Author's response to Interactive comments of Anonymous Referee #2 on "An Aerosol Optical Depth time series 1982–2014 for atmospheric correction based on OMI and TOMS Aerosol Index" by E. Jääskeläinen et al.

We thank the referee for careful reading of our manuscript and for the detailed comments. We will incorporate these comments to the revised manuscript. Below, we list referees' comments followed by our answers (in blue). The pages and lines included in our answers refer to the revised manuscript.

The paper describes a method to produce a long-term aerosol optical depth (AOD) dataset reaching back to 1982. The purpose of this AOD dataset is its use for atmospheric correction of an AVHRR surface albedo time series 1982 - 2014. This is an important application and the AOD dataset is of high value, since no global AOD dataset suitable for this purpose exists over this long period. The title of the paper clearly states this specific limitation to one intended application of the AOD dataset. The final results for atmospheric correction prove the potential of the created dataset. However, the paper text is too short and needs to be extended to clearly describe the method used to produce the AOD dataset (e.g. snow / ice discrimination method, gap filling method, sub-class building, exclusion of low / high aerosol index, : : :).

Many figures need to be described in the text: what the reader can see, what conclusion is drawn from them, what statement shall be highlight with it.

There are now more descriptions of the figures (Figures 3, 4, 7, 8, 11, 18-22).

In addition, the discussion of the impact of assumptions and the results achieved needs to be largely enhanced (e.g. fixed AOD over snow and ice, impact of differences between morning and afternoon orbits, : : :).

The AOD time series presented in the manuscript is the first version of it, so some of the most difficult aspects (like AOD values over ice and snow or the impact of the different orbits) are left to be solved in the future. The atmosphere is thin over the poles so the accurate magnitude of the atmospheric correction is not that critical. Hence, we don't currently calculate AOD over the permanently snow covered areas. Because we cannot define the AOD values precisely over the AVHRR orbit to be used in the surface albedo retrieval, the first step to that direction is the daily value.

In particular the sensitivities of the AI to other parameters (foremost aerosol layer height, but also surface albedo, geometry, used UV wavelength pair) needs to be discussed.

We added more text about the sensitivities of the AI to other parameters to the manuscript, P. 5/L. 29-31.

Additionally, the omission of non-absorbing aerosols as part of total AOD by the AI needs to be discussed.

Even though the negative AI values are discarded, the regressions are made by using the total AOD from the OMI instrument together with the land use classification information. So the use of the absorbing AI values still produce AOD values which are more close to the total AOD than to only absorbing AOD values. Text added to the manuscript, P. 4/L. 8-10, P. 10/L. 27-30.

Furthermore, a conclusion section of only 7 lines is not suitable for a scientific paper. You need to summarize / discuss: impact of most critical assumptions, what have you achieved, what does a general reflectance increase mean, for which application is the mean validation on large aggregates sufficient, OMI AOD is not perfect - but taken here as truth...

The conclusions section is now expanded a little bit and a new section, Section 7: Discussion and conclusions is added to the manuscript; both changes will now address the mentioned matters in more detail.

I therefore recommend a major revision of the paper.

I recommend to start out from a discussion of the required accuracy for an AOD dataset to be used for atmospheric correction; this would then more clearly distinguish the atmospheric-correction AOD dataset from an AOD dataset for aerosol studies. In particular it should be stated which use of the AVHRR albedo dataset the authors have in mind (e.g. change detection of more qualitative and step-wise large differences over time, climate monitoring with small trends only to be detected in a noisy but stable time series), because this will determine the needed albedo accuracy and consequently the required AOD accuracy and stability. In the final discussion the achieved AOD accuracy can then be assessed in comparison to the assumption of a fixed AOD = 0.1.

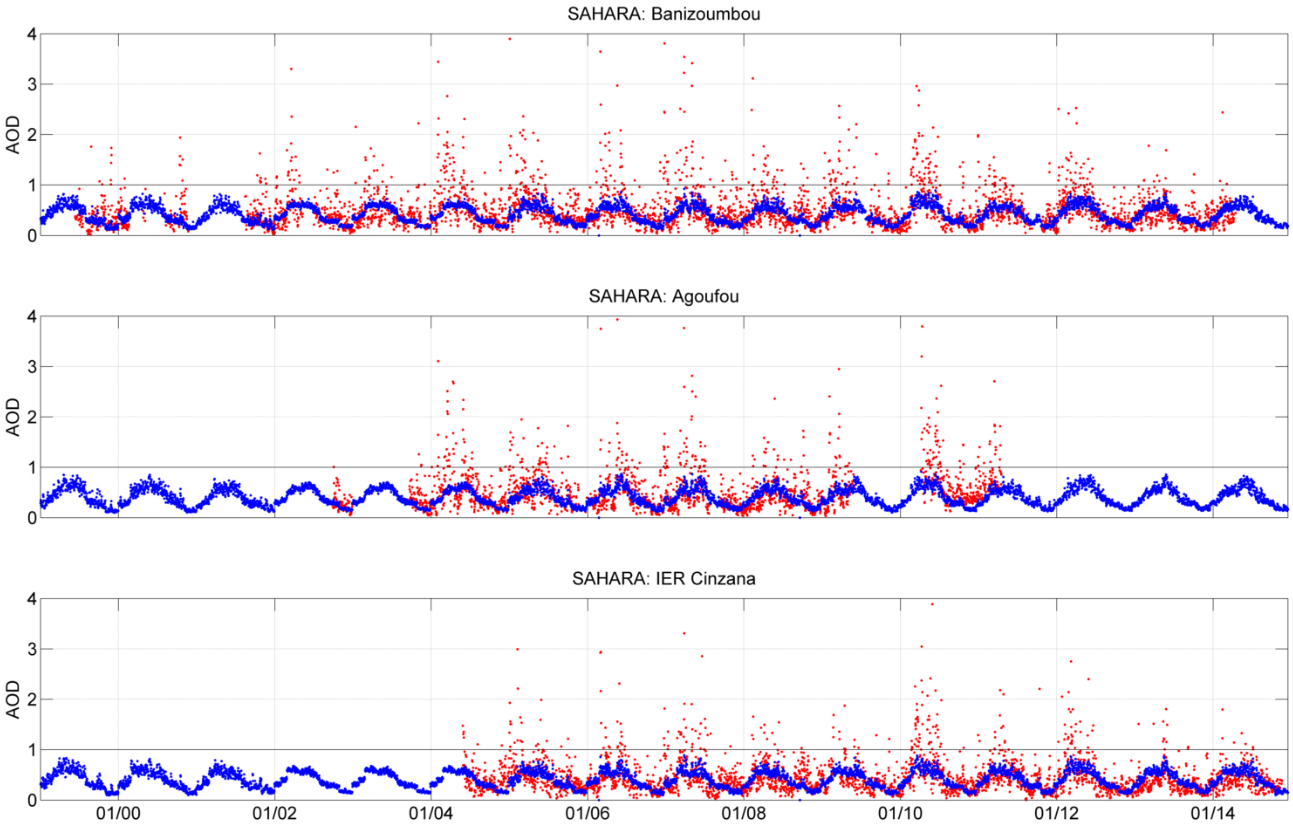
The minimum requirement was to achieve a better accuracy by using the constructed AOD time series compared to the constant AOD 0.1. Because the CLARA-A2 SAL data are used for climate models, it is more important that the overall quality and levels of the data set are right than that all the dynamic changes are detected. For the next versions of the AOD time series, this matter will be taken into account, a note of this is added to the new section in the manuscript, section 7 titled as Discussion and conclusions.

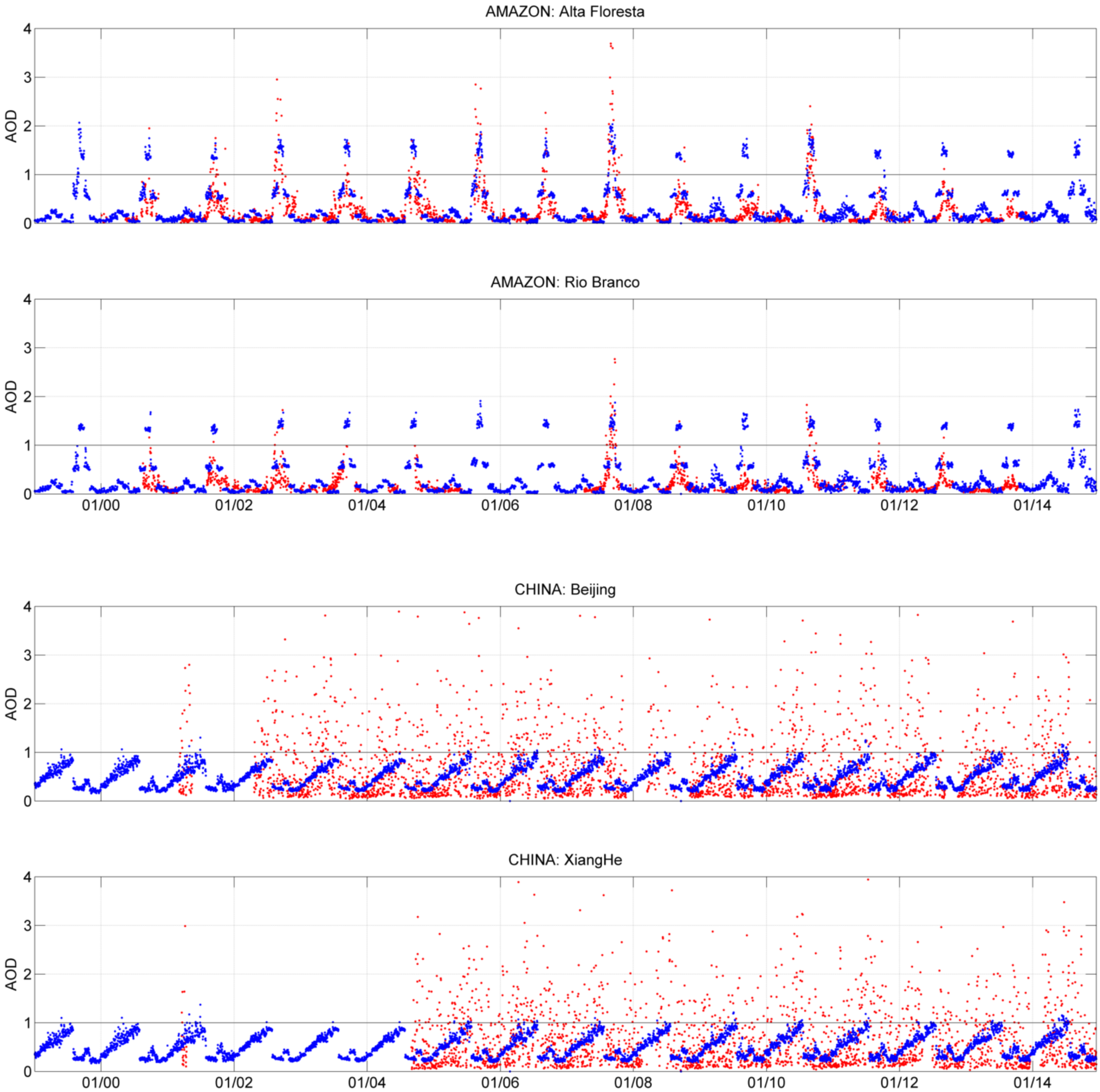
It needs to be discussed in how far the method does only correct for absorbing aerosols (excluding AI < 0) and how this will affect the AOD and albedo values. Also the impact of the difference between total AOD from MODIS and absorbing-aerosol AOD from AI in the regression of the method needs discussion.

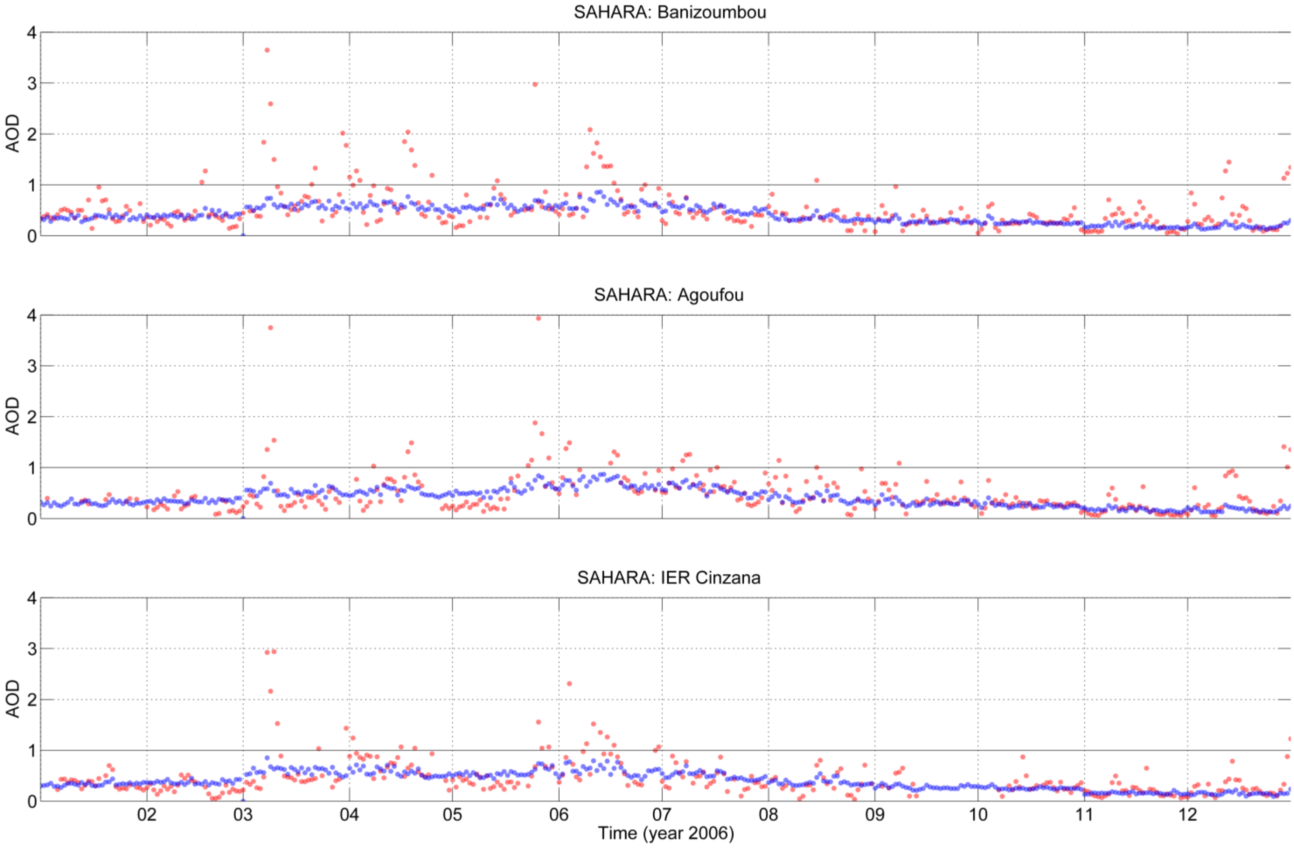
This matter is now assessed in the manuscript, P. 4/L. 8-10, P. 10/L. 27-30. Even though the AOD values from the constructed AOD time series are calculated by using the absorbing AI values, they are not exactly absorbing AOD values, because the regressions are derived for AI and total AOD. For this reason, it is not so problematic to compare the constructed AOD time series to the total AOD from MODIS.

The evaluation is too much done with global / zonal and long-term averages – the added value of the AOD daily maps lies in the spatial and temporal patterns for the atmospheric correction. Also, providing those daily maps contains the risk of introducing additional noise into the datasets – this needs to be assessed, at least with exemplary studies.

The evaluations could have been done more from daily values, that's true. Below there are some example figures. In the first three figures, the AOD data from the seven AERONET stations (red) are compared to the constructed AOD time series data (blue). The data looks very similar with the monthly mean figure (Fig. 17 in the manuscript). The fourth figure is similar as the first figure, but there are data only from one year, 2006, to show better the daily values. The constructed AOD values have clearly less variability than the in situ AOD values.
The CLARA-A2 SAL product is reported in pentad and monthly means. The AOD daily maps may introduce additional noise, but the pentad and the monthly means of albedo product will filter some of the noise away. We added a note of this to the new section in the manuscript, section 7 titled as Discussion and conclusions.




AMAZON: Alta Floresta
AMAZON: Rio Branco
CHINA: Beijing
CHINA: XiangHe
AOD
0
1
2
3
4
01/00
01/02
01/04
01/06
01/08
01/10
01/12
01/14



On what spatial and temporal scales would one expect to reproduce realistic aerosol variability, where one you expect to smoothen them?

The daily values from three AERONET stations in Africa over one year, 2006, are compared to the constructed AOD values from the same locations, and the results are shown in the figure above. The constructed AOD have clearly less variability than the in situ AOD values, but the constructed AOD values are not smooth over the year. The variability in the constructed AOD values is not as distinct as it is in the in situ AOD values, but it is still faithful to the main features of the in situ AOD.

Further comments: To make up for longer text, some of the figures are not necessary and can be deleted or combined. The authors should consider reducing figures: 1 (describing the main stability over long time but regional seasonal cycles in the text will suffice), 2 (one of the two maps is sufficient, aren't they adding up to 100% ?), 3 (can be explained in text), 4 (better describe in text the principles for building the sub-classes), combine fig 8 and 10 into one flow chart with optional boxes; and tables: 5 (can be explained in 1 or 2 sentences in the text).

Figure 1: Removed and the content are now explained in the text, P. 4/L. 2-5
Figure 2: The positive values map is removed and the text edited accordingly, P. 4/L. 5-8.
Figure 3: Removed and text added, P. 4/L. 24.
Figure 4: Because the division to the subclasses is made somewhat subjectively, the figure makes it easier to grasp that how the division is done.
Figures 8 and 10 combined.
Table 5: Removed the table and the contents are now explained in the text, P. 13/L. 26-28, P. 14/L. 6-7.

The authors should make clearer in the title and text that they are discussing a time series of global maps (i.e. with regional AOD variability) to distinguish from a global averaged time series. This will then support the added value discussion of providing spatial information for the atmospheric correction.

Mention of this matter is added to the text, P. 2/L. 13, and we modified the title of the manuscript.

Spatial resolution of all datasets needs to be provided.

The spatial resolutions of the TOMS, OMI and MODIS data sets are collected in the Table 2, and the resolutions of the land use data are added to the text, P. 3/L. 23-24

English usage needs to be improved by involving a native English speaking person; e.g. articles are often miss-used, the word "manifolding" should be replaced (several times).

All detailed suggestions of the reviewers have been carried out. However, it is not easy to satisfy all reviewers/readers simultaneously, because the use of English language has so large variation. Namely, the language of the original manuscript had been officially checked by a native English speaker who is doing language checking professionally. We shall give the feedback of the reviewers to the language checking company.

Reword "TOMS- homogenize" (p. 7 / l. 13).

We changed the word "TOMS-homogenize", P. 7/L. 28-29, P. 8/L. 1, 17, 19

There are a number of vague statements which should be made more precise / quantitative; e.g. "sufficient" (p.1 / l.7), "long enough (p. 1 / l.16), "a little bit too coarse (P. 3 / l. 23), "by a little" (p. 4 / l. 15), "some local inspections" (P. 5 / l. 15), "not so much" (p. 5 / l. 23), : : :

These statements are made more precise.

Detailed comments: The last paragraph of section 1 (structure of the paper) should be shortened to only give one main heading for each section; further detail is not needed here.

The paragraph is shortened as requested.

Section 2.1: EP-TOMS is not used and therefore needs not to be discussed at all.

The AI data from EP-TOMS is used for the time period 1996-2001, Subsection 2.1: P. 3/L. 4, Subsection 4.2.

P. 3 / l. 17: MODIS AOD is a retrieval, not an estimation (higher accuracy).

The word is changed accordingly in the manuscript.

P.3 / l . 21 -24: which land cover dataset do you use?

As mentioned on P. 3/L. 21-25, the used land cover data sets are the AVHRR Land Use Classification, where the spatial resolution is 1 degree, and Global Land Cover 2000.

Section 3.1: where do tau-UV and alpha come from?

The tau-UV comes from the OMAEROe product, AOD at [342.5nm, 388nm, 442nm, 463nm, 483.5nm] and the alpha is calculated from these values. The Ångström exponent is calculated once from each wavelength pair (ten combinations altogether) and these exponents are then used to estimate the AOD at 550 nm by using the AOD values at suitable wavelengths (again, ten AOD values at 550 nm altogether). Text updated P. 5/L. 6-9

End of section 3.2 and later on: you mix up “areas” and “classes” – please be consistent to avoid confusing the reader.

The word “area” is now replaced by the word “subclass” whenever necessary.

P. 5 / l. 15/16: I do not understand these statements – please explain what you mean.

The subclass division is used for helping for example in deseasonalization and when determining the regression coefficients. Some land cover types (such as deserts) are related to certain aerosols. Text edited, P. 5/L. 26

Section 3.2: this is very important to discuss the limitations / assumptions, but needs extension

We added more text of the sensitivities of the AI to other parameters to the manuscript, P. 5/L. 31-32

P. 5 /l. 26: give minimum and maximum number of pixels;
Added as requested, P. 6 /L. 7-8.

also l. 28 Fig. 6: better show results with AI * cos (theta), since you use this quantity; also better colour bar should be used to show variability where most data points lie (e.g. between 0.5 and 0.8)

The figure in question changed along with a better colour bar as requested, and the text updated as well (P. 6 /L. 9-10).

P. 5/ l. 29/30: Correlations of 0.5 are still quite weak – I would thus be more cautious and rather conclude, that the method can only be used for parts of the dataset to construct reliable AOD

It is true, the correlation of 0.5 is weak. The text is edited (P. 6/L. 12-14) as requested by concluding that the method provides probably more accurate AOD values in certain areas of the globe.

P. 6 / l. 10: I do not understand why you need the ordering – isn’t this just the weighted average?

The ordering indicates the process how the weights are obtained. Text clarified, P. 6 / L. 26.

P. 6 / l. 13: a vector of what?

A vector where each value is added multiple times. Text edited, P. 6 / L. 28.

P. 6 / l. 18: explain “after additional restrictions”

Text edited P. 7 / L. 2, additional restrictions referenced to restrictions of AOD < 1 and SZA < 70.

P. 6 / l. 22: explain how you divide them

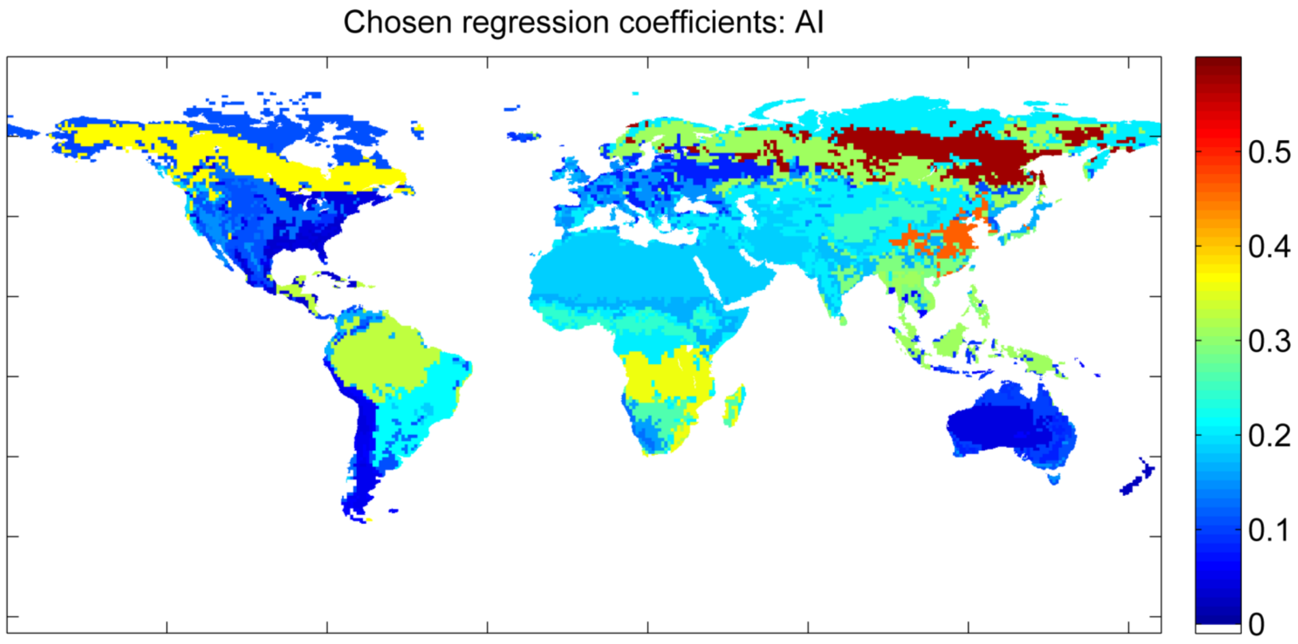
Word “divided” changed to the word “processed.” P. 7/ L. 6

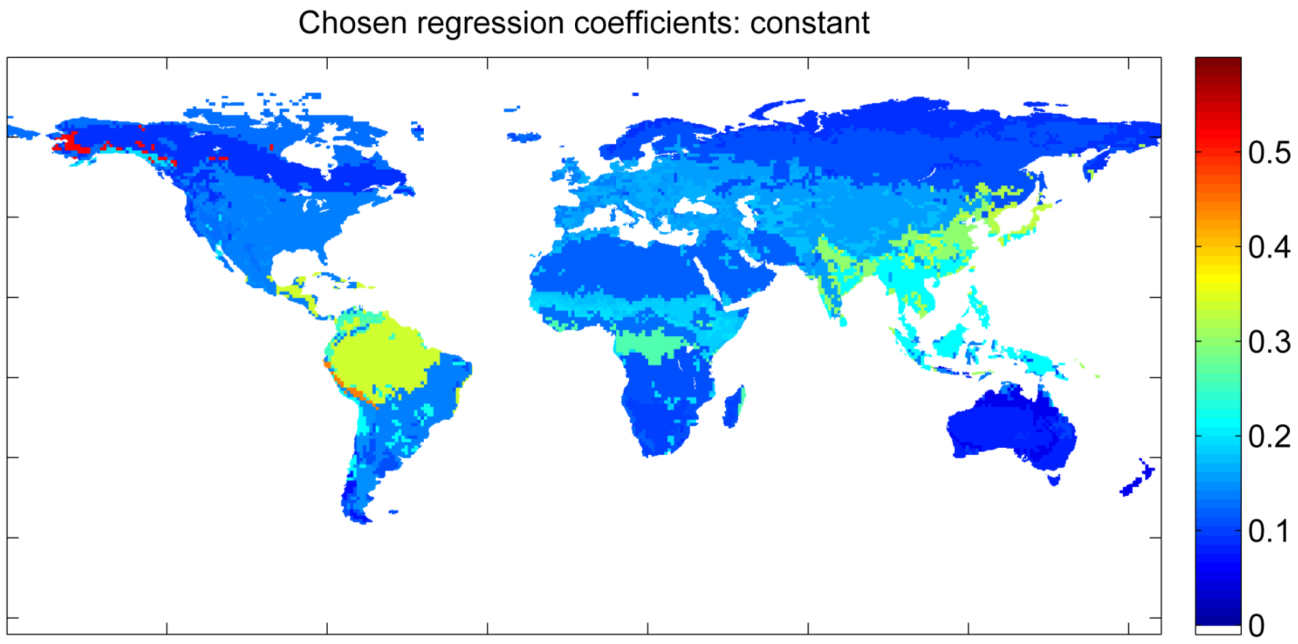
P. 6 / l. 28/29: I do not understand this sentence; is the simplest also the best one or at least equally good as others?

The simplest one was the best model of those which were equally good. Text edited P. 7/L. 12-13

I suggest to show one example time series over those steps to illustrate better what you do; also a map of regression coefficients could be illustrative

Below are the maps of chosen regression coefficients.





Start of section 4: motivate, why you need two different approaches

Text added P. 8/L. 8-10

P. 7 / l. 30: how exactly do you treat cases with AI outside the range [0.5, 4.5]? omit, set to 0.5 and 4.5, respectively, : : :

The AI outside the range [0.5, 4.5] are omitted, text added P.8/L. 16

P. 7 / l. 29-31: why do you use two steps of spatial regridding?

The data from the OMI instrument have different resolution that the data from the TOMS instrument (Table 2). The The data are homogenized by using the resolution of TOMS and it is done, because we want to avoid the difference in the data when using the AI data from two different data sets. The second spatial regridding is done to change the data back to the original resolution which is the resolution of CLARA-A2-SAL. Text added, P.8 / L. 15-18

Fig. 9 needs discussion: many values too high (e.g. Scandinavia, California, Siberia, SouthEastAsia, Tibetan plateau, Himlaya, : : :), mountains come out, compare to OMI AOD retrieval map

We added more details of about the figure, P.8 / L. 28-32, P.9 / L. 1-4, and added an additional figure of about the absolute differences between the constructed AOD and MODIS-AOD. We compared the constructed AOD from the May 2005 to the MODIS-AOD from the same month and year. The accuracy requirement is not included for the mountains in CLARA-A2 SAL product, so it doesn't produce a problem if the AOD time series has difficulties in those areas.

P. 8 / l. 15-18: I am not convinced why you use 3 years before and after the gap – motivate and explain

One year is not enough to catch the variability of the AI, so three years are a safer solution. This gap filling method is the most simple and robust solution for the time being. Text added P. 9/L. 13

P. 9 / l. 8: if the annual cycle was the same overall years, then you could produce one long-term climatology dataset, but there are intra-annual variations, one potential strength of your dataset

This is true, text edited P. 10/L. 8-10.

P. 9 / l. 9: Tropic of Capricorn is the Southern – you want to point to the Northern (sub-) tropical maximum over the Sahara latitude?

No, Tropic of Capricorn is right. The sentence was unclear, the highest AOD are due to the high AOD values from the Amazon area in September. Text is modified to be more clear, P. 10/L. 11.

P. 9 / l. 12-22: this is not very clear (why should the more accurate MODIS dataset have less seasonality)

MODIS-AOD have less seasonality in a sense of global monthly means. The standard deviations vary more (the grey area behind the lines). Text modified P. 10/L. 22-23.

P. 9 / l. 26-28: a difference of 0.3 is very large (given mean global AOD over land of 0.2); also next paragraph: you should talk of large differences, but say better, that they are still smaller than with assuming a fixed AOD=0.1

Yes, it is true that the difference of 0.3 is quite large. We modified Figure 11 by adding the absolute differences of zonal monthly means between OMI-AOD and the constant AOD value 0.1, where one can easily see, that the differences are smaller between OMI-AOD and the constructed AOD time series compared to the differences between OMI-AOD and the constant value. Also, text edited accordingly, P. 11/L. 6-11.

P. 10 / l. 3 onward: please state in how far the 3 example classes are representative for your analysis of all classes. Do they show best, worst or typical results?

The results in Amazon subclass are weaker as expected compared to the other two subclasses, especially in the season SON, because the OMI-AOD vary a lot, from 0.2 to over unity and the linear regression cannot predict that well. In the subclass covering the Sahara and the Middle East, the aerosols are typically dust and there the linear regression using AI provides more accurate AOD. The Mainland Southeast Asia subclass is something between Sahara and Amazon. It is more typical subclass compared to all the others and the results are hence more typical. Text added as requested, P. 12 /L. 11-17.

P. 10 / l. 32: please add AERONET reference: Holben, B.N.; Eck, T.F.; Slutsker, I.; Tanré, D.; Buis, J.P.; Setzer, A.; Vermote, E.; Reagan, J.A.; Kaufman, Y.J.; Nakajima, T.; et al. AERONETâAˇTA federated instrument network and data archive for aerosol characterization. Remote Sens. Environ. 1998, 66, 1–16.

We added the reference, P. 12/L. 20.

p. 11 top: typical satellite AOD validation uses a window of 50x50 km2 for spatial matching; you need to discuss whether you are not creating artificial variability on pixel level

The authors did not understand this point. In this section we discuss the problem of comparing point wise in situ measurements and large satellite pixels, which we solve by comparing a distribution of several in situ measurements within the large pixel of the satellite. We are not doing any processing in this section that could cause pixel level variation.

P. 11 / l. 10 onward and fig. 10+11: use more specific names, not the continents, where the small test regions lie in - this is misleading

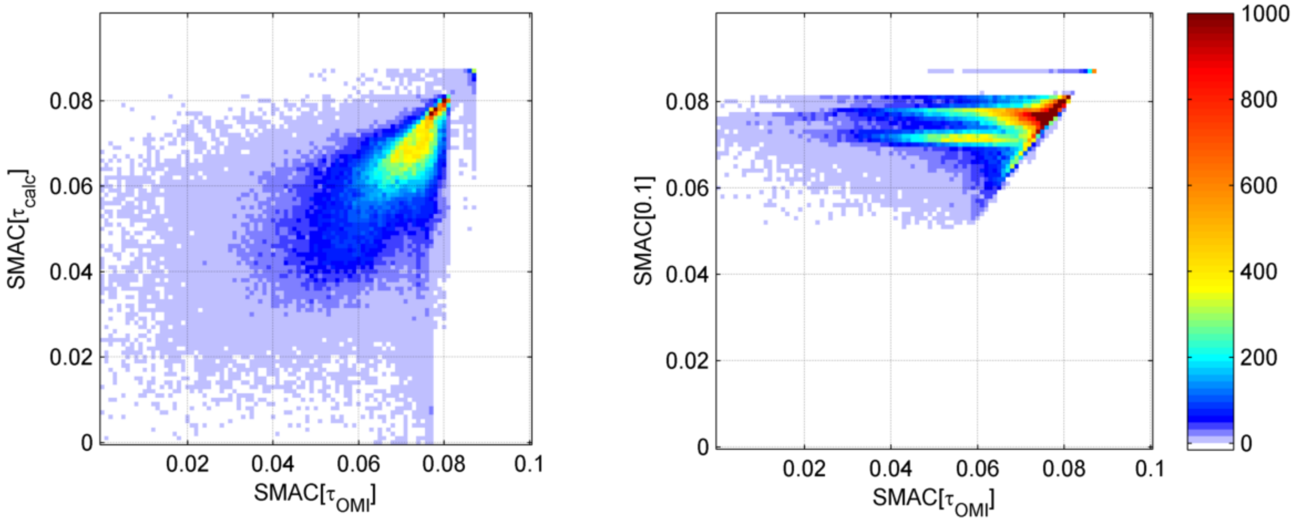
The names are changed to more specific ones, i.e. P. 12/L. 32-35, P. 13/L 1-2.

Fig. 22: better show absolute differences, not relative – otherwise you highlight larger relative errors over dark surfaces

Figure updated as requested, also text changed, P. 13/L. 22-23, 28-32.

Fig. 22: figure title should be “relative difference of corrected reflectance values” (“magnitude values“ is inappropriate terminology); better show scatter plots; I would prefer to see absolute values of reflectance differences; use better colour bar: large areas go from pale yellow to dark yellow (become worse, hard to be seen), some areas become better (from dark red to pale red); I would distinguish negative and positive values

Figure updated as requested (from relative differences to absolute differences), also text changed, P. 13/L. 22-23, 28-32. We decided not to do scatter plots for the manuscript, because it won’t show where the differences are spatially. Negative and positive absolute differences are shown in Figure 20. The scatter plots of the means of the simulated SMAC values are shown below.



P. 12 / l. 14: can you draw a quantitative conclusion rather than saying that reflectances tend to be higher?

The quantitative conclusion added, P. 14/L. 22.

Fig. 11: why do you not make a scatter plot of AODs?

The point of that figure is to show how homogeneous the constructed AOD time series is.

Add discussion in the text: El Nino Indonesia fires can be seen in 1997, lat 60N much too high, Sahara under-estimated/ biomass burning over-estimated, : : :

We added more discussion about the zonal mean figure (Figure 8 in the revised manuscript) to the text, P. 10/L. 9-15.

Fig. 12: why are there several curves for each category?

There is a curve for each year (2005-2014), added the year information to the figure caption.

Fig. 13: global mean AOD over land is 0.2 - so you cannot make it that crude - you have extreme differences + and - 0.7 or so; better show the range -0,25 to 0.25 and exclude the other regions

The point of the figure is to show where the large positive AOD values occur in the calculated AOD time series in relation to MODIS-AOD. Below is the same inspection, but now only the differences from the range [-0.25, 0.25] are shown.

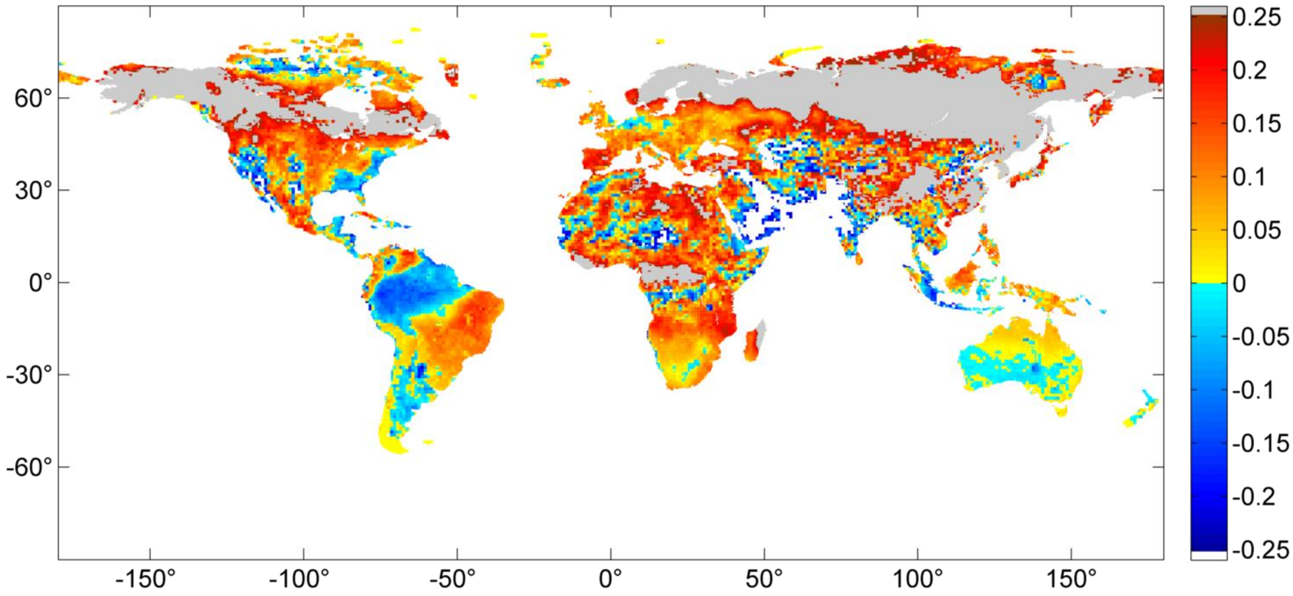


Fig. 18: you show partly very large differences: peaks, distribution shapes, double peaks; how can AOD be >1 with your method? Use a better-suited x-axis (e.g. 0-1)

The x-axis of Figure 18 is updated as requested. In the regressions the AI and AOD data were limited to the values where AOD < 1, but in the calculation of the daily AOD maps from the AI data there were no such limits.

Fig. 19 / text: discuss whether those 6 regions are suited to grasp all global variability of aerosol and surface conditions

The chosen regions do not cover all the possible aerosol scenarios, but they offer enough variability, that we can assess the use of the constructed AOD time series in the atmospheric correction. Added text to the manuscript, P. 13/L. 2-7.

Fig. 21: state in text partly significantly wrong seasonality (thus limiting the capabilities for atmospheric correction to use for assessing seasonal changes)

Text changed accordingly, P. 13/L. 13-17.

Fig. 22-24: which wavelength or band refletances?

The used wavelength band was 0.725-1.000 µm, text added: P. 13/L. 19-20.

Fig. 24: why not again year 2010?

That's true, it should have been year 2010. The figure in question is now updated by using the data from the year 2010.