Visibility-derived aerosol optical depth over global land from 1980 to 2021

Response to RC1

We thank the referee for the constructive and helpful comments. We carefully thought about the comments and made corresponding revisions to the manuscript, which have substantially improved the manuscript.

1. Main modifications to the content:

(1) Simplified the introduction.

(2) Modified the introduction of meteorological data (section 2.2).

(3) Modified the temporal matching method with AERONET in section 2.5. Added expected error (Eq. 14).

(4) Added error analysis at global, regional, and site scales (sections 3.3.1-3.3).

(5) Added uncertainty analysis with AERONET AOD (section 3.3.4).

(6) Added trend analysis for India (Section 3.6).

(7) Improved dataset files.

2. Modifications to the chart:

(1) Modified the Figure 5.

(2) Merged the original Figures 6 and 7 into Figure 6.

(3) Added Table 1, evaluation results for each region.

(4) Added Figure 7, evaluation results at site scale.

(5) Added Figure 8, uncertainty analysis.

(6) Added Figure 12, average AOD for different latitude ranges.

(7) Added India in Figures 1 and 13-15.

Comment 1. Is the article itself appropriate to support the publication of a data set?

The data set is not a novel idea, and its premise is rooted with errs in assumptions. The article does not indicate how it improves over data assimilation techniques already implemented by various numerical weather prediction and reanalysis systems.

Vertical resolution of the aerosol is needed to adequately link the surface visibility to the total column. The visibility from AOD technique suffers similarly from the same fundamental problem investigated for more than two decades to accurately determine PM2.5 from space in the presence of aerosols above the surface boundary layer. This data set continues to have the same drawbacks as techniques used in the 2000s and the results show the Vis AOD data set does not address this key aspect with the stated machine learning methodology. When providing this data set, this major drawback must be clearly stated to the user when considering using these data. The use of vertical aerosol measurements may improve this data set and make it more useable.

Response 1.

- We agree with the referee on the difference between surface visibility and aerosol

optical depth, which is discussed in the 9th paragraph of the introduction. To address this issue, we included the boundary layer height into our method, as the most aerosols locate in the boundary layer. The evaluations in this study demonstrate that this method can provides reliable dataset in depicting annual and long-term trend. As we know, this study provides the first global AOD over land from 1980 to 2021 based on visibility.

- The primary purpose of the study is to provide the global AOD over land for the time period before the EOS satellite era (i.e., before 2000) and analyze trends in global land and regions, which provide a unique data for broad user. We did not intend to improve data assimilation techniques implemented by various numerical weather prediction and reanalysis systems.

- We agree with that the profile of aerosols is the best way to estimate AOD based on surface visibility, however, the reliable dataset on aerosol profile is not globally available. We have discussed the possible errors in section 3.3.4.

Comment 2. Is the data set significant – unique, useful, and complete?

The data is available via HTTP download with a short delay but data is not bundled optimally. The data set provides a zip file. Inside the zip file, it provides "Gridded_Monthly_AOD_1980_2021" and another zip file "Station_Daily_AOD_1980_2021.zip."

Comment 2.1

Inside the "Gridded_Monthly_AOD_1980_2021" directory, a file is provided in netCDF4 format with compatibility for HDF5. The netCDF file has four variables, time "T", Latitude "Lat", Longitude "Lon", and visibility to AOD "VIS_AOD". The netCDF file lacks various metadata such as standard_name, _FillValue, valid_range, long_name, and coordinates. The shape "VIS_AOD" is 504 (Time), 721 (Longitude), and 291 (Latitude). The latitude dimension has an unusual shape since the manuscript defines the latitude range from -60N (or 60S) to 85N. This spatial extent is not provided in the metadata or the data landing page for TPDC.

Response 2.1

- Thank you for your suggestion. We follow the referee suggestions and re-uploaded the dataset. (1) The variables in the netCDF file have been modified to "time", "longitude", "latitude", and "VIS-AOD". (2) We have added the missing metadata data in the netCDF file. (3) The shape (-60°N to 85°N) of latitude dimension is based on the distribution of meteorological stations. (4) We have added the spatial extent in the metadata and the data landing page in TPDC, which is waiting for the data center manager's approval.

Comment 2.2

Inside Station_Daily_AOD_1980_2021.zip, a directory called "Station_Daily_AOD_1980_2021" has ASCII text files that are named for the ASOS station with Vis AOD data in which the "VIS_AOD" column does not indicate the wavelength (i.e., 550nm). A separate "0A0A-Station_Information.txt" file provides the ASOS station names and associated longitude, latitude, and elevation (note that data units are not provided) in this file.

Response 2.2

- Thank you for your suggestion. We have added data units in the file of station information and uploaded to the data center.

Comment 2.3

The manuscript states at line 462, "The more sample data input, the better the model performs." Typically, 80% should be used for training and 20% for testing. Using all of the input data leads to weaker determination of the performance of the model.

Response 2.3

- Thank you for your comment. We agree with referee and have attempted to use 80% for training and 20% for testing. The reasons for not using this method are:
  (1) Figure 3 shows that models perform better using more sample data.
  (2) More importantly, we have two independent datasets for evaluation (Terra and AERONET AOD).

Comment 2.4

Figure 6 shows VIS_AOD has wide variability from zero to near ~0.7 when compared to all three input sources. These differences could lead to very large variations in daily climatology analysis at specific locations. In Figure 7, the correlation improves by increasing the temporal average window which acts to smooth errors in the model. However, in Figure 7, the bottom right panel showing Annual AERONET (labelled incorrectly as TERRA), still indicates a much weaker correlation as errors in modelling and systematic biases and uncertainties with MODIS AOD input products are still evident.

Response 2.4

Thanks for your comments. The type error has been corrected. According the comments from referee #1 and #2, we made the following revision:

- We have made modifications in original Figure 6 and Figure 7 (Figure 6 in the revised version). 15-minute AERONET AOD (550nm) is used to validate and the expected error ($\pm$ (0.05+0.15 * $AOD_{AERONET}$)) is used to evaluate VIS_AOD, which show higher correlation coefficient in daily scale.
- We added more discussion. In Section 3.3, the error analysis results show that VIS_AOD is underestimated in heavy pollution. It has led to significant changes in daily scale climate analysis at specific locations in section 3.3.3 Validation at a site scale. The discussion on averaging over time scales (Schutgens, et al., 2017) was also added. The error analysis and limitations for VIS_AOD are discussed in section 3.3.4 (Eck et al., 2023; Levy et al., 2013; Levy et al., 2018; Li et al., 2020; Wei, et al., 2019a; Wei, et al, 2020; Zhang et al., 2020).

Comment 3. Is the data set itself of high quality?

Comment 3.1

Checking Vis AOD with AERONET AOD daily averages shows an example of deviations that often occur when elevated aerosol layers impact both AERONET measurements and MODIS retrievals compared to the surface visibility.

For more background on this case, see the following article (Eck et al., 2023): https://doi.org/10.1016/j.atmosenv.2023.119798

Aerosol transport cases from biomass burning smoke and dust affect many sites around the globe. These episodic smoke transport events are increasing as drought severity due to global warming continues to promote drier conditions in vegetated and desert regions. The world's deserts such as Saharan, Gobi, and Thar, desert dust continue to be key sources of aerosols lofted above the boundary layer and transported 100s of kms. Other elevated aerosols include volcanic ash and gas-to-aerosol conversion of sulfur dioxide to sulfates in the upper and lower stratosphere such as the recent Hunga Tonga volcanic eruption.

Response 3.1

- Thank you for using our data and providing an example of the bias between daily VIS_AOD and AERONET AOD during aerosol transport events. We agree with your point and have carefully read the background information of the case. The difference between surface visibility and column aerosols is a key factor causing this deviation. We discussed it in section 3.3.4 and cited the reference (Eck et al., 2023). Figure 5 show in the revised version show that VIS_AOD has a comparable accuracy of TERRA AOD and AQUA AOD at daily, monthly and yearly time scales.

Comment 3.2

The manuscript indicates ASOS data (https://www.weather.gov/asos/faq.html#12) are retrieved from the Iowa State University (https://mesonet.agron.iastate.edu/ASOS/). This Iowa State University web page indicates the following information on the U.S. ASOS network which does not refer to the international weather observation stations that are not maintained by the U.S. NWS, FAA, or DOD. While the Iowa State University search tool indicates "ASOS" for the international locations, it is incorrect to assume that they are managed by the U.S. Each international network has their own method for monitoring surface weather using different instrumentation and methods; however, some stations such as international airports need to meet internationally mandated regulations (e.g., WMO: https://community.wmo.int/en/implementation-areas-aeronautical-meteorology-programme), but these may not apply to other stations within the country. Therefore, the comprehensiveness of surface weather data at the international locations may not be as robust quality or measurement accuracy as those collected at international locations outside international airports. Assertion that the global network of surface meteorological measurements is managed by the U.S. ASOS network is misleading.

Response 3.2

- Thank you for your correction. We agree with your opinion. The introduction of meteorological data has been modified in section 2.2.

Comment 3.3

AERONET sites often are not collocated with ASOS measurements. The representativeness of the AERONET site compared to the ASOS site needs to be considered. Did the Authors consider the elevation of the AERONET site with respect to the ASOS location? Oftentimes, AERONET sites can be elevated and further above the ASOS location in urban, forested, mountainous or marine locations in which they

are placed on towers or buildings 10s of meters high above ground. The difference in elevation can have a significant effect on the visibility relationship to the AOD measurement. Therefore, the measurement of the aerosols may be less at the measurement altitude of AERONET site compared to the elevation of the ASOS location.

Response 3.3

- Thank you for your suggestion. We have added error analysis and uncertainty analysis for VIS_AOD in section 3.3.4.2-3.3.4.4 and a new figure was added (Figure 8 in the revised version) to show: (1) Elevation of AERONET site and bias (Figure 8 (b)), (2) Elevation difference between AERONET site and meteorological station and bias (Figure 8 (c)), (3) Distance between AERONET and meteorological station and bias (Figure 8 (d)).

Comment 4. Is the data set publication as submitted of high quality?

Comment 4.1

At lines, 511-513, the statement: "However, the AERONET AOD results are slightly inferior to those of Aqua and Terra AOD, which could be caused by the representativeness of the AERONET station spatial coverage and measurement error (Holben et al., 1998)" is interpreted out of context. AERONET AOD are superior in the determination of AOD. The determination of the spatiotemporal representativeness is within the purview of the methodology utilized to perform the matchup and the justification for such a methodology to improve the spatiotemporal representativeness. Therefore, the fact that MODIS AOD follows closely to the Vis AOD suggests that this input is significantly weighted to it. This is obvious from the variation in the stated correlation coefficients between the model and the input data sets where AERONET is much weaker, and MODIS is much stronger.

Response 4.1

- Thanks for your comments. We follow the referee's suggestion to revise the matching method between AERONET and visibility derived AOD for comparison. We used a 15-minute AERONT AOD for spatiotemporal matching and validation. Figure 5 show in the revised version show that VIS_AOD has a comparable accuracy of TERRA AOD and AQUA AOD at daily, monthly and yearly time scales. We added new discussion on the differences between VIS_AOD and MODIS AOD and AERONET AOD in section 3.3.4.

Comment 4.2

Lines 513 – 515 states "Nevertheless, the results indicate the high reliability and strong predicted capability of the model, and the visibility-derived AOD can be used for aerosol climatology." A major issue with this statement is that areas of the world are affected by transported aerosol above the boundary.

Response 4.2

- Thank you for your correction. We agree with your opinion. We have made the modifications and discussed in section 3.3.4.

Comment 4.3

What is the uncertainty of the boundary layer height for the ERA5 reanalysis? Please state.

Response 4.3

- Thank you for the suggestion. We have added the uncertainty of the boundary layer height for the ERA5 reanalysis in section 2.3.

Comment 4.4

The netCDF file lacks various metadata such as standard_name, _FillValue, valid_range, long_name, and coordinates.

Response 4.4

- Thank you for the suggestion. We have added the lack metadata in the netCDF file and reuploaded the dataset to the data center.

Comment 4.5

Figure 2 lists Aqua MODIS twice and does not indicate Terra MODIS.

Response 4.5

- Thank you for the correction. We have made an adjustment in Figure 2.

Comment 4.6

Figure 7 shows the lower right panel the same x-axis title as the middle right panel.

Response 4.6

- Thank you for the correction. We have removed Figure 7 and replaced it with Figure 6.