measurements: A comparison study, J. Geophys. Res., 111, D20212, doi:10.1029/2006JD007334, 2006.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC algorithm, Atmos. Meas. Tech., 11, 5741-5765, https://doi.org/10.5194/amt-11-5741-2018, 2018.

Martonchik, J. V., Kahn, R. A., and Diner, D. J.: Retrieval of Aerosol Properties over Land Using MISR Observations. In: Kokhanovsky, A.A. and G. de Leeuw, ed., Satellite Aerosol Remote Sensing Over Land. Springer, Berlin, pp.267-293, 2009.

Michou, M., Nabat, P., and Saint-Martin, D.: Development and basic evaluation of a prognostic aerosol scheme (v1) in the CNRM Climate Model CNRM-CM6, Geosci. Model Dev., 8, 501-531, https://doi.org/10.5194/gmd-8-501-2015, 2015.

Mishchenko, M. I., Geogdzhayev, I. V., Cairns, B., Carlson, B. E., Chowdhary, J., Lacis, A. A., Liu, L., Rossow, W. B., and Travis, L. D.: Past, present, and future of global aerosol climatologies derived from satellite observations: A perspective. J. Quant. Spectrosc. Radiat. Transf., 106, 325-347, doi:10.1016/j.jqsrt.2007.01.007, https://pubs.giss.nasa.gov/docs/2007/2007_Mishchenko_mi01210n.pdf, 2007.

~~Mishchenko, M.I., Geogdzhayev, I. V., Cairns, B., Carlson, B. E., Chowdhary, J., Lacis, A. A., Liu, L., Rossow, W. B., and Travis, L.D.: Past, present, and future of global aerosol climatologies derived from satellite observations: A perspective. J. Quant. Spectrosc. Radiat. Transf., 106, 325-347, doi:10.1016/j.jqsrt.2007.01.007, https://pubs.giss.nasa.gov/docs/2007/2007_Mishchenko_mi01210n.pdf, 2007.~~

Nabat, P., Somot, S., Mallet, M., Chiapello, I., Morcrette, J. J., Solmon, F., Szopa, S., Dulac, F., Collins, W., Ghan, S., Horowitz, L. W., Lamarque, J. F., Lee, Y. H., Naik, V., Nagashima, T., Shindell, D., and Skeie, R.: A 4-D climatology (1979–2009) of the monthly tropospheric aerosol optical depth distribution over the Mediterranean region from a comparative evaluation and blending of remote sensing and model products, Atmos. Meas. Tech., 6, 1287-1314, https://doi.org/10.5194/amt-6-1287-2013, 2013.

Naeger, A. R., Gupta, P., Zavodsky, B. T., and McGrath, K. M.: Monitoring and tracking the trans-Pacific transport of aerosols using multi-satellite aerosol optical depth composites, Atmos. Meas. Tech., 9, 2463-2482, https://doi.org/10.5194/amt-9-2463-2016, 2016.

O'Neill, N. T., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics: Empirical tests using multi-year, multi-site AERONET sun-photometer data, Geophys. Res. Lett., 27(20), 3333–3336. doi:10.1029/2000GL011581, 2000.

North, P.: Estimation of aerosol opacity and land surface bidirectional reflectance from ATSR-2 dual-angle imagery: Operational method and validation, J. of Geophys. Res., 107(D12), 2002.

North, P., Briggs, S., Plummer, S. and Settle, J.: Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery, IEEE Transactions on Geoscience and Remote Sensing 37(1), 526-537, 1999.

Penning de Vries, M. J. M., Beirle, S., Hörmann, C., Kaiser, J. W., Stammes, P., Tilstra, L. G., Tuinder, O. N. E., and Wagner, T.: A global aerosol classification algorithm incorporating multiple satellite data sets of aerosol and trace gas abundances, Atmos. Chem. Phys., 15, 10597-10618, https://doi.org/10.5194/acp-15-10597-2015, 2015.

Peyridieu, S., Chédin, A., Capelle, V., Tsamalis, C., Pierangelo, C., Armante, R., Crevoisier, C., Crépeau, L., Siméon, M., Ducos, F., and Scott, N. A.: Characterisation of dust aerosols in the infrared from IASI and comparison with PARASOL, MODIS, MISR, CALIOP, and AERONET observations, Atmos. Chem. Phys., 13, 6065-6082, https://doi.org/10.5194/acp-13-6065-2013, 2013.

Pinty, B., Taberner, M., Haemmerle, V.., Paradise, S. R., Vermote, E., Verstraete, M. M., Gobron, N., Widlowski, J. L.: Global-Scale Comparison of MISR and MODIS Land Surface Albedos, J. of Climate, 24(3), 732-749, 2011.

Popp, T., de Leeuw, G., Bingen, C., Brühl, C., Capelle, V., Chedin, A., Clarisse, L., Dubovik, O., Grainger, R., Griesfeller, J., Heckel, A., Kinne, S., Klüser, L., Kosmale, M., Kolmonen, P., Lelli, L., Litvinov, P., Mei, L., North, P., Pinnock, S., Povey, A., Robert, C., Schulz, M., Sogacheva, L., Stebel, K., Stein Zweers, D., Thomas, G., Tilstra, L.G., Vandenbussche, S., Veefkind, P., Vountas, M., Xue, Y.: Development, Production and Evaluation of Aerosol Climate Data Records from European Satellite Observations (Aerosol_cci), Remote Sens., 8, 421, 2016.

Sayer, A. M., Thomas, G. E., and Grainger, R. G.: A sea surface reflectance model for (A)ATSR, and application to aerosol retrievals, Atmos. Meas. Tech., 3, 813-838, https://doi.org/10.5194/amt-3-813-2010, 2010.

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Ahmad, Z., Holben, B. N., Smirnov, A., Thomas, G. E., and Zhang, J.: SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets, J. Geophys. Res., 117, D03206, doi:10.1029/2011JD016599, 2012a.

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, M.-J., Holben, B. N., and Zhang, J.: Global and regional evaluation of over-land spectral aerosol optical depth retrievals from SeaWiFS, Atmos. Meas. Tech., 5, 1761-1778, https://doi.org/10.5194/amt-5-1761-2012, 2012b.

Sayer, A. M., Munchak, L. A., Hsu, N. C., Levy, R. C., Bettenhausen, C., Jeong, M. J.: MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations.Journal of Geophysical Research-Atmospheres, Vol.119, pp. 13965-13989, doi: https://doi.org/10.1002/2014jd022453, 2014.

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, M.-J., and Meister, G.: Effect of MODIS Terra radiometric calibration improvements on Collection 6 Deep Blue aerosol products: Validation and Terra/Aqua consistency, J. Geophys. Res. Atmos., 120, 12,157–12,174, doi:10.1002/2015JD023878, 2015.

Sayer, A. M., Hsu, N. C., Lee, J., Carletta, N., Chen, S.-H., and Smirnov, A.: Evaluation of NASA Deep Blue/SOAR aerosol retrieval algorithms applied to AVHRR measurements, J. Geophys. Res. Atmos., 122, 9945– 9967, doi:10.1002/2017JD026934, 2017.

Sayer, A. M., Hsu, N. C., Lee, J., Bettenhausen, C., Kim, W. V., and Smirnov, A.: Satellite Ocean Aerosol Retrieval (SOAR) algorithm extension to S-NPP VIIRS as part of the "Deep Blue" aerosol project, J. of Geophys. Res., Atmospheres, 123, 380–400. https://doi.org/10.1002/2017JD027412, 2018a.

Sayer, A. M., Hsu, N. C., Lee, J., Kim, W. V., Dubovik, O., Dutcher, S. T., Huang, D., Litvinov, P., Lyapustin, A., Tackett, J. L., and Winker, D. M.: Validation of SOAR VIIRS over-water aerosol retrievals and context within the global satellite aerosol data record, J. of Geophys. Res., Atmospheres, 123, 13,496– 13,526, https://doi.org/10.1029/2018JD029465, 2018b.

Sayer, A. M., Hsu, N. C., Lee, J., Kim, W., and Dutcher, S.: Validation, stability, and consistency of MODIS Collection 6.1 and VIIRS Version 1 Deep Blue aerosol data over land, J. Geophys. Res. Atmos., 124, https://doi.org/10.1029/2018JD029598, 2019.

Sayer, A. M. and Knobelspiesse, K. D., How should we aggregate data? Methods accounting for the numerical distributions, with an assessment of aerosol optical depth, Atmos. Chem. Phys., https://doi.org/10.5194/acp-2019-372, in press, 2019.

Shi, Y., Zhang, J., Reid, J. S., Hyer, E. J., Eck, T. F., Holben, B. N., and Kahn, R. A.: A critical examination of spatial biases between MODIS and MISR aerosol products -- application for potential AERONET deployment, Atmos. Meas. Tech., 4, 2823-2836, https://doi.org/10.5194/amt-4-2823-2011, 2011.

Shi, Y. R., Levy, R. C., Eck, T. F., Fisher, B., Mattoo, S., Remer, L. A., Slutsker, I., and Zhang, J.: Characterizing the 2015 Indonesia fire event using modified MODIS aerosol retrievals, Atmos. Chem. Phys., 19, 259-274, https://doi.org/10.5194/acp-19-259-2019, 2019.

Schutgens, N. A. J.: Site representativity of AERONET and GAW remotely sensed AOT and AAOT observations, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-767, in review, 2019.

Smirnov, A., Holben, B. N., Slutsker, I., Giles, D. M., Mc-Clain, C. R., Eck, T. F., Sakerin, S. M., Macke, A., Croot, P., Zibordi, G., Quinn, P. K., Sciare, J., Kinne, S., Harvey, M., Smyth, T. J., Piketh, S., Zielinski, T., Proshutinsky, A., Goes, J. I., Nelson, N. B., Larouche, P., Radionov, V. F., Goloub, P., Krishna Moorthy, K., Matarrese, R., Robertson, E. J., and Jourdin, F.: Maritime Aerosol Network as a component of Aerosol Robotic Network, J. Geophys. Res.-Atmos., 114, D06204, doi:10.1029/2008JD011257, 2009.

Sogacheva, L., Kolmonen, P., Virtanen, T. H., Rodriguez, E., Saponaro, G., and de Leeuw, G.: Post-processing to remove residual clouds from aerosol optical depth retrieved using the Advanced Along Track Scanning Radiometer, Atmos. Meas. Tech., 10, 491-505, doi:10.5194/amt-10-491-2017, 2017.

Sogacheva, L., de Leeuw, G., Rodriguez, E., Kolmonen, P., Georgoulias, A. K., Alexandri, G., Kourtidis, K., Proestakis, E., Marinou, E., Amiridis, V., Xue, Y., and van der A, R. J.: Spatial and seasonal variations of aerosols over China from two decades of multi-satellite observations – Part 1: ATSR (1995–2011) and MODIS C6.1 (2000–2017), Atmos. Chem. Phys., 18, 11389-11407, https://doi.org/10.5194/acp-18-11389-2018, 2018a.

Sogacheva, L., Rodriguez, E., Kolmonen, P., Virtanen, T. H., Saponaro, G., de Leeuw, G., Georgoulias, A. K., Alexandri, G., Kourtidis, K., and van der A, R. J.: Spatial and seasonal variations of aerosols over China from two decades of multi-satellite observations – Part 2: AOD time series for 1995–2017 combined from ATSR ADV and MODIS C6.1 and AOD tendency estimations, Atmos. Chem. Phys., 18, 16631-16652, https://doi.org/10.5194/acp-18-16631-2018, 2018b.

Tang, Q., Bo, Y., and Zhu Y.: Spatiotemporal fusion of multiple-satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method, J. Geophys. Res. Atmos., 121, 4034–4048, doi:10.1002/ 2015JD024571, 2016.

Thomas, G. E., Carboni, E., Sayer, A. M., Poulsen, C. A., Siddans, R., and Grainger, R. G.: Oxford-RAL Aerosol and Cloud(ORAC): aerosol retrievals from satellite radiometers, in: Satellite Aerosol Remote Sensing over Land, edited by: Kokhanovsky, A. and de Leeuw, G., Springer Praxis Books, Springer, Berlin,Heidelberg, 193–225, https://doi.org/10.1007/978-3-540-69397-0_7, 2009

Toth, T. D., Zhang, J., Campbell, J. R., Reid, J. S., Shi, Y., Johnson, R. S., Smirnov, A., Vaughan, M. A., and Winker, D. M.: Investigating enhanced Aqua MODIS aerosol optical depth retrievals over the mid-to-high latitude Southern Oceans through intercomparison with co-located CALIOP, MAN, and AERONET data sets, J. Geophys. Res. Atmos.,118, 4700–4714, doi:10.1002/jgrd.50311, 2013.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis, *J. Geophys. Res.*, 103(D14), 17099– 17110, doi:10.1029/98JD00900, 1998.

~~Torres, O., Bhartia, P. K., Herman, J. R., Sinyuk, A., and Holben, B.: A long-term record of aerosol optical thickness from TOMS observations and comparison to AERONET measurements, J. Atmos. Sci., 59, 398–413, 2002.~~

Torres, O., Bhartia, P. K., Sinyuk, A., Welton, E. J., and Holben, B.: Total Ozone Mapping Spectrometer measurements of aerosol absorption from space: Comparison to SAFARI 2000 ground-based observations, *J. Geophys. Res.*, 110, D10S18, doi:10.1029/2004JD004611, 2005.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, *J. Geophys. Res.*, 112, D24S47, doi:10.1029/2007JD008809, 2007.

Torres, O., Ahn, C., and Chen, Z.: Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations, Atmos. Meas. Tech., 6, 3257–3270, doi:10.5194/amt-6-3257-2013, 2013.

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.

Veefkind, J. P., de Leeuw, G., and Durkee, P. A.: Retrieval of aerosol optical depth over land using two-angle view satellite radiometry during TARFOX, Geophys. Res. Let., 25 (16), 3135-3138, 1998.

Virtanen, T. H., Kolmonen, P., Sogacheva, L., Rodríguez, E., Saponaro, G., and de Leeuw, G.: Collocation mismatch uncertainties in satellite aerosol retrieval validation, Atmos. Meas. Tech., 11, 925-938, https://doi.org/10.5194/amt-11-925-2018, 2018.

Wei, J., Peng, Y., Mahmood, R., Sun, L., and Guo, J.: Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products, Atmos. Chem. Phys., 19, 7183-7207, https://doi.org/10.5194/acp-19-7183-2019, 2019.

Witek, M. L., Garay, M. J., Diner, D. J., Bull, M. A., and Seidel, F. C.: New approach to the retrieval of AOD and its uncertainty from MISR observations over dark water, Atmos. Meas. Tech., 11, 429-439, https://doi.org/10.5194/amt-11-429-2018, 2018.

WMO: Guidelines on the Calculation of Climate Normals, WMO-No.1203, 2017.

Zhang, J., and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products, Atmos. Chem. Phys., 10, 10949-10963, https://doi.org/10.5194/acp-10-10949-2010, 2010.

Zhao X. P., Laszlo, I., Guo, W., Heidinger, A., Cao, C., Jelenak, A., Tarpley, D., and Sullivan, J.: Study of long-term trend in aerosol optical thickness observed from operational AVHRR satellite instrument, J. Geophys. Res., 113, D07201, doi:10.1029/2007JD009061, 2008.

Zhao, T. X. P., Chan, P. K., and Heidinger, A. K.: A global survey of the effect of cloud contamination on the aerosol optical thickness and its long-term trend derived from operational AVHRR satellite observations. Journal of Geophysical Research: Atmospheres, 118(7), 2849-2857, doi: 10.1002/jgrd.50278, 2013.