Respond to Reviewer #2

Dear reviewer, thank you for your useful comments. We have carefully studied your comments, and replied to your comments point by point and included corresponding modifications. In the following text, your comments are marked in bold italics, our responses are in black, and the modifications in manuscript are shown in blue.

General comments

The manuscript by Jin et al. evaluated the distributions, trends, and variabilities of aerosol parameters (concentration, sizes, and types) using multi-source, long-term aerosol records, including ground-based observations, satellite products, and atmospheric reanalysis, which ensure the quality of the study and the conclusions. More importantly, geometric mean is used to better describe the lognormal distribution of aerosol, and a TFPW-MK method was introduced to avoid error in trend analysis of time-series data due to probable auto-correlation. Generally, the manuscript is well organized and written. However, before it could be considered to publish, there are still some problems that need clarification.

Thanks for your comments and recognition. We have carefully revised our manuscript following your comments and we believe that our manuscript has been improved greatly.

Specific comments

- Line 59: The statement is unclear for me. How could the short-term variances be beneficial to reveal the natural aerosol emissions? Please give a brief explanation.

Thanks for your comment. We'd like to clarify that the short-term variability in aerosols we are discussing here are intended to reveal some events related to aerosol generation or transfer, and not necessarily actual study on aerosol emissions. Detecting aerosol emissions from satellite data is a complex and challenging issue that may require precise detection techniques and the use of model simulations [1]. Our motivation is to obtain pixel-level variability in the spatiotemporal distribution of aerosols through some simple calculations, which is probably meaningful about future studies on aerosol or particulate matter concentration estimation. For example, the GRASP algorithm we are studying allows for different spatiotemporal variabilities to be defined for each physical model parameter. From the results (Figures 12-13 in the manuscript), it appears that the short-term variability of aerosols we have studied is large in some known regions with significant aerosol emissions, such as the summer in northern Asia, dust source regions, and areas with severe urban pollution in China and India. More detailed study [2] has also given similar results, such as in urban polluted aerosol environments, where aerosol temporal variability is higher when wind speeds are lower due to reduced aerosol particle clearance by winds. So, our calculation of small-scale spatiotemporal variability of aerosols is primarily aimed at displaying a result which may be useful for estimating aerosol particle concentrations, rather than studying aerosol emission sources.

[1] Dubovik, O., Lapyonok, T., Kaufman, Y., Chin, M., Ginoux, P., Kahn, R., & Sinyuk, A.: Retrieving global aerosol sources from satellites using inverse modeling. Atmospheric Chemistry and Physics, 8, 209-250, doi:10.5194/acp-8-209-2008, 2008

[2] Chen, C., Dubovik, O., Schuster, G.L., Fuertes, D., Meijer, Y., Landgraf, J., Karol, Y., & Li, Z.: Characterization of temporal and spatial variability of aerosols from ground-based climatology: towards evaluation of satellite mission requirements. Journal of Quantitative Spectroscopy and Radiative Transfer, 268, 107627, doi:https://doi.org/10.1016/j.jqsrt.2021.107627, 2021

- Line 224: What is the type I error, and please give a brief explanation.

Thanks for your comment. The Type I error means that the null hypothesis is rejected when it is actually true, which results in a false positive result. In this study, Type I errors are mainly caused by the spatiotemporal autocorrelation of the data. The increase in Type I errors means that we obtain false significant trends of aerosol, which may lead to incorrect conclusions. Therefore, we used various methods (TFPW-MK and FDR) to avoid increasing Type I errors and obtain reliable conclusions.

To facilitate readers' understanding, we made the following modifications:

(Lines 225-228) However, the spatiotemporal auto-correlation is usually widespread in observed AOD series (namely, the data are not strictly independent and random). This makes the original MK test usually results the occurrences of type I error (rejecting true null hypothesis) in statistics and a large number of fake significances (Von Storch 1999).

- Line 296: This study used the AOD geometric to investigate the temporal trends, which is an essential point for the work. Meanwhile, it appears they use normal mean for standard deviation calculations. Does this lead to any biases?

Thanks for your comment. Yes, using geometric mean to process data is a feature of our study, but it will not be contradictory to using standard deviation to evaluate small-scale variability of aerosols. Because the data organization method of geometric mean is mainly for studying the spatial distribution and trend of aerosols, which is not a same topic comparing to the small-scale variability. More importantly, the standard deviation is a simple statistical measure that is not related to the data distribution. This means that when we only want to calculate the standard deviation without any additional hypothesis, the distribution of the data does not need to be considered.

- Section 4.1 is helpful to understand the quality of different products before looking at the long-term trends. However, some further analysis/discussion would be beneficial for the readers. For example, how much data is removed due to cloud cover and/or unfavorable land cover such as snow? How does this impact the analysis of the long-term trends? Does this lead to more or less data in some areas during the same season?

Thanks for your comments. Yes, cloud and snow cover or missing data can have an impact on statistical results. For MERRA-2, all data are continuous and complete; for satellite data, the cloud coverage rate is around 60%, and snow usually impacts aerosol retrieval in high latitude area during winter. But these issues do not affect the main results because we have taken into account the seasonal variations in both spatial distribution and trend analysis (using seasons or months is also recommended compared to analyzing on an annual basis [1]). In addition, in the long-term trend analysis based on rank, data missingness is also acceptable due to the advantages of the TFPW-MK method [2].

To avoid unnecessary misunderstandings, we have added some explanations at the relevant locations. The specific modifications are as follows:

(Lines 395-399) Here, it is noted that due to snow cover, there was a large number of missing retrievals in winter of NA, which would lead to an overestimated mean AOD for the whole Asia. Therefore, the specific details of seasonal variations in aerosol should be obtained from the distribution maps of AOD. Moreover, since the rank-based trend analysis method allows for missing data, the missing values caused by cloud cover and snow cover do not have a significant impact on the calculations of trends.

[1] Collaud Coen, M., Andrews, E., Bigi, A., Martucci, G., Romanens, G., Vogt, F.P.A., & Vuilleumier, L.: Effects of the prewhitening method, the time granularity, and the time segmentation on the Mann–Kendall trend detection and the associated Sen's slope. Atmospheric Measurement Techniques, 13, 6945-6964, doi:10.5194/amt-13-6945-2020, 2020

[2] Yue, S., Pilon, P., Phinney, B., & Cavadias, G.: The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16, 1807-1829, doi:10.1002/hyp.1095, 2002

- The AOD was higher in spring than winter due to the emissions associated to heating and industry and lower boundary layer? How much of this is due to cloud cover vs. dust from deserts?

Thanks for your comments. There is a misunderstanding here. It has been shown that industrial emissions and lower boundary layers can exacerbate regional air pollution events, but these issues do not apply to studies of total AOD (total atmospheric column) in Asia. We just want to briefly remind the readers that, due to the missing satellite data, the average total AOD in winter may be biased towards higher values because of the missing data from North Asia (where anthropogenic emissions are relatively low). However, it is actually very difficult to assess the impact of data gaps or dust on total AOD at such a large scale. If we look at it from the perspective of MERRA-2 (where the data is continuous and complete): the average AOD in Asia is highest in spring (0.215) and follow by summer (0.200), autumn (0.140), and winter (0.133); the average dust AOD in Asia is highest in spring (0.059), followed by summer (0.036), autumn (0.023), and winter (0.022).

- As another note on the aerosol type, the authors report decreasing or increasing dust and sea salt. But why these are decreasing or increasing as the emissions of these two aerosol types are natural and not anthropogenic.

Thanks for your comments. Yes, in general, natural sources of aerosols are rarely influenced by human activities. However, some human factors may alter the distribution of aerosols, with the most typical example being Beijing's protective policies against the northwestern transport of dust in the spring. As shown in Figure 10 in manuscript, from a perspective of long-term trend analysis, this has led to a decrease in dust aerosols in the Beijing region.

Additionally, in the previous version, we overlooked the spatial autocorrelation of the data, which could lead to false significant trends. After applying the FDR method, many false trends related to dust and sea salt have been eliminated.

- Line 105-109: Except for the consideration of adjacent pixels, a globally consistent assumption is also very important for continuous distribution in aerosol retrieval.

Thanks for your comment. Yes, in retrieval, the assumption of global consistency is important for the continuous distribution of aerosols, but it is challenging to achieve.

- Section 2.2: Please clarify the estimated uncertainties of different parameters from observation of CE-318.

Thanks for your comment. we have revised the manuscript and added a description of the uncertainties:

(Lines 133-143) The version 3 AERONET aerosol products did a lot of improvements based on the version 2, such as applying a new polarized radiative transfer code, updating datasets of spectral solar flux and surface albedo parameters, modifying temperature and gas absorption corrections, and using stricter quality control rules (Giles et al. 2019; Sinyuk et al. 2020); The bias of AOD estimation is +0.02 and the uncertainty with one standard deviation is ~0.02; The uncertainty of SSA is estimated ~0.03 when the AOD at 440 nm is larger than 0.4. In addition, an extra site established at Wuhan University was also applied in this study (Jin et al. 2021) to supplement the aerosol optical properties in CC area. This site is equipped with the same sun sky photometer as AERONET and calibrated annually using the China Meteorological Administration Aerosol Remote Sensing Network (CARSNET) (Che et al. 2009). The AOD was calculated from the direct sun measurement and other complex properties were retrieved from the sky irradiance under cloudless conditions (Smirnov et al. 2000), with the method of Dubovik and King (2000). The criterion of data is similar to the Level 1.5 Version 2.0 product released by the AERONET.

- How different it is if using a normal distribution in the long-term aerosol studies compared with using a lognormal distribution.

Thanks for your comment. It is difficult to use a unified standard to evaluate the differences between geometric and arithmetic means, and these differences are mainly concentrated in areas with high aerosol loadings. Following report by Sayer [1], the differences between geometric and arithmetic means of aerosol (**Fig. R1**) can change a lot depending on the time scale of data aggregation and the magnitude of AOD. For daily data, the difference is from 0 to ~0.02; for monthly data, the difference can be magnified by orders of magnitude.

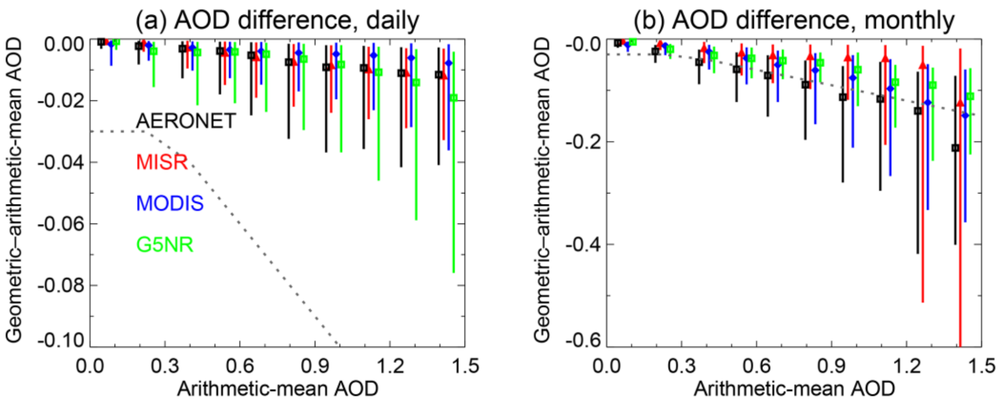

Fig. R1. Difference between geometric and arithmetic means of AOD (Figure 9 in Sayer et al. (2019) [1])

[1] Sayer, A.M., & Knobelspiesse, K.D.: How should we aggregate data? Methods accounting for the numerical distributions, with an assessment of aerosol optical depth. Atmospheric Chemistry and Physics, 19, 15023-15048, doi:10.5194/acp-19-15023-2019, 2019

- Line 280: What do the 'pixel level' and 'region level' refer to? Please clarify.

Thanks for your comments. The 'pixel level' means that trends are estimated based on the grids (such as Figure 9-10 in the manuscript), and 'region level' means that trends are estimated based on sub-regions in Asia, such as EA, SEA, and SA et al.

- Figure 10: Is the aggregation method of aerosol percent also the geometric?

Thanks for your comments. Yes, as shown in Eq. 16 in the manuscript, the mean values and trends are all calculated based on the geometric mean method.

- Table 2: The SS aerosol in SEA showed a slight decrease but this was not revealed in size-segregated AOD.

Thanks for your comments. We want to clarify that in the previous version of our study, the lack of consideration of spatial-correlation in the data may have led to some false trends. With the addition of the FDR method, the trends obtained are more accurate in this version of the manuscript. Therefore, these false trends shown by the SS aerosol have been largely eliminated in this version. Additionally, the MERRA-2 data is more continuous in both time and space, resulting in more detected trends. However, the SS aerosol from MERRA-2 and size-segregated aerosols from MISR are not one-to-one correspondence, so the analysis shows as different a result. Therefore, we used both satellite and reanalysis data to obtain a comprehensive estimation of aerosols.