<u>Responds to anonymous referee 1</u> :

The authors thank the referee for their detailed review of the manuscript and for all their comments and suggestions about comparison with satellite measurements. The inclusion of these airborn REM trend analyses allows a clear improvement of the paper with a more global view of aerosol modifications around the world.

During the review process, the routines for MK trend analysis were translated into R and an error was found in the selection of data for north hemispheric winter season. This error was corrected in the original matlab routines leading to minor changes in slope absolute values for most of the stations, but also sometimes to modification of the statistical significance. The more important changes are:

- ALT was the only station with ss trend in absorption coefficient and this was the only case where there is a strong discrepancy among the analysing methods, MK being ss positive, LMS/log not ss and GLS/day ss negative. The correction leads to MK not ss trend in absorption coefficient at ALT and remove therefore the solely strong discrepancy between the methods.
- MLO has a ss negative trend in scattering coefficient for the last 10 y, leading to a better agreement between scattering and absorption trends. The evolution from positive to negative ss trends is now well established.
- Some other not ss present-day trends are now ss negative (RMN scattering coefficient, CPR absorption coefficient, THD single scattering albedo) or ss positive (PUY single scattering albedo, MSY scattering Ångström exponent, LLN absorption Ångström exponent).
- Some ss trends are now not ss: IZO absorption coefficient,
- One trend (JFJ scattering Ångström exponent for the 20y period) change from ss negative to ss positive trend.
- The statistical significance of some of the 10 y trends of the time evolution analysis (Sect. 3.2) is also modified, but these changes do not impact the results.

The revised manuscript and all tables and figures were corrected in order to take into account the new results.

Answers to specific comments:

I have two major concerns with the manuscript and some other general concerns. The first major concern is overstatement about the global scale of these data. The second is a lack of the context provided by satellite data.

1. First, the manuscript claims in the title to be at a global scale and makes statements about global trends (for example page 25 line 5 "leading to global positive median trend of 0.02%/y"). There is just no way that the stations in the manuscript represent the global scale. In figure 12, there are 46 stations, 32 of which are in North America or Europe. Other figures are very similar. That means about 2/3 of the data are from less than 7% of the area of the Earth. In Figures 4 through 11 there are no stations in South America, none in the vast majority of Africa, and none on the main continent of Australia. You can't claim a global scale when entire continents are missing. Furthermore, the station locations are probably biased to regions with

decreasing trends. Regions with recent decreasing trends in aerosols, such as North America and Europe, are heavily represented. On the other hand, regions with increasing trends in aerosol in the last decade or two, such as India and the Mideast, are not represented in the figures.

The referee is completely right. The word "global" was used for two purposes: 1) this study represents the best "globality" that can be reach with in-situ aerosol measurement and 2) the word "global" is used instead of the word "annual" in the result section. These two inadequate usages of the word "global" were modified in the manuscript. The title is now "Multidecadal trend analysis of in-situ aerosol radiative properties around the world".

2. One specific statement is the conclusion (page 25 line 40) "Results from this study provide evidence that the aerosol load has significantly decreased over the last two decades in the regions represented by the 52 stations" is very misleading and should be changed. It should read something more like "The stations considered confirm decreasing trends in North America and Europe. Trends elsewhere are scattered, with too few stations to understand global trends." The rest of the manuscript should be similarly less definite about a global scale.

The authors agree that the results of this study cannot be considered as global due to the low representativity of stations in Asia, South America, Africa and Australia. The manuscript was consequently modified:
"Results from this study provide evidence that the aerosol load has significantly decreased over the last two decades in North America and Europe. The low number of stations in the other continents means global tendencies cannot be assessed and the results are more variable."

3. Second, the manuscript perpetuates an unfortunate situation in the literature that the in-situ and satellite researchers rarely make use of the other. I often tell satellite researchers they need to consider the in-situ data. Here the in-situ researchers need to consider the satellite data. Why should this manuscript be the one to do that? It claims a "global scale", and that definitely means including some satellite data.

In the present manuscript, satellite data are dismissed saying the ground stations have longer records (line 16, page 3). This is mostly untrue. MISR and MODIS both have 20 years of data, making their record longer than all but a handful of ground stations (and almost all of those handful are in the United States). SeaWifs has an even longer data record (Hsu et al., 2012). For the entirety of satellite data, the manuscript has only one oddly chosen reference about measurements in South Korea.
MISR and the newer MODIS retrievals provide aerosol optical depth over land as well as ocean. They measure more than optical depth. MODIS measures the Angstrom coefficient for scattering. MISR has a measure of the single scattering albedo, with some difficulties in the measurement but good enough for trends in some locations.
I am not asking for a major review of how satellite data relates to long-term, ground-based measurements. I do think it is reasonable to ask you to show a figure with a map of satellite derived trends in optical depth, Angstrom coefficient, and possibly aerosol absorption (optional, since the satellite absorption data are a bit trickier). The period could be something like 2009-2018 or 2004-2018 to match most of the ground sites. Then use that figure to put your ground stations in context. It isn't that hard to produce such figure. Out of 40+ authors there should be somebody who has experience using satellite data. If there isn't, it says something about our field. I'm copying one of your figures next to some satellite context below.
For discussing the context from satellite data, one important reference is Zhao et al. (Environ.Res. Lett, 2017) because it shows trends in not only optical depth but also detailed optical properties

such as single scattering albedo for the Eastern US, Europe, and China. At a quick glance, those trends seem consistent with the ground stations; you can do a better analysis.

Environ. Res. Lett. **12** (2017) 054021

Letters

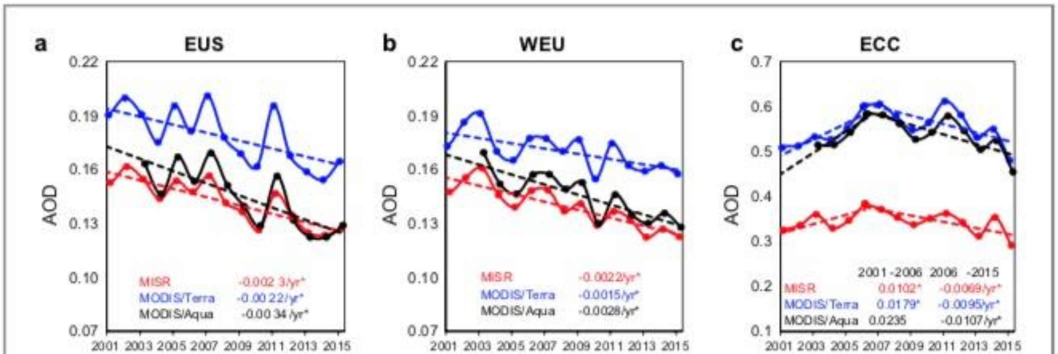


Figure 2. Annual mean AOD during 2001–2015 observed by MISR, MODIS/Terra, and MODIS/Aqua in three target regions. The dashed lines and the numbers in the legends represent trendlines and slopes (yr^{-1}) determined by linear regression, respectively. The slopes are marked with a star (*) if they are statistically significant at the 0.05 level based on Student's *t*-test. The trends and their statistical significance determined by the Mann–Kendall test associated with Sen's slope are shown in table S1.

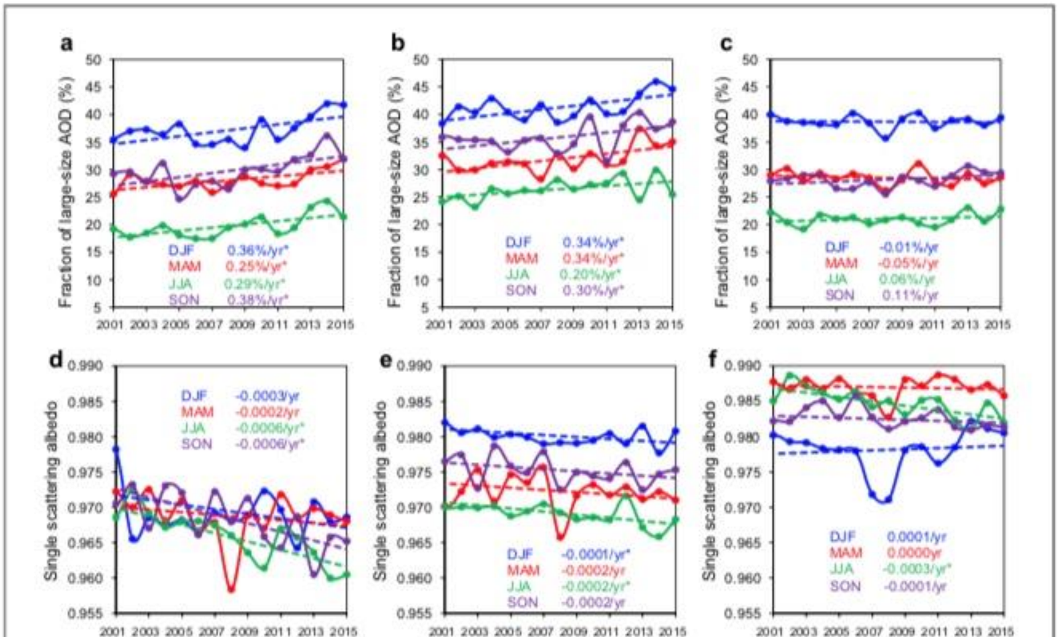


Figure 5. Seasonal mean fraction of the large-size AOD and single-scattering albedo during 2001–2015 observed by MISR in three target regions. The fraction of large-size AOD: (*a*) EUS, (*b*) WEU, and (*c*) ECC. Single-scattering albedo: (*d*) EUS, (*e*) WEU, and (*f*) ECC. The dashed lines and the numbers in the legends represent trendlines and slopes (yr^{-1}) determined by linear regression, respectively. The slopes are marked with a star (*) if they are statistically significant at the 0.05 level based on Student's *t*-test. The trends and their statistical significance determined by the Mann–Kendall test associated with Sen's slope are shown in table S3.

Alfaro-Contreres et al. (2017), Wei et al. (2019), and Murphy (2013) show the context that the region from the Mideast to India (with no ground stations) has had increasing trends in aerosol.

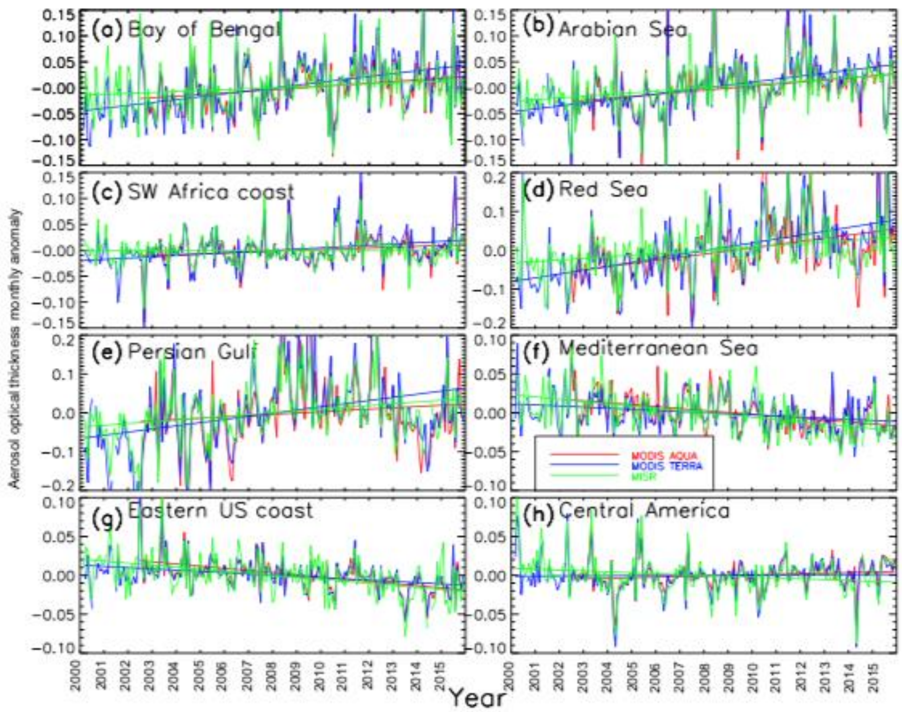
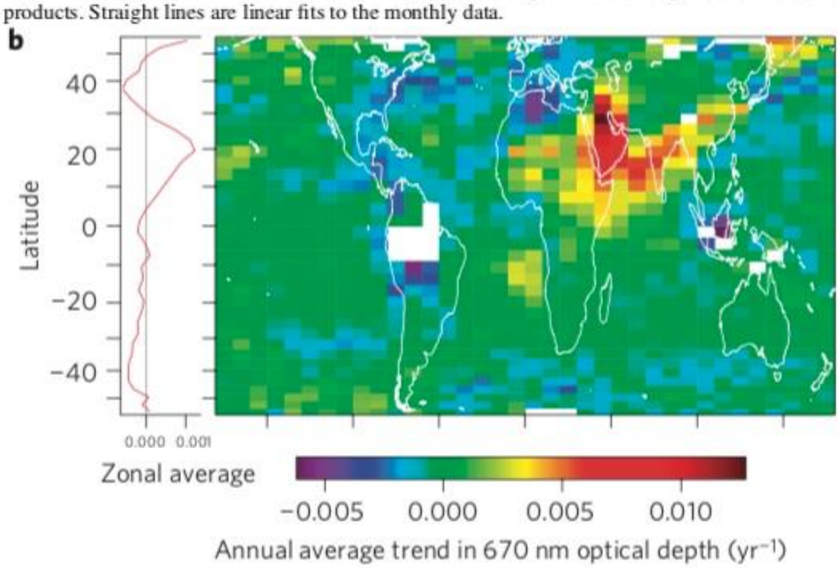


Figure 4. The deseasonalized, monthly and regionally averaged AOTs for eight selected regions utilizing MODIS C6 DT and MISR aerosol products. Straight lines are linear fits to the monthly data.

b



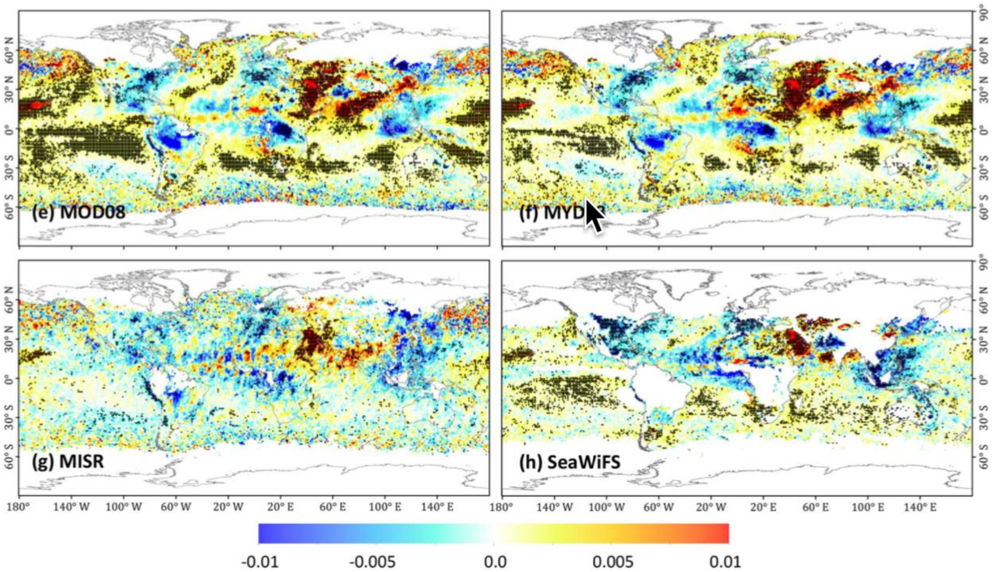
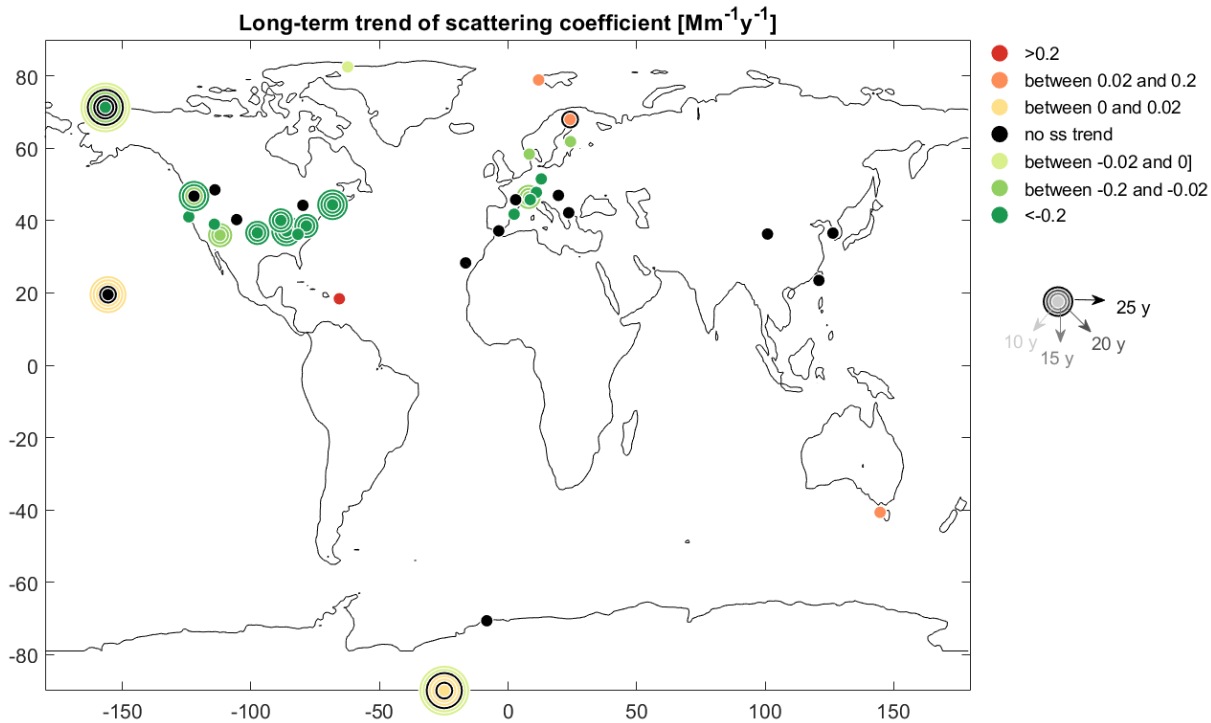


Figure 14. Linear trend based on deseasonalized monthly AOD_S anomalies from 2003 to 2010. Units are AOD yr^{-1}. Black dots indicate a significant trend at the 95 % confidence level ($p < 0.05$).



You can also look at Mehta et al., Remote Sensing of the Environment, 2016 Kahn and Gaitley JGR 2015, Hsu et al., ACP, 2020, “Global and regional trends…”, Wei et al., ACP, 2019, “Intercomparison in spatial distribution and temporal trends”. This is not a comprehensive list.

The authors thank the referee for pointing this lack of results from satellite measurements, for the proposed references and do agree about the unfortunate statement of the relative confinement of both research domains. The authors do further agree with the referee that the in-situ measurements of aerosol properties cannot be called “global” since large

domains of world are under-represented (e.g. Asia) or even not represented at all (e.g. Africa, South America, middle East). The world "global" was then removed from the manuscript.
After a literature review, guided in part by the reviewer's suggestions, we did not find an already published global long-term trend analysis from satellite measurements with the same (or similar) aerosol radiative properties used in our study. Despite the number of co-authors, none has a sufficiently experience to extract aerosol radiative properties from satellite measurements and compute global trends. Moreover, the authors consider that a necessary condition of good comparison between in-situ and satellite aerosol trends is the used of the same trend analysis methodology. If such an analysis would be very relevant, the amount of necessary work to achieve it is too large to be done in the restricted lapse of time imposed by the next IPCC report deadline. The authors are however open for a further collaboration in this domain.

To respond to the referee's requirement, sect. 4.4 was completely re-written in order to include some results from already published long-term trend analysis from satellite measurements:

"4.4 Comparison with other trends and causality

The current study has focused on surface in situ aerosol optical properties at point locations, primarily in North America and Europe, but also in Asia and Polar Regions. Comparison with reported trends from other long-term measurements of aerosol properties (e.g., surface aerosol mass concentrations, surface chemical mass concentrations, ground-based and satellite column optical properties, etc), can provide a more holistic and global view of changes in the atmospheric aerosol. Model simulations of aerosol trends can also supply insight into global impacts of emission changes. We, thus, present a (non-exhaustive) comparison of the trend results from this study with some other relevant aerosol trend studies in the literature. The supplemental materials of Li et al. (2017) include a summary of trends reported in the literature for AOD, PM2.5 and several aerosol constituents (e.g., sulphate, BC, etc.).

There are some important caveats to keep in mind when comparing aerosol trends across platforms and instruments. First, they represent different aspects of the aerosol (chemical, physical, or optical), at different conditions (dry or ambient), different wavelengths (300-1100 nm), different techniques (in-situ, REM) and different locations (ground-based, airborne or satellite). Second, there are differences in the statistical methodologies, both in terms of methods used and data treatment. Third, the periods covered often overlap, but are not the same. Further, some REM measurements can only be made under certain conditions (e.g., daylight and cloud-free conditions versus continuous sampling, over land versus over ocean, etc.), meaning temporal coverage may be quite different. Because of all these differences, we only discuss general tendencies rather than absolute values when comparing trends from different studies. Below we first compare our results with trends from other surface in-situ measurements and REM observations. Finally, we discuss causes of the observed trends and speculate specifically on some of the trends in intensive aerosol properties, which have received less attention in the literature than properties related to aerosol loading.

4.4.1 Comparison with other surface, in-situ aerosol trends

A comparison of the present day trends derived here to our previous trend ending in 2010 (CC2013) demonstrates that the larger number of stations, particularly in Europe, permits a more detailed view of regional trends. The current wide coverage across continental Europe shows decreasing present-day trends. Decreasing σ_{sp}, σ_{bsp} and σ_{ap} trends were confirmed for individual stations (e.g., SMR (Luoma et al., 2019), PAL (Lihavainen et al., 2015b), ARN (Sorribas et al., 2019)), as well as at ACTRIS sites including JFJ, HPB, IPR, IZO, PAL, PUY, SMR and UGR (Pandolfi et al. (2018)).There are some discrepancies in the trends between our current study and Pandolfi et al. (2018) that seem to be principally due to differences in the analyzed periods. Three additional years of data were included in this study and some older periods included in Pandolfi et al. (2018) were invalidated following the evaluations described in Sect. 2.4. The European b and $å_{sp}$ trends computed by Pandolfi et al. (2018) are similar to the results of this study for most of the stations, in that they also found a general ss increase of b and variable $å_{sp}$ trends. In North America the ss decreasing trends in aerosol extensive properties observed in CC2013 are found to have continued in this work with the extended data sets. These results are confirmed by the two other trend studies for in-situ aerosol optical properties in North America. While the methodology and time period of Sherman et al. (2015) were different, the sign and ss of their σ_{sp}, b,and $å_{sp}$ trends for BND and SGP were the same as reported here. White et al. (2016) found a decreasing trend in absorption coefficient (estimated from light transmittance measurements on 24 h filter samples) at 110 IMPROVE stations for the 2003-2014 period. SPO σ_{sp}, b and $å_{sp}$ trends for the 1979-2014 period (Sheridan et al., 2016) do agree with CC2013 results, whereas the 1979-2018 trends reported in this study suggest an evolution towards more ss positive trends. The very low aerosol concentrations in Antarctica and the difference in the MK algorithm could however also explain the differences amongst these three analyses.

There have been multiple trends studies on carbon species (also referred to as black carbon (BC), elemental carbon, equivalent black carbon, brown carbon or other terms) which is closely related to aerosol absorption. A decreasing trend in BC concentration is found in Europe (Singh et al., 2018, Kutzner et al., 2018, Grange et al., 2019) related primarily to traffic emission decreases rather than changes in wood burning and/or industrial emissions. Similarly, Lyamani et al. (2011) noted a decrease in BC in southern Spain due to the 2008 economic crisis. In contrast, Davuliene et al. (2019) reported an increasing trend in equivalent black carbon (eBC) for the Arctic site of TIK. In North America, White et al. (2016) found that the decreasing elemental carbon trend at IMPROVE sites was larger than the aerosol absorption trend at the same sites due to the impact of Fe content in mineral dust. BC trends in the Arctic have been extensively studied (e.g., AMAP, 2015; Sharma et al., 2019; and references therein) and suggest a decreasing trend. This is consistent with our general trend in absorption for polar regions (Table 4), although for individual stations most trends were statistically insignificant.

Particulate mass (PM) and visibility are other metrics for atmospheric aerosol loading that can be most readily compared with our trends in aerosol scattering. Tørseth et al. (2012) detailed decreases in PM across Europe while Hand et al. (2014, 2019) report significant decreases in PM2.5 mass across the US with larger trends in eastern than in western US. Both these trends were also confirmed by the PM trend analysis in Mortier et al. (2020) and are consistent with our reported scattering trends. Li et al. (2016) used visibility to assess trends in atmospheric haze and aerosol extinction coefficient around the world. The time delay in when the trends switch sign between North America (late 1970s), Europe *(early 1980s) and China (mid 2000s) correlates with SO_2 trends and the trend differences between eastern and western part of US and Europe are consistent with what is presented in our study.*

Many atmospheric aerosols are formed in the atmosphere rather than being directly emitted, so understanding trends in aerosol precursors is also relevant for understanding changes in the atmospheric aerosol. Our study found similar results for scattering as have been found for sulphate trends (Aas et al., 2019), i.e., decreasing sulphate trends across Europe and the US, albeit with the sulphate decrease in Europe beginning before the decrease was observed in the US. Aas et al (2019) also describe potential increases in sulphate in India and increases followed by decreases in SE Asia. Vestreng et al. (2007) monitored the sulphur dioxide emission reduction in Europe and concluded that SO_2 emission reductions were largest in the 1990s with a first decrease in Western Europe in the 1980s followed by a large decrease in Eastern Europe in the 1990s. Similarly Crippa et al. (2016) simulated a larger impact of policy reduction in Western than in Eastern Europe for NO_x, CO, PM_{10} and BC between 1970 and 2010. Likewise, Huang et al. (2017) simulated the non-methane volatile organic compounds emissions and found a rapid decrease in Europe and in North America since the 1990s, whereas the emission of Africa and Asia clearly increased between 1970 and 2012.

4.4.2 Comparison with remote sensing trends

A significant advantage of many REM platforms is their global coverage. Satellites often provide coverage over both land and ocean and the major ground-based REM network AERONET (Holben et al., 1998) is more globally representative than the sites used in this study. However, there are some inherent limitations in comparing aerosol optical property trends from REM retrievals with surface in-situ trends. Our study used aerosol optical measurements made at low RH (typically RH<40%) at the surface, while column aerosol optical retrievals are made at ambient conditions and represent the atmospheric column including layers aloft. Only in the situation of a well-mixed atmosphere, will it be reasonable to compare trends in surface in-situ optical properties with those obtained by ground-based or satellite retrievals. It has also to be mentioned that satellite measurements are less sensitive to the near ground layers containing the greatest aerosol load. Thus, while our trends can be compared with those for column aerosol properties, there is no reason to expect them to be in complete agreement. Below we discuss trends in PM, AOD, column σ_{ap} and column SSA.

Satellites have been used to assess the decreasing PM trends in North America and Europe and also to estimate PM trends in other regions with sparse surface measurements. For example, Nam et al. (2017) evaluated the trend in satellite-derived PM10 over Asia and reported mixed annual trend values depending on the subregion they looked at. Li et al. (2017) found satellite-derived PM2.5 to continuously increase in some parts of Asia (e.g., in India) for the 1989-2013 period - we also find an increasing trend (for aerosol absorption) at the one site we studied in India (MUK). For China, Li et al. (2017) report that the PM2.5 trend transitions from an increasing to a decreasing trend with the transition occurring in the 2006-2008 time period similar to the sulphate trend pattern reported by Aas et al. (2019). The in-situ measurements from China (WLG) and Taiwan (LLN) used in our study are not long enough to detect this transition.

Multiple ground-based REM studies (e.g., Yoon et al., 2016, Wei et al., 2019, Mortier et al., 2020,) report decreasing trends in AOD over the US and Europe with larger decreasing *trends over Europe than over the US, which is the case in our study (see Table 4) as well. The lack of measurements in many regions similar to the lack of representativeness in the surface in-situ aerosol sites discussed in this study (Asia, Africa, South America, etc) are also emphasized. Ningombam et al. (2019) analyse AOD 1995-2018 trends from 53 remote and high altitude sites, of which 21 had ss negative trends. Regionally, Ningombam found primarily negative trends at sites in the US, Europe and polar regions. Their findings for sites in China and India suggested mixed trends with some being positive and some negative in those regions. Some of the sites in Ningombam et al. (2019) were also involved in our study. The trends they find for AOD at LLN and MLO are similar to ours (i.e., not ss trends) at SPO (i.e., ss increasing) and at SGP (i.e., decreasing (note: they refer to SGP as 'car')). Their results are different for IZO (we found no ss trends for scattering while they reported ss decreasing AOD) and at BRW and BIR (we found ss decreasing scattering trends but they found not ss AOD trends).*

Satellite retrievals can offer an even more global picture of aerosol trends than the surface based REM data. Various satellite trend analyses present a picture of trends in aerosol optical depth for different regions of the world that is quite consistent across satellite (and ground-based) AOD datasets. For example, for the satellite literature that we surveyed, all found decreases in AOD over the US and Europe (e.g., Hsu et al., 2012, Mehta et al., 2016, Zhao et al., 2017, Alfaro-Contreras et al., 2017, Wei et al., 2019) consistent with what we have reported for the AOD from ground-based, REM instruments. As we note above, this is also consistent with surface in-situ scattering trends.

There are some discrepancies in the various satellite derived AOD trends over Asia that are likely due to differences in time period of analysis, trend methodology, regional definitions and/or perhaps satellite data product. Nam et al. (2017) found AOD trends varied depending on what part of Asia was being evaluated. Zhao et al. (2017) reported an increasing then decreasing trend over China, which was also suggested by others (e.g., Sogachova et al., 2019; Alfaro-Contreras et al., 2017). Wei et al. (2019) found a slightly negative but statistically insignificant AOD trend for China. Our study found statistically insignificant trends in aerosol loading for both the high-altitude surface site in China (WLG) and in Taiwan (LLN), perhaps because measurements at both these sites span the AOD increase/decrease periods mentioned by Zhao et al. (2017). Over India, increasing trends in satellite AOD were reported by all the literature we surveyed (e.g., Wei et al. 2019; Mehta et al., 2016; Hsu et al., 2012; Alfara-Contreras et al., 2017). This is consistent with our finding of an increasing trend for aerosol absorption for the one Indian site (MUK) in our study.

The satellite measurements also enable evaluation of aerosol loading changes in regions with few to none long-term surface in-situ aerosol optical property measurements. The Middle East exhibited an increasing trend in AOD, while South America exhibited variable trends (e.g., Wei et al., 2019; Metha et al., 2016, Hsu et al., 2012; Alfaro-Contreras et al., 2017). Wei et al. (2019) found a statistically insignificant trend in South America and suggested it was due to complex and changing aerosol sources. Mehta et al. (2016) looked specifically at Brazil and found a decreasing annual AOD trend, but an increasing AOD trend in springtime. Decreasing AOD trends were found over central Africa (Wei et al., 2019), over the African deserts (Metha et al., 2016) and on African coasts (Alfaro-Contreras et al., 2017) regardless if they are dominated by smoke aerosols (southwest) or dust (northwest).

In addition to AOD, trends for other column aerosol property such as column σ_{ap} *and column SSA can be considered. While there appear to be many investigations focusing* *on trends in column aerosol properties other than AOD at individual sites, there are only a few papers that take a more global, multi-site approach (e.g., Li et al., 2014; Zhao et al., 2017; Mortier et al., 2020). There have been several studies related to changes in column σ_{ap} using AERONET REM retrievals. For example, Li et al. (2014) suggest an increase in column σ_{ap} over the US and a decrease over Europe and at most sites in Asia. More recently, Mortier et al. (2020) found ss decreasing σ_{ap} trends in Europe, North America and ss increasing σ_{ap} trends in Asia and Africa. Zhao et al. (2017) used satellite retrievals and reported decreasing trends of column σ_{ap} over both the US and Europe and a not ss column σ_{ap} trend over China. Nam et al. (2018) suggested there was an increasing trend in column extinction Ångström exponent over Asia based on satellite observations. These findings are mostly consistent with our results (Table 4) which indicated decreasing $å_{sp}$ trends in the US and Europe, but perhaps an increasing trend in Asia.*

Comparisons of in-situ and column ω_0 trends are more fraught, because, in addition to the above mentioned caveats related to comparing surface and column measurements, column ω_0 can only be obtained from REM techniques under higher aerosol loading conditions. For example, Kahn and Gaitley (2015) indicate that MISR SSA retrieval requires AOD>0.15-0.2. Similarly, AERONET retrievals require AOD (at 440 nm) > 0.4 (Dubovik et al., 2000). This limits the sites for which column SSA can be retrieved. Andrews et al. (2017) present a plot derived from global model simulations suggesting more than 80% of the globe has annual AOD values below 0.2, and, indeed, many of the surface in-situ sites discussed here are in remote locations with annual AOD consistently below 0.2. Andrews et al. (2017) also suggest there is a systematic variability of SSA with loading that might result in column SSA biases if retrievals are constrained to higher levels of AOD. With these caveats in mind, we can compare our surface ω_0 trend results with satellite column ω_0 trends.

Li et al. (2014) studied 2000-2013 trends in column ω_0 at select AERONET sites. Their findings suggest that column ω_0 is increasing in the US, Europe and Asia. However, they noted the uncertainty in these trends is high because they used level 1.5 data (AOD<0.4) in order to have enough data points for their analysis. Zhao et al. (2017) utilized satellite retrievals and reported decreasing trends in column ω_0 over the eastern US and Europe and a not statistically significant trend over China for the 2001-2015 period. Their results over the US and western Europe are consistent with the overall regional ω_0 trends reported in this study (i.e., Table 4), although Figure 7 suggests there is a fair amount of variability in the surface ω_0 trends at the individual sites in these two regions. Our study found an increasing trend in ω_0 at the surface in Asia (based on 3 sites), which is consistent with Li et al.'s column ω_0 trend but not with the lack of trend in column ω_0 over China suggested by Zhao et al. (2017). But, as noted above, remote sensing retrievals of column SSA should be considered with caution and, clearly, further effort in column SSA trend analysis is warranted.

4.4.3 Causality

While it is beyond the scope of this effort to explore in depth the causes of the observed trends of aerosol optical properties, some general comments can be made. First, tendencies in regional trends for variables representing aerosol loading (e.g., surface in-situ aerosol scattering, PM, and AOD) are generally consistent across multiple datasets. Overall, the main cause of observed decreasing trends in loading is likely strong reduction of both primary aerosols and precursors of secondary aerosol formation connected to mitigation strategies on regional to continental scales (e.g., Huang et al., 2017; Crippa et *al., 2016; Pandolfi et al., 2016; Vestreng et al., 2007). Detailed analysis of PM reductions and composition changes in Europe and the US have enabled attribution of the trends to changes in source types and emission levels (e.g., Hand et al., 2019; Pandolfi et al., 2016, Ealo et al., 2018).*

The explanations of the trends based on long-term measurements are supported by modelling efforts. Like many satellite retrievals, model simulations also provide global coverage and, in addition, can be used to investigate reasons for observed changes in aerosol. Model simulations described in Li et al. (2017) suggested that the decrease in PM2.5 in western and central Europe is principally due to sulfate, whereas in eastern Europe decreases in organic aerosol also plays a role. The EMEP status report (2019) notes that the difference in emissions trends between western and eastern Europe has become more significant since 2010. Further, the EMEP status report suggests that estimated increasing emissions of all pollutants since 2000 in the eastern Europe are mainly influenced by emission estimates for the remaining Asian areas in the EMEP modelled domain. Similarly, Zhao et al. (2019) used a model to attribute the AOD, ω_0 and $å_{sp}$ decreases in North America and Europe to considerable emission reductions in all major pollutants except in mineral dust and ammonia.

For Asia, modelling by Li et al. (2017) suggests aerosol changes are principally related to increases in organic aerosol and secondary inorganic aerosol, whereas the increases in BC, nitrate and ammonium are comparably moderate. Yoon et al. (2016) use a model to ascribe the observed increases in AOD over India to increases in BC and water soluble materials - both related to anthropogenic emissions. Over China, Yoon et al. (2016) observe a disconnect between the model chemical composition and the measured AOD which they explain by noting that the measurement sites they rely on in the region are far from the population centers where most of the emissions occur. Zhao et al. (2019) use a model to attribute the increase in AOD followed by a decrease in AOD to emission increases induced by rapid economic development until 2008-2009 followed by decreases in both anthropogenic primary aerosols and aerosol precursor gases.

Zhao et al. (2017) suggest that the larger reductions in aerosol precursors (e.g., SO2 and NO_x emissions) rather than primary aerosols, including mineral dust and black carbon can explain the decreases in ω_0 and $å_{sp}$ observed over Europe and the US. This is because the secondary aerosols formed from such precursors tend to be primarily scattering, so less secondary aerosol would change the relative balance between scattering and absorption driving ω_0 down. Similarly, secondary aerosol particles tend to be small so a decreasing trend in secondary aerosol would change the relative contribution of small to large particles in the aerosol size distribution and lead to a decreasing trend in $å_{sp}$. In contrast, in Asia simultaneous increases in aerosol precursors and BC before 2006, and a simultaneous decrease after 2011 explains the trends ω_0 and $å_{sp}$ they observed there. Modifications in emissions of aerosol precursors also impact the atmospheric chemistry leading to non-linear response of the formation of secondary inorganic aerosol (Banzhaf et al., 2015).

While regional changes in emissions are one driving factor in trends, because of long-range transport, out of region changes in sources also have the potential to affect trends. For example, Saharan dust impacts CPR, IZO and UGR (e.g., Denjean et al., 2016; Rodriguez et al., 2011; Garcia et al., 2017; Lyamani et al., 2008) and its emissions may change (decrease) in a warmer world (Evan et al., 2016). Other examples of sites clearly impacted by long range transport include IZO impacted by North African pollution due to developing industries (Rodriguez et al., 2011), and the high altitude station of MLO which *is impacted by Asian pollution (e.g., Perry et al., 1999). Mountainous stations can also be affected by modifications of the planetary boundary layer or of the continuous aerosol layer heights responding to ground temperature or mesoscale synoptic weather changes (e.g., Collaud Coen et al., 2018 and references therein).*

The oscillation in trend sign for several variables at the Arctic sites is potentially caused by the very low aerosol loading, but the Arctic region is changing rapidly and the impact of evolving transport patterns, atmospheric removal processes or local sources cannot be excluded (e.g., Willis et al., 2018) and requires closer study.

While both increasing and decreasing levels of aerosol due to changes in anthropogenic emissions have been observed, the role of non-anthropogenic sources may become more important in the future. For example, climate change also affects soil drought and the positive feedback between drought and wildfires can also affect aerosol optical properties (Hallar et al., 2017, McClure and Jaffe, 2018). The number and intensity of wildfires is increasing in several regions (e.g., Moreira et al., 2020; Turco et al., 2018; Hand et al., 2014). McClure and Jaffe (2018) confirmed an increasing trend of $PM_{2.5}$ 98 percentiles in northwest US due to an increase in wildfires superimposed on the global decrease in anthropogenic emissions. Yoon et al. (2016) also note an increase in extreme AOD events in the western US, which they hypothesize could be due to wildfires. Another example of potential changes in natural aerosol may take place in the Arctic, where decreases in sea ice coverage might play a role in natural aerosol increases in the region (e.g., Willis et al., 2018) (decreases in sea ice coverage may also lead to enhanced anthropogenic emissions due to increased human activity (e.g., Aliabadi et al., 2015)). Whether such changes in natural aerosol emissions lead to observable changes in overall aerosol trends or trends at the extremes of aerosol loading is something to look for in future trend analyses.

Detailed studies at each station are necessary to discriminate between direct causes like changes in anthropogenic emissions, and indirect causes related to general climate changes such as drought, changes in surface albedo, biogenic aerosol concentration, atmospheric chemistry, sea ice coverage or atmospheric circulation patterns. The availability of the homogenized data set from this study will provide a useful tool for these types of analyses.

In order to get a truly global overview of aerosol trends, surface in-situ measurements need to be paired with model simulations and satellite observations. This will enable evaluation of the uncertainty in regional and global trends based on deficiencies in spatial and/or temporal coverage. Satellites and models are able to fill the gaps in coverage from ground-based measurements, but both rely on surface measurements for ground truth."

4. Finally, I have two lesser general concerns. The first is to think about how extreme events can affect trends. There are aerosol events that are so large they can change even a decade-long trend with a single event. I would bet that the recent Australian fires were big enough to change the trends for that region, indeed large swaths of the Southern Hemisphere, for an entire decade. The 1997-98 Indonesian fires also were big enough. There may be other such events in your data series. This doesn't disprove the validity of your statistical analysis, just look at the time series and comment if appropriate.

Extreme events have an increasing occurrence caused by the global warming. It would have been very interesting to study the long-term trends of the extremes of in-situ aerosol parameters. Some cited studies in the western USA (McClure and Jaffe, 2018) show that the average $PM_{2,5}$ were not affected by the increasing biomass burning events but that the extremes (98 percentiles) were largely affected. This study was not designed to study regional effects but to bring together all the in-situ aerosol long-term trends results in a solely and homogeneous study. The time delay and the already large size of the paper do not allow us to study the trends of the extreme events. Similarly, trend in coarse mode aerosol (PM_{10}-PM_1) would also be very interesting, but were skipped for the same reason.

5. The other general comment is that most of section 4.4 is speculation without supporting evidence.

    As already mentioned as an answer to comments 3, section 4.4 was completely re-written in order to include more satellite and model studies. Sect 4.4 is now divided in sub-sections dealing with 1) the comparison with other in-situ results, 2) the comparison with remote sensing instruments and finally 3) the causality. Model studies are mostly cited in the last sub-section. These modifications allow to better comparing the results of this study with global studies.