Anonymous Referee #1

Received and published: 27 October 2017

This is original analysis air pollution trends in China based on aerosol extinction measurements. The paper is well structured and clearly written. I do not find any scientific errors in methods or data interpretation. I recommend this paper to published in ACP after considering the following minor issues.

Thank you very much to the reviewer for his/her encouraging comments on our paper! We have replied to the specific comments below, and have also revised the manuscript accordingly.

Scientific issues:

Lines 53-55: There is at least one more point why remote sensing observations are problematic here: they do not easily distinguish between different mixed-layer height, which is a major parameters affecting surface air pollution.

Thanks a lot for this note. We added a discussion here as follows:

"Moreover, remote sensing techniques cannot recognize mixed layer height, a major parameter affecting surface air pollution, which make them unsuitable for air quality studies."

Lines 61-63. The authors compared visibility observations against remote sensing here. How about in situ measurements of air pollution vs. visibility observations? I suppose that there are clear differences in terms of both spatial coverage and length of time series. I would like to see in situ measurement shortly (couple of lines) mentioned in this context as well.

We did compare visibility converted AEC data against in situ $PM_{2.5}$ and PM_{10} observations in our previous paper:

Li, J., C. Li, C. Zhao, and T. Su (2016), Changes in surface aerosol extinction trends over China during 1980–2013 inferred from quality-controlled visibility data, *Geophys. Res. Lett.*, 43, 8713–8719, doi:10.1002/2016GL070201.

The figure is attached below:

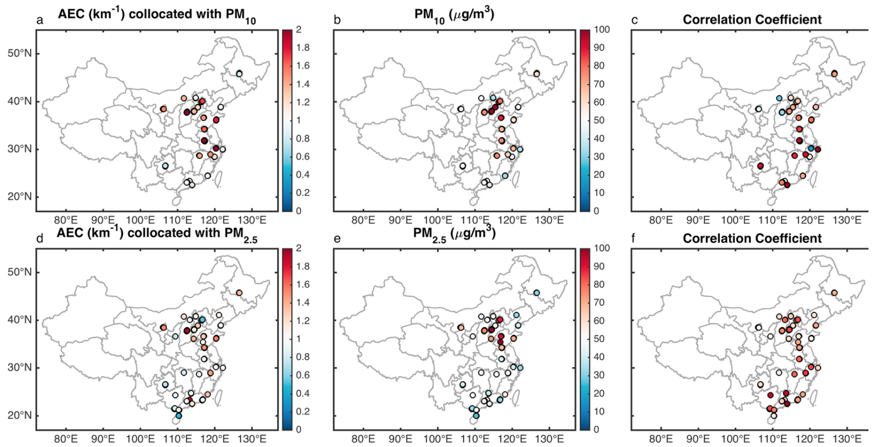


Figure R1. Validation of visibility converted aerosol extinction coefficient (AEC, km^{-1}) against PM_{10} (μg/m3) and $PM_{2.5}$ (μg/m3) concentration. (a–c) The spatial distribution of AEC collocates with PM_{10} stations, distribution of PM_{10}, and the correlation between AEC and PM_{10} calculated using daily mean data. (d–f) The same information for AEC and $PM_{2.5}$. (Figure source is *Li et al.*, 2016)

We also added a phrase here that the visibility data used in this study compare well with surface air quality measurements. In the end of Section 2.1 we also mentioned that "This AEC dataset has also been validated against surface $PM_{2.5}$ and PM_{10} measurements. Please refer to *Li et al.* (2016) for detailed description of the correction and validation processes."

Technical issues:

The use of tense is not in a good balance in the abstract. I would recommend the authors to consider this point carefully and make the necessary revisions.

Both AEC and its trend have a unit. It seems that these units have been scaled out somehow from figures 1-4, making it impossible interprete the real magnitude of AEC (or its trend) from these figures. The authors should add this information.

We are sorry for the confusion. The reviewer is correct in that the unit of AEC has been scaled out, i.e., the trends in this paper actually refer to relative changes. This is because the absolute magnitudes of mean and extreme AEC are different and thus the absolute trends cannot be directly compared. We have ensured that the term trend is consistent throughout the manuscript. We have also added an explanation in the abstract that the magnitude of the trends are “expressed in terms of relative changes”.

line 266-267: studies – > studied...remains to be understood whether...

Corrected.

Anonymous Referee #2

Received and published: 19 December 2017

GENERAL

The paper presents trend analyses of aerosol extinction coefficient at numerous measurement sites in China. Different methods for calculating trends are compared. The analysis also compares trends in major areas of China. The paper is very interesting, I can definitely recommend publishing it in ACP. I did not find any very big errors in the paper. However, there are some points that need to be explained in more detail and some points that should be changed. The changes I am suggesting are minor, mainly clarifications.

We thank the reviewer for his/her positive and helpful comments on our paper! We have addressed the detailed comments point-by-point below, and have also revised the paper accordingly.

DETAILED COMMENTS

The most important point that should be changed is this: the Aerosol Extinction Coefficient (AEC) that is used for the analyses is not unitless like it is presented all over the paper and the supplement. AEC comes from the Koschmieder formula (visibility = 3.912/AEC) and visibility is given in units of length. So, the unit of AEC is inverse units *of length, for instance inverse meters or inverse kilometers or inverse megameters. In polluted areas of China extinction coefficient is typically in the range of some hundreds of inverse megameters. Go through the paper and the supplement and present the units of AEC everywhere both in the text and the figures. This is important also since the AEC values are something that link the paper's values better to the rest of the world.*

We are sorry for the confusion. Yes, AEC has unit km^{-1}. However, because the absolute value of extreme and mean AEC can be orders of magnitude different which makes their absolute trends incomparable, we report all trends in terms of relative changes. Therefore, the trends throughout the paper are unitless. We have clarified this point in both the abstract (line 13) and the main text (lines 158-160).

There are no tables. Give the main results in 1 or 2 tables. For instance trends within each major region obtained with the different methods. Tables give you also more references because they can easily be compared with by other authors.

Thanks for this suggestion. We added a table displaying the extreme and mean trends and their differences for the major regions. We also added the same table for seasonal trends in the supplementary material.

In the figures, give units for the color bars, if they have a unit. And if they are unitless, *give an explanation of the color scales in the captions. Now there are no explanations of the colorbars in any figure.*

OK. We added statements in the figure captions that the trends are unitless (line 437).

L108 " ... annual or seasonal time series of the 95th percentile of the extinction coefficients ..." There is nowhere mentioned, what is time resolution of the data. So, does this mean the 95th percentile of one-minute or hourly or daily averaged AEC in any given year? How do you define each season?

Again sorry for the confusion here. We use daily averages (averaged from hourly data) and the percentiles refer to those of daily averages. We added an explanation as follows in lines 110-111:

"The hourly AEC is first averaged to daily values and 95th percentile (50th percentile for the median trend) is then calculated for each year or each season for the seasonal analysis."

The seasons are defined as: March, April and May for spring, June, July August for summer, September, October and November for autumn and December, January and February for winter, which are stated in the parenthesis of lines 260-263.

L111. In eq (1) there is Xi and Xj. They must be the AEC values at i and j. Then b has the same unit as AEC. Or is it - as I would assume - that b has the units of AEC divided by the units of time, for instance inverse meters in a year if i – j in eq. (1) means time step. Does it?

b has unit of AEC divided by units of time. We use annual percentiles of AEC so the unit of b is km^{-1}/yr. However, b is not the final trend reported as we converted it to relative changes, which is b times the number of years divided by the AEC value of the starting year (1980). This point is clarified in lines 158-160.

The quantile regression has the formula (9). Is the beta in formula 9 the trend? If it is, write it explicitly out. If it is not, in which formula is it? And further, does it also have units? It should if it is to be compared with b of eq. (1).

Yes. The beta in (9) is the slope of the trend and its unit is also km^{-1}/yr. The final trend of the quantile regression model is also reported as relative changes.

Figure 2 shows the probability density functions of AEC in different regions. It is very interesting. But the same issue applies to this plot also: units. AEC has units of inverse length, e.g., inverse megameters. So, I would recommend presenting the picture so that you simply show the x-axis as inverse megameters but use a logarithmic scale.

That would also help in comparing the data with the rest of the world. Another issue in this figure is the values of the pdfs. The integral of a pdf should equal 1. Now there are values larger than 1 so the integrals are definitely > 1. Explain in detail what the y axes mean. And do corrections if needed. Further in the same figure: if you calculate a pdf like that, the data are divided into bins of AEC and then you present how large a fraction of data is in each bin. What is the bin division you used?

Figure 2 shows the distribution of the absolute AEC values so they have unit km^{-1}. We have changed the x axis to inversion megameters in logarithmic scale and added the unit in the figure caption. We have double checked that the integral of all pdf functions are indeed 1. Values greater than 1 are reasonable to appear in the pdf and this does not necessarily mean the integral is greater than 1 (for integral you have to multiply the x axis). In fact, pdf can take any non-negative values as long as the integral equals to 1. The pdf only reflects the probability density for some interval rather than the real probability. For example, in one pdf in Figure 2 the y axis is 1.35 and the x axis is 6.21. This means the probability between some very small interval around 6.21, say between 6.20 and 6.22 is 1.35*(6.22-6.20)=0.027. This number must not exceed 1. To produce Figure 2 we use 20 bins ranging from the minimum log(AEC) to the maximum log(AEC) to calculate each pdf.

Different trends between extreme and median surface aerosol extinction coefficients over China inferred from quality controlled visibility data

Jing Li[1,*], Chengcai Li[1], Chunsheng Zhao[1]

Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China, 100871

Abstract

Although the temporal changes of aerosol properties have been widely investigated, the majority focused on the averaged condition without much emphasis on the extremes. However, the latter can be more important in terms of human health and climate change. This study uses a previously validated, quality-controlled visibility dataset to investigate the long-term trends (expressed in terms of relative changes) of extreme surface aerosol extinction coefficient (AEC) over China, and compare them with the median trends. Two methods are used to independently evaluate the trends, which arrive at consistent results. The sign of extreme and median trends are generally coherent, whereas their magnitudes show distinct spatial and temporal differences. In the 1980s, an overall positive trend is found throughout China with the extreme trend exceeding the mean trend, except for Northwest China and the North China Plain. In the 1990s, AEC over Northeast and Northwest China starts to decline while the rest of the country still exhibits an increase. The extreme trends continue to dominate in the south while it yields to the mean trend in

the north. After year 2000, the extreme trend becomes weaker than the mean trend overall in terms of both the magnitude and significance level. The annual trend can be primarily attributed to winter and fall trends. The results suggest that the decadal changes of pollution in China may be governed by different mechanisms. Synoptic conditions that often result in extreme air quality changes might dominate in the 1980s, whereas emission increase might be the main factor for the 2000s.

1. Introduction

As a by-product of the rapid industrial and economic development, China has been faced with a serious issue of air pollution. The variability and trends of China's air quality or aerosol properties have become the focus of numerous past studies (*Jinhuan and Liquan*, 2000; *Che et al.*, 2007; *Deng et al.*, 2008; *Streets et al.*, 2008; *Yoon et al.*, 2011; *Guo et al.*, 2011; *Zhang et al.*, 2015). While many of these works reached important conclusions about the temporal evolution of China's pollution, the majority only analyzed the arithmetic means (e.g., monthly or annual means of aerosol optical depth), with little attention paid to the extreme values. However, it is often these extremes that are responsible for many health and climate related aftermaths. Additionally, considering that the distribution of aerosol optical properties, such as aerosol optical depth (AOD) and extinction coefficients, are often highly right-skewed (*O'Neill et al.*, 2000; *Collaud Coen et al.*, 2013; *Yoon et al.*, 2016), analyzing the arithmetic mean tends to discard the large portion of information in the long tails, thus biasing the result. Moreover, as indicated by previous studies, extreme pollution events are often associated with abnormal synoptic conditions (*Zheng et al.*, 2015; *Ye et al.*, 2016), whereas the mean should be more prone to changes in the emission which increases pollution level overall. Therefore, analyzing the changes in both the mean and extreme values would help understand the factors influencing the variability of pollution.

For the few studies that did address temporal changes in the percentiles of aerosol loading, usually either satellite or surface based remote sensing measurements are used, such as Aerosol Optical Depth (AOD) retrievals from Moderate Resolution Imaging Spectroradiometer (MODIS, *Sullivan et al.*, 2015) or the Aerosol Robotic Network (AERONET, *Xia*, 2011; *Yoon et al.*, 2016). Nonetheless, remote sensing data is not ideal for extreme analysis, mainly because it frequently misses heavy pollution conditions likely associated with strict cloud screening (*Lin and Li*, 2016). Moreover, remote sensing techniques cannot recognize mixed layer height, a major parameter affecting surface air pollution, which make them unsuitable for air quality studies. As a result, the "real" extremely high aerosol loadings cannot be well detected using remote sensing. On the other hand, surface visibility observations that do not require cloud screening or other retrieval assumptions, can serve as a suitable alternative for pollution related research. After eliminating fog, rain or snow conditions, degradation of surface visibility can be mainly attributed to aerosol extinction and are thus closely related to air quality (*Husar et al.*, 2000). Moreover, since routine visibility observation started as early as 1970s for many sites, these data can offer a much longer time series for trend analysis than remote sensing products. Previously, *Li et al.* (2016) used a quality controlled (by comparing against surface PM_{10} and $PM_{2.5}$ measurements) visibility converted Aerosol Extinction Coefficient (AEC) dataset to study temporal changes of monthly mean surface aerosol extinction in China for the past 30 years and found that there are obvious shifts in the trends for different time periods. However, it still remains to understand whether the extreme values change faster or slower than the mean.

In this paper, we use the same dataset as in *Li et al.* (2016) to further investigate the trends of extremely high (defined as the 95th percentile) surface aerosol extinction coefficients and compare them with the median trends representing averaged condition. Although a threshold visibility value is often used in previous studies to define extreme events (e.g., *Fu et al.* 2013 define extreme pollution as visibility lower than 5 km conditions), the same threshold does not apply to all sites since their reporting conventions may be different. We thus believe a percentile criterion would be more appropriate. In addition to estimating the linear trend of the 95th percentile value itself, we also use a novel method proposed by *Franzke* (2013) based on quantile regression with surrogate data testing for significance, who used this method to test for significant trends in extreme temperatures. To our knowledge, this method has not been applied to aerosol related research, and the independent application of two methods increases the robustness of the results.

In section 2, we describe the data and method used in this study. The analysis results are presented in section 3, followed by the conclusions and a brief discussion in section 4.

2. Data and Methods

2.1 Visibility data

Here we use the same visibility dataset as in *Li et al.* (2016). This hourly surface visibility dataset is obtained from the National Centers for Environmental Information (NCDC, http://www1.ncdc.noaa.gov/pub/data/noaa/) of the National Oceanic and Atmospheric Administration (NOAA). The data selection criteria and quality control procedure strictly follows those implemented by *Li et al.* (2016). Briefly, data before 1980 is not used because of different reporting standard (*Che et al.*, 2007; *Wu et al.*, 2012). Those after 2013 are also excluded because many sites have replaced human observation with automatic visibility sensors. Then the eight quality assurance steps proposed by *Li et al.* (2016) is applied to the dataset. A total of 272 sites are selected for China, whose data have been manually inspected to show no observable jumps or spikes.

The visibility is further converted to Aerosol Extinction Coefficient (AEC) using the Koschmieder formula (*Koschmieder*, 1926), and corrected for relative humidity effects according to *Husar and Holloway* (1984) and *Che et al.* (2007). This AEC dataset has also been validated against surface $PM_{2.5}$ and PM_{10} measurements. Please refer to *Li et al.* (2016) for detailed description of the correction and validation processes.

2.2 Trend analysis methods

We define extremes as the 95th percentile of the visibility converted surface AEC. The hourly AEC is first averaged to daily values and 95th percentile (50th percentile for the median trend) is then calculated for each year or each season for the seasonal analysis. To estimate trend of the extremes, we use two independent methods. The first is to obtain an annual or seasonal time series of the 95th percentile of the extinction coefficients and then perform a *Sen*'s slope (*Sen*, 1968) estimate of its linear trend. The *Sen*'s slope b is calculated as

$$b = \mathrm{Median}\left(\frac{X_i - X_j}{i - j}\right) \forall j < i \qquad (1)$$

where X_i and X_j are the ith and jth value in the time series respectively.

Then the Mann-Kendall statistical test (*Mann*, 1945; *Kendall*, 1975) is applied to test whether the trend is significant at 95% level. The test statistic is calculated as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(X_j - X_i) \qquad (2)$$

Where n is the number of data points, and sgn is the sign function:

$$\mathrm{sgn}(X_j - X_i) = \begin{cases} +1 & \text{if} \quad X_j > X_i \\ 0 & \text{if} \quad X_j = X_i \\ -1 & \text{if} \quad X_j < X_i \end{cases} \qquad (3)$$

The variance of S is given by

Formatted: Superscript

Formatted: Superscript

$$\mathrm{Var}(S) = \frac{1}{18} n(n-1)(2n+5) \tag{4}$$

If the sample size n>30, which is well satisfied in our case, the standard normal test statistic ZS is computed using:

$$ZS = \begin{cases} \frac{S-1}{\sqrt{\mathrm{Var}(S)}} & \text{if } S>0 \\ 0 & \text{if } S=0 \\ \frac{S+1}{\sqrt{\mathrm{Var}(S)}} & \text{if } S<0 \end{cases} \tag{5}$$

According to the normal distribution table, the 5% significance level is satisfied if $|ZS|>1.96$.

The second approach is quantile regression, which is a well established method used in many previous studies (*Koenker and Hallock*, 2001; *Hannachi*, 2006; *Barbosa et al.*, 2011; *Donner et al.*, 2012; *Franzke*, 2013) to estimate extreme trends of climate data.

For regular linear least square regression, the model can be expressed as

$$E[y \mid \mathbf{X}] = \beta\mathbf{X} + \varepsilon \tag{6}$$

where y is the response variable conditioned on $\mathbf{X}$, and the β's satisfy the minimization of the summed error function

$$err = \min \sum_i \xi(y_i - \beta X_i) \tag{7}$$

where

$$\xi(u) = u^2 \tag{8}$$

For linear quantile regression, the response variable becomes the τ th ($\tau \in [0,1]$) quantile of y conditioned on $\mathbf{X}$,

$$Q_\tau[y \mid \mathbf{X}] = \beta\mathbf{X} + \varepsilon \qquad (9)$$

where the β s still satisfy equation (2), but equation (3) now becomes

$$\xi_\tau(u) = \begin{cases} u\tau & u \ge 0 \\ u(\tau - 1) & u < 0 \end{cases} \qquad (10)$$

Note that ξ_τ is symmetric when $\tau = 0.5$, rotated to the right when $\tau < 0.5$ and to the left when $\tau > 0.5$. The quantile regression problem can be numerically solved by linear programming (*Koenker and Hallock*, 2001). Here we use the R package "quantreg" to solve for the regression coefficients of daily mean AEC. Trends for both the 95th and 50th (median) percentiles are estimated and the trends are compared. To test for significance of the quantile regression trends, we adopt the bootstrap approach proposed by *Franzke* (2013), who used surrogate data generated with the same autocorrelation function and the same probability density function as the original dataset. The detailed generation procedure can be found in *Schreiber and Schmitz* (1996) and *Franzke* (2013). Here we generate 1000 surrogate time series to represent the intrinsic variability of the AEC time series.

In addition, we also calculate the trends for the median AEC (50th percentile) using the above two methods, and compare them with the extreme trends. All trends are normalized and expressed as relative changes per decade, calculated as trend slope times the length of the time series divided by the corresponding AEC percentiles of the initial year. Therefore the trends reported are unitless.

Deleted: percentage

Figure S1 in the supplement shows an example of the trend analysis using these two methods. In the following text, to save space we only present trends using quantile regression, whereas the *Sen's* slope results, which agree well with the former, are presented in the supplement material.

3. Results

3.1 Trend Maps

We first examine the distribution and temporal changes of trends for all sites in China. As indicated by *Li et al.* (2016), there are significant temporal shifts of the magnitude and sign of monthly mean AEC trends for different decades. We thus also respectively examine the extreme and median trends for three consecutive decades: 1980-1990, 1991-2000 and 2000-2013. The overall trends for the 1980 to 2013 period are weakly positive for the majority of the sites (see Figure S2).

The three columns in Figure 1 show the distribution of extreme trend (upper row), median trend (middle row) and their differences (extreme minus median, bottom row) for the 272 sites for the three periods respectively. To avoid the confusion caused by positive and negative signs of the trend, the difference here are calculated using the absolute value of the extreme and median trends. Larger dots in black circles mean that the trends are statistically significant at 95% level. Figure 1 is the results from quantile regression, whereas the trends using *Sen*'s slope is presented in Figure S3, which shows largely consistent pattern. It is seen from Figure 1 that the sign of median and extreme trends mostly agree throughout China. An extensive positive trend is observed all over China in the 1980s. During the 1990s, many sites, especially those in north China, began to experience a decreased AEC. After year 2000, the north China sites continue to show decreasing trends whereas AEC over many south China sites started to rise again.

However, a detailed comparison between median and extreme trends reveals distinct spatial and temporal differences. Focusing on the bottom three panels of Figure 1 (g-h), it is clear that in the 1980s, the extreme trends exceed the median trend throughout China, with some differences as large as 50% (northwestern sites). The number of sites showing significant extreme trend (178) is also greater than those with significant median trend (91). Note that the number of significant sites can be different between quantile regression and *Sen*'s slope results, because (1) quantile regression is applied to daily data while *Sen*'s slope uses annual or seasonal percentiles and (2) quantile regression uses bootstrap method to test for significance while *Sen*'s slope uses MK test. Nonetheless, the spatial patterns of the two methods are consistent. In the 1990s, the distribution of the trend differences switched to a north-south "dipole" pattern, with negative values in the north and positive in the south in general, i.e., extreme trends are weaker than the median trend in the north but stronger in the south, with a rough separation at 33°N marked by the horizontal black line on Figure 1h. In the north, the sites showing significant extreme trends also becomes fewer than those with significant median trends in the north. Even in the south, the difference between the extreme and median trends is much smaller compared to the 1980s, indicating a slowdown of the increase in the extreme values. After year 2000, almost the entire China exhibits a “blue” pattern as opposed to the “red” pattern in the 1980s. Except for a few sites in central south China, the majority exhibits a weaker extreme trend than the mean trend. There are also fewer sites showing significant trends in the extreme (52) than in the median (119). This feature is particularly strong for northeast, northwest and south China. Although east and south China still show positive AEC trends, this result suggest that in this decade, the extreme pollution conditions have not increased as much as the mean or background pollution.

In short, the positive trends in the 1980s over China can be primarily attributed an increase in the extremes. The 1990s experienced with a transition, with extreme trends becoming weaker than the median trend in the north and only slightly stronger in the south. Finally in the 2000s, the extreme trends largely yield to the median trends.

3.2 Regional Trends

To examine the spatial and temporal changes in more detail, we further divide the country into six representative regions, marked by black rectangles on Figure 1b. Three of these regions: the North China Plain (NCP), Yangtze River Delta (YRD), and Pearl River Delta (PRD) are the major urban conglomerates in China. Since the change in the extreme and median is essentially related to the shift of the distribution, we first evaluate the regional AEC (/Mm^{-1}) distributions for the three decades. Figure 2 plots these distributions by region on logarithmic scale, as AEC is usually considered to follow a lognormal distribution (*Collaud Coen et al.*, 2013). The dashed lines in Figure 2 indicate location of the 95th percentile. For all regions, there is a rightward extension of the tail of the distribution from 1990s to 1980s, implying an increase of the extremes, which is also characterized by the rightward shift of the 95th percentile line. NCP, YRD and NW China also show a rightward shift of the distribution peak. From 1990s to 2000s, although the distribution peak shifts to the right for PRD, YRD, SW China and NE China, there is no obvious shift in the tail for these four regions. For the other two regions, NCP and NW China, there is a leftward shift in both the peak and the tail, but the shift of the peak is stronger. Overall, we can roughly conclude that the 1980s' AEC trend is characterized by a change of the extremes, while in the 2000s the median dominates the trend.

Consistent with *Li et al.* (2016), we also calculate trends successively for all periods starting each year from 1980 to 2004 and ends in 2013 with 10-year increments. Figure 3 shows the temporal evolution of the quantile regression trend differences with x axis indicating the trend calculation start year and y axis indicating the length of the time series, with its counterpart using *Sen's* slope shown by Figure S4. To save space, only the absolute differences between the extreme and median trends are presented in Figure 3, while their respective values are shown in Figures S5 and S6 for quantile regression and Figures S7 and S8 for *Sen*'s slope. Table 1 displays the regional extreme and median trends using the two methods and their differences for the three periods: 1980-1990, 1991-2000, 2001-2013. Note although the absolute values of Sen's slope and quantile regression trends can be different, their signs are consistent. The time series and linear trends for each region are presented in Figure S9. Because in Figure 3 the trends are calculated successively for each period, it helps to examine the time node of the changes more precisely. For example, although Figures 1 and 2 both indicate that the extremes increase more rapidly in the 1980s, for YRD and PRD, the duration is short with the extreme trend exceeding the median trend since around 1982, while for the rest four regions the change happened around 1986 or later. YRD, PRD and NE China experienced a short period of stronger extreme trend from ~1994 to 1996, whereas the other three regions show weaker extreme trends. After 2002, SW and NW China display a slightly higher extreme trend, which is different from the rest four regions. These features suggest that there can be minor differences when the trends are examined for different time periods.

The seasonal time series of the difference between extreme and median quantile regression trends are plotted in Figure 4, with a 4-year moving average to smooth out small wrinkles (its counterpart using *Sen's* slope is shown in Figure S10). Note that Figure 4 shows the evolution of the trend difference for every ten-year period from 1980 to 2004 (i.e, 1980-1989, 1981-1990,…, 2004-2013). An outstanding feature in Figure 4 is that for all regions, the summer (JJA) trend difference (indicated by red curves) exhibit quite different, or even reversed variability from the other three seasons and the annual result. For NE, NW China and the PRD, spring (MAM) trends also have relatively larger departure. In general, winter (DJF) and fall (SON) trends agree better with the annual trend. Since these two seasons are dominated by anthropogenic aerosols such as sulfate, nitrate, black and organic carbon throughout China (*Cao et al.*, 2007; *Wang et al.*, 2007; *Wang et al.*, 2015), the results indicate that changes in anthropogenic aerosol loading are primarily responsible for the observed extreme and median trends. In the spring many regions are influenced by dust, and in the summer, the relative humidity effect may significantly enhance aerosol extinction. Both are natural factors and should have minor contribution to the annual trend according to Figure 4.

4. Conclusions and Discussion

While the trends of aerosol pollution in China have been studied extensively, it remains to understand whether the extreme conditions have changed and whether their changes are faster or slower than the mean. In this study, we use a quality controlled visibility dataset to examine decadal trends of extreme values of surface aerosol extinction coefficients. Quantile regression and *Sen*'s slope estimates are jointly used to estimate the trends to improve its robustness. Our analysis reveals that in general, the extreme and median trends agree in terms of the sign, but they can differ significantly in terms of the amplitude. During the 1980s, the extremes increased faster than the median for most China except for a few north and northwest sites. The 1990s experienced a transition with extreme trend becoming weaker than median trend in the north but still slightly stronger in the south. Then in the 2000s, the majority of the country exhibited a

Deleted: s

weaker extreme trend than the median trend. Seasonally, winter and fall trends are the most consistent with annual trends, while the summer trend shows the largest departure from the annual trend.

This study uses daily mean daytime AEC without accounting for its diurnal variability. Nonetheless, visibility can still change considerably in the course of a day (*Deng et al.*, 2011). To examine this effect we repeat the analysis using daily minimum and daily maximum AEC respectively. Their counterparts of Figure 1 are shown in Figures S11 and S12. A brief comparison indicates high resemblance of these two figures to Figure 1 that uses daily mean data, albeit with some reasonable differences in the amplitude.

The reason for the different behaviors between the extreme and median trends still needs further investigation, and will be the topic of our future study. Some implication is that in the 1980s and part of 1990s, synoptic conditions might be playing a major role in modulating aerosol variability. For example, several extremely heavy pollution events are believed to be linked to stagnant weather (*Tao et al.*, 2014; *Zheng et al.*, 2015). After mid 1990s, emission might become more dominate which tends to increase both the extreme and the mean. But since it is a relatively uniform background change, the signal might be more prominent in the mean condition. On the other hand, aerosol properties can also be potentially influenced by decadal or interannual climate variability (*Chen and Wang*, 2015; *Wang and Chen*, 2016), whose footprint may be embedded in these extreme and mean trends. However, the mechanism that they impact on the extremes and the mean still need to be understood, and likely require a comprehensive study using both observations and model simulations. This also requires the models to accurately simulate the extreme events, which is a challenging task.

Admittedly, the visibility data is not ideal for aerosol-related studies, given its various sources of uncertainties as discussed in *Li et al.* (2016). However, it is a currently best compromise since there is lack of reliable long-term aerosol observation datasets. Moreover, remote sensing produces are vulnerable to extreme pollution, making them unsuitable for extreme trend studies. For example, as discussed in *Lin and Li* (2013), MODIS frequently misses the heavy haze over north China likely due to cloud screening algorithm. Sun photometers will also stop working when the sun is blocked by the heavy pollution. This also suggests that current remote sensing instruments and retrieval algorithms need to be improved to observe these extreme events.

Acknowledgements

We thank the NOAA NCDC database for providing the hourly visibility measurements used for this study. The data is downloaded from the NCDC public ftp at http://www1.ncdc.noaa.gov/pub/data/noaa/. This work is funded by National Science Foundation of China Grants No. 41575018 and No. 41530423, and the 1000-Young Talent program of China.

References

Barbosa, S. M. (2011), Testing for Deterministic Trends in Global Sea Surface Temperature, *J. Climate*, *24*, 2516–2522, doi:10.1175/2010JCLI3877.1.

Cao, J. J., et al. (2007), Spatial and seasonal distributions of carbonaceous aerosols over China, *J. Geophys. Res.*, 112, D22S11, doi:10.1029/2006JD008205.

Che, H., X. Zhang, Y. Li, Z. Zhou, and J. J. Qu, (2007). Horizontal visibility trends in China 1981–2005. *Geophys. Res. Lett.*, *34*(24).

Chen, H. P., and H. J. Wang (2015), Haze days in North China and the associated atmospheric circulations based on daily visibility data from 1960 to 2012. *J. Geophys. Res. Atmos.*, 120(12), 5895-5909.

Collaud Coen, M., Andrews, E., Asmi, A., Baltensperger, U., Bukowiecki, N., Day, D., Fiebig, M., Fjaeraa, A. M., Flentje, H., Hyvärinen, A., Jefferson, A., Jennings, S. G., Kouvarakis, G., H. Lihavainen, C. Lund Myhre, W. C. Malm, N. Mihapopoulos, J.V. Molenar, C. O'Dowd, J. A. Ogren, B. A. Schichtel, P. Sheridan, A. Virkkula, E. Weingartner, R. Weller, P. and Laj, P. (2013), Aerosol decadal trends – Part 1: In-situ optical measurements at GAW and IMPROVE stations, *Atmos. Chem. Phys.*, 13, 869-894, doi:10.5194/acp-13-869-2013.

Deng, J., T. Wang, Z. Jiang, M. Xie, R. Zhang, X. Huang, and J. Zhu (2011). Characterization of visibility and its affecting factors over Nanjing, China. *Atmos. Res.*, *101*(3), 681-691.

Deng, X., X. Tie, D. Wu, X. Zhou, X. Bi, H. Tan, F. Li and C. Jiang (2008). Long-term trend of visibility and its characterizations in the Pearl River Delta (PRD) region, China. *Atmos. Environ.*, *42*(7), 1424-1435.

Donner, R. V., R. Ehrcke, S. M. Barbosa, J. Wagner, J. F. Donges, and J. Kurths (2012), Spatial patterns of linear and nonparametric long-term trends in Baltic sea-level variability, *Nonlin. Processes Geophys.*, *19*, 95–111, doi:10.5194/npg-19-95-2012.

Fu, C., Wu, J., Gao, Y., Zhao, D., and Han, Z. (2013). Consecutive extreme visibility events in China during 1960–2009. *Atmos. Environ.*, *68*, 1-7.

Franzke, C. (2013), A novel method to test for significant trends in extreme values in serially dependent time series, *Geophys. Res. Lett.*, 40, 1391–1395, doi:10.1002/grl.50301.

Guo, J. P., X. Y. Zhang, Y. R. Wu, Y. Zhaxi, H. Z. Che, B. La, W. Wang and X. W. Li, (2011). Spatio-temporal variation trends of satellite-based aerosol optical depth in China during 1980–2008. *Atmos. Environ.*, *45*(37), 6802-6811.

Hannachi, A. (2006), Quantifying changes and their uncertainty in probability distributions of climate variables using robust statistics, *Clim. Dyn.*, *27*, 301–317, doi:10.1007/s00382-006-0132-X.

Husar, R. B., and J. M. Holloway (1984), The properties and climate of atmospheric haze, in Hygroscopic Aerosols, edited by L. H. Ruhnke and A. Deepak, pp. 129–170, *Deepak Publ.,* Hampton, Va.

Husar, R. B., J. D. Husar, and L. Martin (2000). Distribution of continental surface aerosol extinction based on visual range data. *Atmos. Environ.*, *34*(29), 5067-5078.

Jinhuan, Q., and Liquan, Y. (2000). Variation characteristics of atmospheric aerosol optical depths and visibility in North China during 1980–1994. *Atmos. Environ.*, *34*(4), 603-609.

Kendall, M. G. (1975), Rank Correlation Methods, *Griffin*, London.

Koenker, R., and K. F. Hallock (2001), Quantile regression, *J. Economic Prespectives*, *15*, 143–156.

Koschmieder, H. (1926). Theorie der horizontalen Sichtweite. *Beitsaege Physik zur Atmosphere* 12, 33-55.

Li, J., C. Li, C. Zhao, and T. Su (2016), Changes in surface aerosol extinction trends over China during 1980–2013 inferred from quality-controlled visibility data, *Geophys. Res. Lett.*, 43, 8713–8719, doi:10.1002/2016GL070201.

Lin, J.-T., and J. Li (2016), Spatio-temporal variability of aerosols over East China inferred by merged visibility-GEOS-Chem aerosol optical depth, *Atmos. Environ.*, 132, 111-122, doi:doi:10.1016/j.atmosenv.2016.02.037.

Mann H. B. (1945). Nonparametric tests against trend. *Econometrica* **13**: 245–259.

O'Neill, N. T., A. Ignatov,B. N. Holben, and T. F. Eck (2000). The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data. *Geophys. Res. Lett*, *27*(20), 3333-3336.

Schreiber, T., and A. Schmitz (1996), Improved surrogate data for nonlinearity tests, *Phys. Rev. Lett.*, 77, 635–638.

Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's tau, *J. Am., Stat. Assoc.*, 63, 1379–1389.

Streets, D. G., C. Yu, Y. Wu, M. Chin, Z. Zhao, T. Hayasaka and G. Shi (2008). Aerosol trends over China, 1980–2000. *Atmos. Res.*, *88*(2), 174-182.

Sullivan, R. C., R. C. Levy and S. C. Pryor (2015). Spatiotemporal coherence of mean and extreme aerosol particle events over eastern North America as observed from satellite. *Atmos. Environ.*, *112*, 126-135.

Tao, M., L. Chen, X. Xiong, M. Zhang, P. Ma, J. Tao and Z. Wang (2014). Formation process of the widespread extreme haze pollution over northern China in January 2013: Implications for regional air quality and climate. *Atmos. Environ.*, *98*, 417-425.

Wang, G., K. Kawamura, X. Zhao, Q. Li, Z. Dai and H. Niu (2007). Identification, abundance and seasonal variation of anthropogenic organic aerosols from a mega-city in China. *Atmos. Environ.*, *41*(2), 407-416.

Wang, Q. Y., R.-J. Huang, J. J. Cao, X. X. Tie, H. Y. Ni, Y. Q. Zhou, Y. M. Han, T. F. Hu, C. S. Zhu, T. Feng, N. Li, and J. D. Li (2015), Black carbon aerosol in winter northeastern Qinghai–Tibetan Plateau, China: the source, mixing state and optical property, *Atmos. Chem. Phys.*, 15, 13059-13069, doi:10.5194/acp-15-13059-2015.

Wang, H. J. and H. P. Chen (2016), Understanding the recent trend of haze pollution in eastern China: roles of climate change. *Atmos. Chem. Phys.*, 16, 4205-4211.

Wu, J., C. Fu, L. Zhang and J. Tang, J. (2012). Trends of visibility on sunny days in China in the recent 50 years. *Atmos. Environ.*, *55*, 339-346.

Xia, X. (2011). Variability of aerosol optical depth and Angstrom wavelength exponent derived from AERONET observations in recent decades. *Environ. Res. Lett.*, *6*(4), 044011.

Ye, X., Y. Song, X. Cai and H. Zhang (2016). Study on the synoptic flow patterns and boundary layer process of the severe haze events over the North China Plain in January 2013. *Atmos. Environ.*, *124*, 129-145.

Yoon, J., W. von Hoyningen-Huene, M. Vountas and J. P. Burrows (2011), Analysis of linear long-term trend of aerosol optical thickness derived from SeaWiFS using BAER over Europe and South China, *Atmos. Chem. Phys.*, 11, 12149-12167, doi:10.5194/acp-11-12149-2011, 2011.

Yoon, J., A. Pozzer, D. Y. Chang, J. Lelieveld, J. Kim, M. Kim, Y. G. Lee, J.-H. Koo and K. J. Moon (2016). Trend estimates of AERONET-observed and model-simulated AOTs between 1993 and 2013. *Atmos. Environ.*, *125*, 33-47.

Zhang, X., L. Wang, W. Wang, D. Cao, X. Wang and D. Ye (2015). Long-term trend and spatiotemporal variations of haze over China by satellite observations from 1979 to 2013. *Atmos. Environ.*, *119*, 362-373.

Zheng, G. J., F. K. Duan, H. Su, Y. L. Ma, Y. Cheng, B. Zheng, Q. Zhang, T. Huang, T. Kimoto, D. Chang, U. Pöschl, Y. F. Cheng and K. B. He (2015), Exploring the severe winter haze in Beijing: the impact of synoptic weather, regional transport and heterogeneous reactions, *Atmos. Chem. Phys.*, 15, 2969-2983, doi:10.5194/acp-15-2969-2015.

Table 2. Regional extreme and median trends

Region	1980-1990						1991-2000						2001-2013					
	SL* 95th percentile	SL Median	Difference	QR* 95th percentile	QR Median	Difference	SL* 95th percentile	SL Median	Difference	QR* 95th percentile	QR Median	Difference	SL* 95th percentile	SL Median	Difference	QR* 95th percentile	QR Median	Difference
NE China	.97	.73	.24	.87	.58	.29	-.26	-.27	.02	-.24	-.31	.08	-.16	-.13	-.03	-.16	-.22	.06
NCP	.67	.70	-.03	.59	.71	-.11	-.14	-.02	-.12	-.13	-.17	.04	.15	.27	-.12	.16	.28	-.12
NW China	.88	.79	.11	.91	.87	.04	.12	-.19	.32	-.01	-.22	.21	-.15	-.23	.08	-.09	-.24	.15
SW China	.55	.15	.40	.46	.13	.33	.04	-.03	.07	.00	-.05	.05	.19	.07	.12	.16	.07	.09
YRD	.68	.32	.36	.59	.33	.26	.05	-.11	.16	.08	-.02	.10	.13	.32	-.19	.14	.33	-.19
PRD	.76	.17	.59	.66	.18	.48	.16	.12	.05	.13	.07	.06	.43	.54	-.11	.40	.53	-.13

* SL refers to Sen's slope and QR refers to quantile regression

Formatted: Section start: New page, Width: 21.59 cm, Height: 27.94 cm

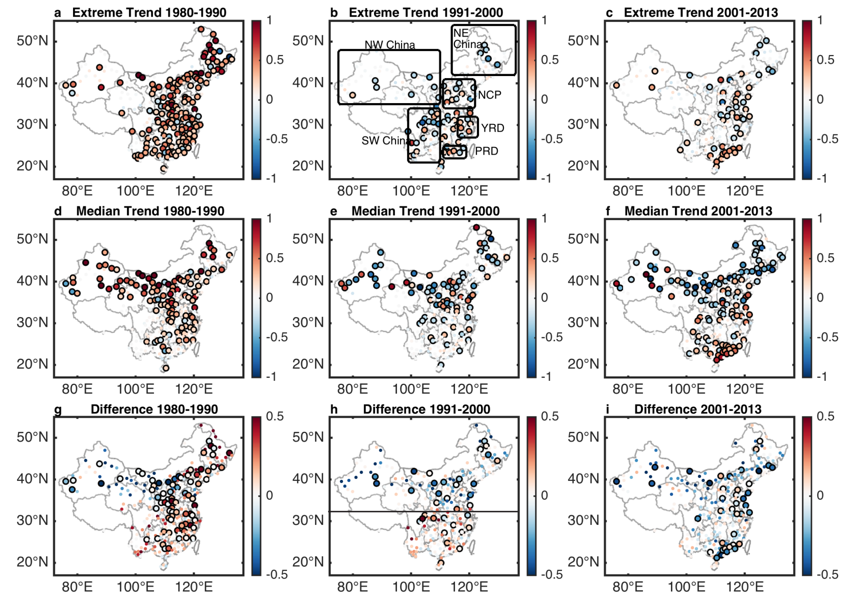


Figure 1. The first row: extreme trends estimate using quantile regression for the three decades, 1980-1990 (a), 1991-2000 (b), 2001-2013 (c); The second row: median trends estimated using quantile regression for the three decades; Bottom row: the difference between the values of the extreme trends and median trends, calculated as the extreme minus median. All trends are unitless and expressed as relative changes.

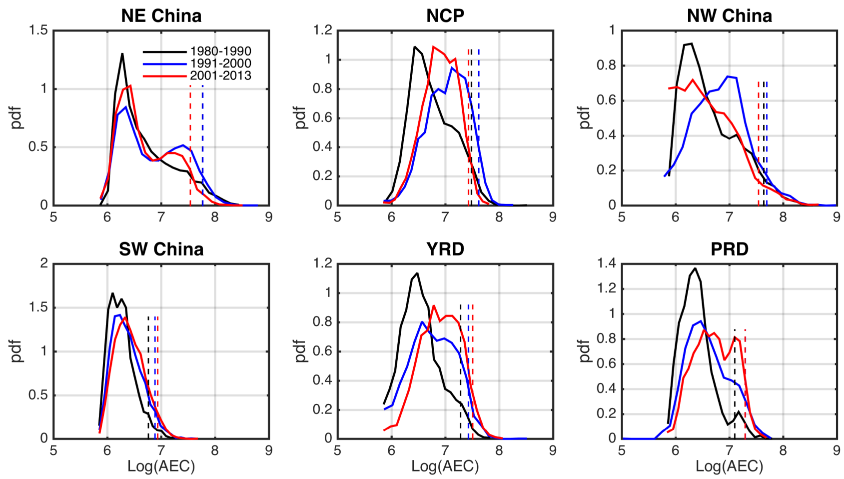
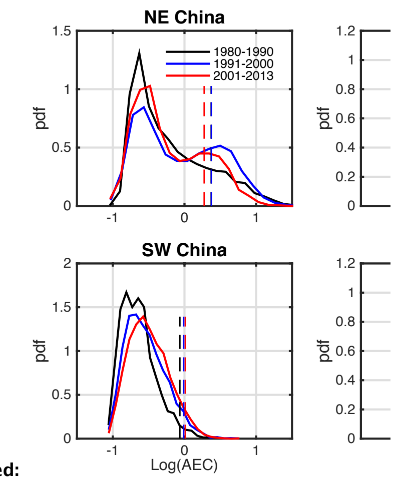


Figure 2. Probability distribution function (pdf) of AEC (megameter^{-1}) for the three decades over the six representative regions marked on panel b of Figure 1. The AEC has been converted to logarithmic scale.

Deleted:



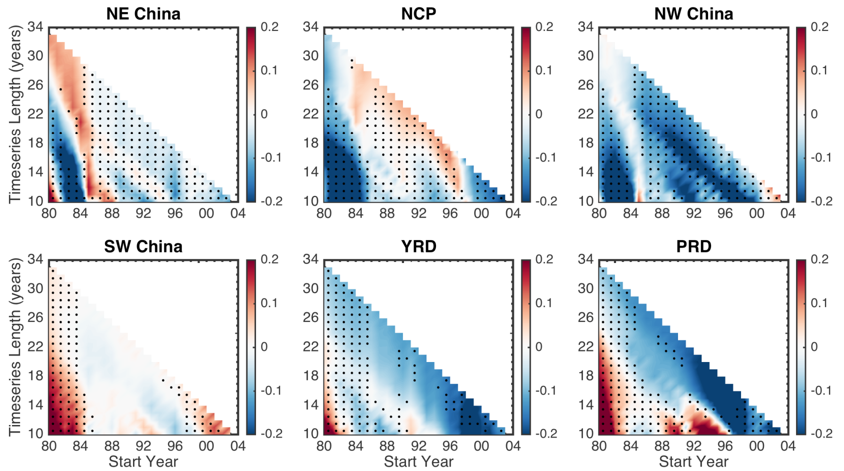


Figure 3. Difference between extreme and median trends calculated using quantile regression for the six representative regions marked on panel b of Figure 1. Trends are between each year from 1980 to 2004 and the end of the record, with 10 minimum. The x axis indicates the staring year, and the y axis indicates the length of the time series to calculate the trend.

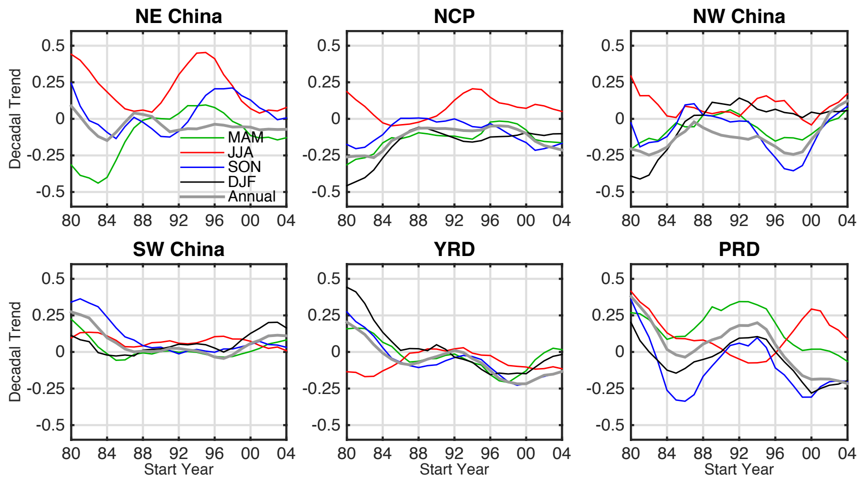


Figure 4. Seasonal time series of the difference between the extreme and median trends. The trends are calculated for each 10 year period starting form 1980 to 2004 (x axis), i.e., the first point is the trend difference for the 1980 to 1989 period, the second from 1982 to 1990, etc.