ANSWERS TO REFEREE #1

First of all, we thank Referee #1 for these constructive remarks and comments. The comments have been addressed below and have been taken into account for revising a part of the text following recommendations of the referee. The responses to the reviewer points are below after the reviewer points that are in italics.

Comment. Information of long term series of atmospheric data is an essential task in order to understand various processes going back to more than 20-30 years. Aerosol optical depth investigated here, is a parameter that for surface based measurements is not available before the early 1990's. So the motivation of the paper to use SD mainly and other data in order to retrieve AOD in the past is solid and the results valuable for the atmospheric community. As the authors quote, this is not the first paper approaching this issue (AOD retrieval) using SD data, however a new and other methods are compared here.

My point of view is that there are two major aspects that are not mentioned in this work and without addressing them the work limits its credibility.

The first one is the uncertainty estimation. *This is a difficult task but in order for the AOD data that will be calculated from past SD data to be useful, authors have to provide an uncertainty estimation. There are a lot of aspects linked with this uncertainty that have to do with:*

-The input data.

There are various technical aspects mentioned and referring the Sanchez–Lorenzo work but the effect is not quantified. In addition to them, the use of daily information increases the uncertainty related with the presence of clouds. Datasets that were used for cloud cover assessment include various uncertainties, the simplest one being that their temporal resolution within one day cannot ensure for the 100%-non presence of clouds.

-With the method itself

This is addressed but not quantified through the aeronet based comparison. Temporal issues from the use of daily data too.

The second one already mentioned partly, is the role of clouds.

a. *Cirrus clouds. Cirrus clouds optical thickness (or depth) is in the same magnitude as the one of aerosols. In a number of cases they can be practically invisible and their effect on the SD measurements cannot be distinguished with the one of the aerosols. In the case that instruments like sky cameras or human observations are used, the problem can be partly solved. However, data series going back to the past include mostly or only observations, commonly with a temporal resolution of the order of hours. Moreover, for meteorological station observers it is evident that SD measurements are not affected by partly visible or invisible cirrus clouds.*
b. *Clouds. During the course of the day the presence of clouds will have an effect on the SD duration data. Daily mean cloud cover used here from ECA&D is defined "Daily mean cloud cover: Whenever synoptical cloud cover data is available at 00, 06, 12 and/or 18 UT, mean daily cloud cover is calculated as the average of the available values." So CCT ~0 is only an average of 4 measurements during the day. In addition, WMO has defined the SDF>0.7 for a cloudless condition threshold but of course for aerosols where in most of the cases, the solar attenuation is well within those limits, the threshold serves eliminating only part of the number of cases that can introduce a cloud related uncertainty to the method application.*

Moreover, AERONET related comparisons does not include most of cirrus related and all of thicker clouds cases as these are algorithmically eliminated from aeronet retrievals.

Summarizing, the presence of clouds in some part of the day (with SDF>0.7) can be comparable to the aerosol effect. This causes a small but systematic overestimation of the effect of aerosols in DNI measurements when applying any method based on radiative transfer modeling (where clouds are set to zero) to past series. In addition, periods in the past with changes in cirrus clouds (or increased partly cloudiness) could be wrongfully characterized based in this method, as high aerosol periods.

Thank you for these two main remarks. First about the uncertainty estimation. The key science question, to be answered with our proposed method, is the following: what is the long-term historical evolution of the atmospheric aerosol load? Although the method can provide absolute daily aerosol estimates, the most relevant quantities therefore are the seasonal means and more especially their corresponding anomalies. This latter is widely used because uncertainties and systematic biases can be then minimized, allowing more accurate interpretation of the changes in the past aerosol loads. To make this point clearer also in the manuscript, we have included the following statement in the last paragraph of the "Methodology" section as follows:

"The anomalies represent the quantity of interest in this study, as uncertainties and systematic biases associated with aerosol estimates can be then minimized, allowing more accurate interpretation of the changes in the past aerosol loads."

Nevertheless, in order to quantify uncertainties associated with input data, the method itself and cloud information, diagnostic and prognostic (predictive) methods can be used. Here, we use the diagnostic method for the uncertainty estimates by comparing the sunshine duration retrievals to AERONET data (Sayer et al., 2020). In the revised manuscript, we have carried out the uncertainty estimates, and we have re-written the relevant part of the text after the last paragraph of Section 6.2 (Performance of the three methods) as follows:

"A diagnostic method is used to quantify the uncertainties corresponding to sunshine duration based AOD retrievals (Sayer et al., 2020). AERONET observations are used to derive the diagnostic expected error (EE_{AOD}) envelope for the retrieved AOD. This type of EE envelope uncertainty estimate is similar to the corresponding one of MODIS Dark Target satellite AOD retrievals (Levy et al., 2010). As derived using AERONET AOD as ground-truth, the diagnostic EE envelope includes all possible sources of uncertainties due to for example, changes of the card type, burning threshold, cloud contamination, changes in aerosol properties during the day and uncertainties in sunshine duration measurements. To derive the EE estimates, we use a random subset of about 7000 AOD retrievals from the validation dataset as described previously with all stations. We divide the data into 100 bins, so that each bin contains the same amount of measurements. We compute the standard deviation of the retrieval error and the average retrieved AOD in each bin. We intentionally select the retrieved AOD as the variable to compare the uncertainty against so that it is possible to estimate the retrieval uncertainties also in cases in which accurate or measured AOD is not available. We fit a linear model to the uncertainty data and derive the EE envelope estimate as follows:

$$EE_{AOD} = \pm(0.01 + 0.40 \times AOD_{NHM}) \quad (8)$$

where AOD_{NHM} is the retrieved AOD from the NHM method."

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A. C., Stebel, K., and Witek, M. L.: A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing, Atmos. Meas. Tech., 13, 373–404, https://doi.org/10.5194/amt-13-373-2020, 2020.

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, 10, 10399–10420, doi: 10.5194/acp-10-10399-2010, 2010.

Minor comments

Comment 1. Page2 L15 AERONET has the denser network but looking at the site I can see only four stations with data series more than 15 years. Probably WMO-GAW network or Skynet network can be mentioned too.

Thanks. We have mentioned both WMO-GAW and Skynet networks in the relevant text as follows:

"It has been mainly measured using reference instruments, sun photometers for instance, from various ground-based networks. Among them are the Aerosol Robotic Network (AERONET; Holben et al., 1998), the Global Atmospheric Watch Precision Filter Radiometer network (McArthur et al., 2003) and the SKYradiometer NETwork (Aoki et al., 2006)."

Comment 2. 15 line 15 Ångström

Thanks. Done as requested in the whole manuscript.

Comment 3. Why is the effective wavelength in BAOD close to AOD at 750 nm?

Thanks. Based on our analysis, we acknowledge there is a variability in the effective wavelength of BAOD depending on the atmospheric state. However, we found the assumed 750 nm as effective wavelength in BAOD as a reasonable compromise knowing that the true atmospheric state is not always known, especially during the reconstruction period. In addition, we have already mentioned in the manuscript that other studies (Qiu, 1998; Molineaux et al., 1998) have reported similar assumptions.

Molineaux, B., Ineichen, P., O'Neil, N.: Equivalence of pyrheliometric and monochromatic aerosol optical depths at a single key wavelength, *Appl. Opt.*, 37(30), 7008–7018, doi: 10.1364/AO.37.007008, 1998.

Qiu, J. H.: A method to determine atmospheric aerosol optical depth using total direct solar radiation, *J. Atmos. Sci.*, 55, 744–757, doi:10.1175/1520-0469, 1998.

Comment 4. How can anyone trust Merra data for small to moderate aerosol changes in the past when Pinatubo and El Chichon are not visible in the dataset?

Thanks. We want to stress that in the MERRA data, volcanic sources are included in the aerosol assimilation process. In addition, it is clearly shown in the time series of monthly means over the major aerosol source regions as illustrated in the Figure 13 of Gelaro et al., 2017 or in the winter series of Figure 6 of our manuscript, where the El Chichon and Pinatubo events are visible in the time series. Nevertheless, it is worth mentioning that we have already pointed some limitations of the MERRA AOD dataset in the manuscript in Section 6.3.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M.G., Reichle, R. and Wargan, K.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *J. Clim.*, *30*(14), 5419-5454, doi: 10.1175/JCLI-D-16-0758.1, 2017.