

ANSWERS TO REFEREE #2

First of all, we thank Referee #2 for these insightful remarks and comments on this topic. The comments have been addressed below and prompted changes to parts of the text. Our responses follow the reviewer points in italics

'Comment on egusphere-2025-5950'

In this study, the authors collate sunshine duration measurements going back in some cases to the early 20th century with reanalysis data to estimate the change in aerosol optical depth at observation stations worldwide. Their results suggest that the period 1900-2015 can be divided into periods of increasing and decreasing trends in aerosol optical depths. They especially highlight the period 1900-1959, which they find associated with a decreasing trend.

The paper is well written, and the data collection efforts are very impressive. However, the method and its interpretation do not live up to the objectives of the study. I therefore recommend major revisions before publication to give a more realistic account of the insights that could be taken from the analysis, as commented below.

We thank you for your positive and constructive comments on the manuscript.

Main comments:

- *Motivation of the study: The paper gives the “early brightening” as its main motivation, claiming that proving or disproving that early trend could distinguish between different representations of aerosol history (Page 4, lines 16-19). But it is not that simple. The question is not only whether any claimed early brightening is real, but what caused it. And opposing the method of Smith et al. (2021) to other representations is reductive. First (page 3 lines 15-18), most global aerosol-climate model simulations prescribe aerosol emission inventories, rather than aerosol forcing. Second, although it is true that AeroCom-style simulations often take the difference between present-day and preindustrial simulations, there are time-dependent simulations too, hindcast simulations covering 1980-present or CMIP simulations covering 1850 to present (and future according to agreed scenarios). And the challenge in interpreting long aerosol time series is not only to understand the history of different aerosol compounds (Page 4 line 5) but also to separate forcing (external influences, in this case anthropogenic and volcanic emissions) from feedbacks (in this case, changes in anthropogenic and/or natural aerosol emissions that happen because climate has changed). Reconstructing past aerosol optical depth trends would be welcome but does not address most of those challenges. The authors should be realistic when presenting the strength of their line of evidence in the introduction.*

Thank you for this important comment. We agree that our original Introduction sounded too strong about what can be concluded from our BAOD reconstruction, and that the “early brightening” question is not only about whether a signal exists, but also about its causes. In the revised manuscript, we have softened the motivation and conclusions. At the same time, we

emphasize that it remains important to extend observation-based aerosol-burden information further back in time. The main contribution of this work is a long and unusually large station compilation of SD-based BAOD proxy time series. Obviously, the very limited number of observing stations that existed 100 years ago is not sufficient to generalize the regional results we obtained to the global scale. In sum, new elements are now added in the Introduction as follows:

“A major source of uncertainty in aerosol—climate interactions is related to the limited availability of long, observation-based records of aerosol optical depth (AOD), especially prior to the satellite era. Ground-based sunphotometer AOD observations (e.g., AEROSOL ROBOTIC NETWORK abbreviated AERONET) provide high-quality reference measurements but began only in the early 1990 (Holben et al., 1998), at which time the network counted only a few stations. In parallel, the earliest satellite-based AOD products extend back only to the late 1970s, with rough AOD estimates obtained with retrievals from the Total Ozone Mapping Spectrometer (Torres et al., 2002).

Longer, observation-based proxies for AOD are valuable for multiple purposes, including characterizing historical variability and providing benchmarks for the quantification of aerosol evolution in global climate models. Because aerosols modulate radiative transfer through the atmosphere, the scarcity of long AOD records also limits our ability to interpret multi-decadal changes in surface solar radiation, including the reported dimming/brightening.

Long-term changes in surface solar radiation (“dimming/brightening”) reflect contributions from aerosols, clouds, water vapor, and circulation-driven variability. Because of their complexity, the relative roles of these factors remain challenging to disentangle from the available observations alone. Nevertheless, long time series of aerosol burden can provide valuable complementary evidence by documenting the timing and spatial coherence of changes in AOD. In particular, the early-twentieth-century brightening signals that were reported over some regions motivate an increase in effort to extend observation-based aerosol burden proxies further back in time, while recognizing that burden changes alone do not uniquely determine causation.”

“Although this effort is unprecedented in terms of number of stations investigated in many parts of the world, and in terms of duration of the uncovered time series of the distant past, it is acknowledged that the available datasets only cover a small part of the world. Obviously, the farther through the past an empirical investigation is conducted, the lower number of observing stations is available and the higher the risk is for incomplete or questionable data. Considering this important limitation, it is emphasized that the present regional results of the early 20th century are fragmentary and cannot be truly generalizable to the global scale.”

- *Method: The authors seem to trust data assimilated products just for the reason that they assimilate observations (Page 7 lines 21-23, Page 8 line 17). But there is very little data to assimilate in many regions, in fact in most regions for the early part of the time series. The influence of data assimilation on water vapour and ozone will vary massively over the 20th century. Indeed, long ozone time series are more difficult to get than aerosols! So one would expect uncertainties to explode the further back one goes. It is difficult to trust the authors' results for the first half of the 20th century.*

Thanks for this remark. It is certain that uncertainties are expected to increase when going back to the early 20th century because of the decreasing number of observations used for data assimilations. However, the sensitivity analysis, as shown in Fig. 3b from our previous study, demonstrates that total column amount of ozone (TOC) does not have a significant impact on the estimated AOD (Wandji Nyamsi et al., 2020). Thus, the accuracy of long TOC time series is only a second-order or third-order concern here. Therefore, we added a sentence at the relevant part of the text as follows:

“Overall, it is emphasized that l_o has only a very small impact—if not negligible—on the accuracy of the retrieved AOD (Wandji Nyamsi et al., 2020), therefore the selection and accuracy of its data source are of limited importance.”

In comparison, the total column amount of water vapor (TWV) is a much more important quantity. As additionally developed in the previous study just mentioned, the proposed expected error envelope for the estimated AOD takes into account all possible sources of uncertainty, including that in TWV. In addition, we have acknowledged in the relevant part of the text as follows:

“ERA-20C is produced with a modern data-assimilation system; however, because the density of assimilated surface observations decreases markedly backward in time, uncertainty in model-derived fields is expected to be larger in early decades and in sparsely observed regions.”

- *Results: Clearly, despite the authors’ impressive efforts, data is scarce and trends are made from a very different number of stations between periods and regions. Early trends come from 5-6 stations in Europe – difficult to conclude on an “early brightening” in those conditions. In addition, trends are often caused by local influences, and I welcome the effort of the authors to find those local causes, although the analysis remains often circumstantial (e.g. Page 21 lines 3-6, Page 23 lines 2-5). What can we conclude from such a limited and local dataset? It seems clear that (1) taking the average of trends at different stations to get a continental/regional average is not justified, and (2) the conclusion of the paper should be “we tried hard, but it is not possible to get good trends from our method before the satellite era”.*

Regional trends are not computed as the average of local trends but rather by using bootstraps. We acknowledge that the limitation caused by the spatial scarcity of stations for a given period. Nevertheless, we have intentionally tried to provide a regional trend value. We have rewritten the relevant part of the text as follows:

“For such a period with a very limited number of stations, quantifying a regional trend is highly challenging, mainly because of the spatial scarcity of stations. It is thus to ascertain the effective representativeness of the trend value over that region. This limitation is duly acknowledged, as also mentioned in Section 1. Nevertheless, a tentative regional trend value is provided here, as well as for other similar cases.”

Other comments:

- *Page 3, line 7: Changes in aerosol load are not only due to anthropogenic emissions - they can also be caused by changes in natural aerosol emissions from climate feedbacks.*

We have clearly mentioned this possibility in the Introduction and other relevant parts of the text. For instance, we have added a sentence as follows:

“although it is possible that a part of the observed changes in aerosol load has been caused by changes in natural aerosol emissions from climate feedback.”

- *Page 3, line 24: And the SO₂/BC/OC history might be wrong too!*

Thanks for this remark. The relevant part of the text has been rewritten more precisely.

- *Page 4, line 23: “characterising” – yes, but indirectly.*

We have rewritten the relevant part of the text.

- *Page 4, line 25: AERONET measurements have low uncertainty, but sampling is an issue since AERONET cannot measure AOD in cloudy conditions.*

We have rewritten the relevant part of the text.

- *Page 5, line 15: Romania and Spain are in Europe.*

Rewritten accordingly.

- *Page 7, line 29: It would be good to support the claim of “high quality”.*

We agree and have added a few relevant references demonstrating the high quality of OMI’s total column amount of ozone as follows:

Balis, D., Kroon, M., Koukouli, M. E., Brinksma, E. J., Labow, G., Veefkind, J. P., McPeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, *J. Geophys. Res.*, 112, D24S46, <https://doi.org/10.1029/2007JD008796>, 2007a.

McPeters, R., Kroon, M., Labow, G., Brinksma, E., Balis, D., Petropavlovskikh, I., Veefkind, J. P., Bhartia, P. K., and Levelt, P. F.: Validation of the Aura Ozone Monitoring Instrument total column ozone product, *J. Geophys. Res.-Atmos.*, 113, D15S14, <https://doi.org/10.1029/2007JD008802>, 2008.

Antón, M., López, M., Vilaplana, J. M., Kroon, M., McPeters, R., Bañón, M., and Serrano, A.: Validation of OMI-TOMS and OMI-DOAS total ozone column using five Brewer spectroradiometers at the Iberian peninsula, *J. Geophys. Res.-Atmos.*, 114, D14307, <https://doi.org/10.1029/2009JD012003>, 2009.

- *Page 10, line 6: But doesn't that remove inhomogeneities caused by rapid industrialization?*

We acknowledge that it might happen that our homogeneity algorithm (HA) yields rejection of the null hypothesis for a change point caused by rapid industrialization. This is a delicate matter. It is understandable that automatic and accurate selection of homogeneous time series from a single algorithm is very difficult due to many possible reasons or situations. In the mentioned case, our HA becomes more restrictive by selecting lesser time series and could offer more confidence in the fact that the time series is homogeneous.

The performance of the HA for these cases is beyond the scope of this study. That is still relevant, though, and thus could further be investigated in future dedicated work when/if a reasonable time series of meteorological sites is available with their complete respective metadata information.

- *Page 10, line 26: It would be more interesting to test an inhomogeneous time series, trying to explain change points that are not due to volcanism.*

Thanks for this suggestion. We have found time series of stations in Japan and Europe where detected inhomogeneities were caused by, for instance, the station relocation or instrument change at various times, well in agreement with metadata information. So, it would have been possible to present such results in addition to the analyzed time series for Potsdam, as in our study. The paper is already long and with many significant details here and there. Because of that, we cannot present additional results for many case studies. We decided to present results for Potsdam because of its long time series, going back to the late 19th century. Moreover, it is certainly one of the most known and investigated meteorological stations, whose data time series have been scrutinized in numerous papers of the literature. In addition, since the other reviewer agrees with our view and is satisfied with the proposed illustration, we have kept it as is.

- *Page 13, line 24: That number varies with the solar cycle in a way that is well determined. It seems unnecessary to neglect that variation.*

Although spaceborne observations of total solar irradiance (TSI) support in providing a nearly continuous record of the Sun's output, the TSI time series still lacks temporal coverage or has significant uncertainty because of conflicting observations from different sensors, especially before 1985 (Gueymard, 2018). Hence, reconstructions of pre-1985 satellite-based TSI time series present unavoidable uncertainties and inhomogeneities. All these uncertainties are somewhat partially taken into account in the derived expected error envelope for estimated AOD. In any case, the daily variations in TSI caused by short-term solar activity can be larger than those induced by the long-term solar cycle (see, e.g., Figs. 1–3 in the aforementioned reference). Consideration for smooth effects caused by the solar cycle would introduce variations of only $\pm 0.1\%$ in TSI. Daily variations of up to $\pm 0.3\%$ are possible during high-activity periods, but even the latter modulation is significantly smaller than the overall uncertainty in retrieved BAOD.

Gueymard, C. A.: A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations, *Solar Energy*, 168, 2–9, <https://doi.org/10.1016/j.solener.2018.04.001>, 2018.

- *Page 14, lines 15-18: The original method of Wandji Nyamsi et al. (2020) was applied to measurements that rounded down to a cloudiness of 0 okta. In the present study, the rounding is up to 1 okta, which if it happens to be in the direction of the Sun would cause much contamination of the results. Are those comparisons still valid?*

Thanks for noticing. The relevant part of the text was misleading. We have rewritten it.

- *Page 15, line 15: Note that the dimming and brightening periods were not uniform across surface flux stations either. Those trends are fundamentally local.*

We recognize that these phenomena are fundamentally local. We have rewritten the relevant parts of the text accordingly, starting with the Introduction.

- *Page 17, line 12: “two thirds” is a complicated way to say “four” here.*

Corrected as requested.

- *Page 17, lines 18-20: Aren’t those stations the same that were just discussed. How do that generalise things?*

To avoid any confusion, the first part of the sentence has been removed.

- *Page 19, line 12: Here, four stations is one third of all stations, so the “only” sounds rather optimistic.*

This has been corrected accordingly.

- *Page 23, lines 12-13: Perhaps not from anthropogenic sources (but then why not?) but there could be other influences.*

We have made things more precise by adding “at least related to anthropogenic sources” at the end of the relevant sentence.

- *Page 24, line 6: Typo “for ass sites”*

Done.

ANSWERS TO REFEREE #3

First of all, we thank Referee #3 for these constructive remarks and comments on this topic. The comments have been addressed below and prompted changes to parts of the text. Our responses follow the reviewer points in italics.

General comments

Overall, the paper is of good quality and well written. It addresses an interesting topic with a clear structure and relevant analysis.

We appreciate the hypothesis connecting the retrieved BAOD trends with natural and socio-economic factors.

The findings are interesting, particularly the contrasting BAOD decrease in Japan and the early brightening signal in Europe.

We thank you for your positive overview on our work.

However, some parts of the methodology could be improved for better readability and clarity:

- *The methodology sections are a little difficult to follow in places.*
 - *In Section 3.1, especially the second paragraph discussing the six-year analysis (?) and the reduced time series, the explanation is not fully clear. It is appreciated that Section 3.2 illustrates the homogeneity tests. However, “selecting a homogeneous time series” does not sufficiently explain how the time series is actually reduced in cases of detected inhomogeneity.*

Thanks for this remark. We have rewritten the relevant part of the text in more detail as follows:

“The median year that is categorized as a change point for one analysed variable is retained only if that year is not amongst potential years of previously listed climatic factors. Then, the four homogeneity tests are applied anew on each of the five remaining variables, thus resulting in a total of six analysed variables as mentioned earlier.

Let Y_i denotes the retained median year of the i -th analysed variable, $\forall i \in \{1, 2, \dots, 6\}$. Y_i is collected and gathered in a set $\{Y_i\}$. At this step, two situations exist. First, the set is empty, i.e. no retained years were found; the SD time series is thus considered homogeneous. Second, the set is not empty so that an action is executed as follows. The number of occurrences of Y_i is counted. If Y_i or Y_i+1 had two occurrences at maximum, the SD time series is considered homogeneous too. For at least three occurrences in total, the SD time series is considered inhomogeneous. In that case, the SD time series is split into two sub-SD time series from the most recent year (reaching up to three occurrences) (1) backwards and (2) onwards. Only onwards sub-SD time series as the reduced time series, is retained because, following the development of ground- and satellite-based instruments, recent data are generally more homogeneous than old data.”.

- *In Section 3.3, without prior knowledge of the method or reference to external resources cited in the text, it is not straightforward to grasp the basic principle of how sunshine duration (SD) measurements translate into aerosol optical depth (AOD)—which is the core aim of the paragraph titled “estimating daily AOD from SD measurements.” For example, it would be helpful to clarify that estimations are performed at sunrise/sunset using the heliograph burning threshold.*

Thanks for this remark. We have explained the concept of the method in a step-by-step manner in the relevant section as follows:

“The concept of this method is elaborated in a step-by-step approach as follows. In the context of historical observations, SD is the total length of the burned trace on an appropriate card of the Campbell-Stokes heliograph. Each card represents data over one specific day. The burned trace appears when DNI is greater than the burning threshold of the card. During a cloudless day, — the sky condition of interest in this paper—the burned trace is expected to be continuous; it starts not long after sunrise and stops just before sunset. Assuming the diurnal uniformity of daily ozone and water vapor conditions, the SD-derived AOD can also be considered diurnally uniform. Consequently, SD can be transformed to an equivalent hour angle (ω) such as $\omega = 15(\frac{SD}{2})$ in degrees. From an hour angle, in turn, with the latitude of the station and solar declination angle at a given day, both accurately known, the solar zenith angle (Θ_s) can be computed. Then, Θ_s is used to compute air masses and all atmospheric transmittances using the mathematical equations explicitly given in Appendix A. These equations are used to finally derive AOD as described hereafter. With this approach based on the computed Θ_s , the SD-derived AOD is estimated at those instants close to sunrise/sunset, when the burned trace becomes visible/extinct, respectively.”

Specific comments

- *Page 2, line 6: “Concentrations” is used for the sake of conciseness. However, aerosol concentrations cannot be derived straightforwardly from AOD series, as aerosol scattering properties do not necessarily remain constant over the analyzed time period.*

Thanks for this remark. To align ourselves with the initial wording, we have replaced “Aerosol concentrations” by “Mean-annual BAOD” in the relevant part of the text.

- *Page 7, line 27: The subsection title suggests that only observational data is used for ozone, but the paragraph describes the use of model data.*

Thanks for this remark. We have rewritten relevant parts of the text accordingly.

- *Page 9, line 15: “Daily scale”? This should likely be “yearly scale,” as the analyzed time series rely on yearly averages/standard deviations.*

Thanks for this remark. The correction has been done.

- *Section 4 (Results and discussion):*

- *First period: 1900–1925*
 - *It is surprising that no impact from the Santa María eruption (1902) is apparent in the retrieved BAOD at Zagreb-Grič, whereas a significant increase is observed at other European stations.*

Thanks for noticing this apparent contradiction. One likely reason for this situation is the possible low quality of available daily cloud cover observations (TCC) or their daily mean ($\overline{TCC}$) before 1950. Such old TCC data series are obtained from synoptic observations made by human observers, commonly with a temporal resolution of 3 hours or coarser. Hence, the average can be computed from only a very few observations during the day. For instance, four observations are used if synoptical cloud cover data is available at 00, 06, 12 and 18 UT or eight observations if available at 03, 06, 09, 12, 15, 18, 21, 24 UT.

Both Zagreb-Grič and Potsdam TCC time series belong to European Climate Assessment & Dataset (ECA&D). Potsdam's $\overline{TCC}$ is computed based on three observation times (07, 14, 21 UT) before 1950, whereas Zagreb-Grič's $\overline{TCC}$ is computed from observations over an unknown interval as reported in the data description from ECA&D. This uncertainty can realistically alter the quality of $\overline{TCC}$ in some cases. Nevertheless, a peak value of 0.3 from the estimated daily BAOD at Zagreb-Grič has been observed on March 27, 1903. That occurred within about one year from the volcanic eruption Santa María in 1902. Therefore, we have added one sentence to the relevant part of text as follows:

“Impacts on BAOD of massive volcanic eruptions might not be perceptible at yearly scale but can be seen at least at daily scale. This is, for instance, the case for Zagreb-Grič, where the impact of Santa María is not detectable in the annual BAOD mean time series in Fig. 3. At daily scale, however, an elevated peak value of 0.3 has been observed on March 27, 1903, for example.”

Technical corrections

- *P24l6: for ass sites over?*

Thanks for this remark. It has been corrected accordingly.

- *P28l12: thstudied → the studied*

Thanks for this remark. Done as requested.