

Review of Li et al., “Evaluation and application of multi-decadal visibility data for trend analysis of atmospheric haze”

This manuscript represents an admirable effort to reduce, analyze, and interpret a global visibility data set in the context of changes in global air quality and haze. Specifically, human-observed visibility data from thousands of stations were combined with other measured aerosol parameters and SO<sub>2</sub> emission data to assess trends in global haze since the mid-1990s over the United States and since the early 1970s in Europe and Asia. Trends in visibility generally were consistent and validated with trends in SO<sub>2</sub> emissions and other ground-based aerosol data sets, such as reconstructed aerosol extinction and PM<sub>2.5</sub> mass. Human-observed visibility data are particularly challenging for trend analysis due to the many possible associated biases. The authors approached these challenges in a careful, objective, and consistent manner. The paper was well-written and the methods were generally well-described. The analysis and results are important contributions to the understanding of global haze trends in the context of changing emissions, especially with respect to the response of haze to regulatory and economic activity on a global scale. A weakness of the data set is its inability to resolve clean conditions, for example, across the western United States. This weakness biases the results to regions of higher haze degradation. The shortness of the trend periods is also a weakness; longer periods would provide more robust trends, as the authors indicated. It is not clear whether analysis of the shorter time periods were always necessary, especially over Asia. Overall, I recommend the paper be published after addressing specific comments provided below.

*RESPONSE: Thanks very much for affirming the value of this work. We do hope that both the value and limitation of visibility (Vis) data could be revealed and recognized through the analysis and presentation. As is mentioned several times in the paper, we discovered that the adopted Vis data cannot adequately represent variation in air quality over relatively clean regions due to threshold and discreteness issues. This is why the remaining stations after screening and the analysis in the paper are biased to the regions with high haze level.*

*We agree that long-term trends would be more reliable if no trend reversal is contained. This paper separately calculates trends for several periods of 8-10 years to allow possible trend reversal and to include stations with short-term data. We do agree that for some regions and periods, e.g. over Northern and Southern China, the trends over 1981-1996 could be merged due to similar station coverage and similar magnitude of the trends. We have tested the long-term trends for similar scenarios, and found generally consistent results with the separated short-term trends. These results are presented in the revised text. Specifically, the revisions are:*

*Page 12, Line 12-14: “Over these regions, we also calculate several time series and trends for longer merged periods with consistent station coverage and similar trends, to assess the consistency of the short-term trends.”*

*Page 16, Line 10-12: “Over 1954-1973, the long-term trend of 1/Vis is 1.2% yr<sup>-1</sup> ( $p < 0.001$ ), lying between the separated short-term trends.”*

*Page 17, Line 26-28: “Long term 1/Vis trend over Western Europe for 1981-2011 (insufficient qualified stations after 2011) is -1.8% yr<sup>-1</sup> ( $p < 0.001$ ), consistent with the separate short-term trends.*

*Page 19, Line 22-24: “The long-term trend over 1981-1996 for Northern China (0.5% yr<sup>-1</sup>,  $p < 0.001$ ) also exceeds that for Southern China (0.2% yr<sup>-1</sup>,  $p = 0.04$ ).”*

*Thank you for the comments and suggestions below which are very useful to improve the quality of the manuscript. We have taken them into full consideration during the revision. Responses to these comments are provided below. All the page and line numbers refer to the revised manuscript.*

**Specific comments:**

Page 3, line 16-26: This paragraph is a bit disjointed. It switches from variability in haze trends due to changing emissions to inherent uncertainties in Vis.

*RESPONSE: Agreed. This paragraph is to outline two possible limitations in previous studies. One is the usually long period employed that might miss reversals in trends, and the other is the inherent uncertainties in Vis data that were not well resolved. We have inserted a sentence at the front of this paragraph (Page 3, 16-17) to initiate the discussion*

*“Despite the abundance of the above mentioned studies, the interpretation of Vis data and its trends might be limited by insufficient data processing or poor data quality.”*

Page 3, line 27: Are the data used in this paper the same as those reported by Wang et al. (2012), and if so, is the inherent discrepancy resolved?

*RESPONSE: As mentioned in Section 3.1, the Vis data used in this study is the raw observations (with various report frequencies) from the ISD archive, while the data adopted by Wang et al. (2012) is from the “Global Summary of the Day (GSOD)” data which are daily means of the raw hourly data. The limitation of arithmetic means is discussed in detail in Section 3.1.2 (See Page 6, Line 18–24). Besides adopting nonparametric monthly statistics from the raw data, we also introduce the threshold filtering and change point separation which is different from Wang et al. (2012). All these efforts are made to resolve and reduce the inherent data uncertainties, and in the results of this study, more consistency with other in situ data and historical SO<sub>2</sub> emission trends validate these processing efforts. The contrasting 1/Vis trends over the US in Wang et al. (2012) is due to the detection limit and improved air quality, which makes Vis data incapable of capturing the variation and trends in air quality over the US after mid 1990s, as discussed in Section 4.1.*

*We added several lines in Page 14, Line 25-28 to clarify this:*

*“Thus the apparent discrepancy in sign of trends in 1/Vis (Wang et al., 2012) with trends in other aerosol measurements (Attwood et al., 2014; Hand et al., 2012a; Hand et al., 2014; US EPA, 2012) is resolved by more comprehensive data processing and screening.”*

Page 5, line 19: Consider removing Figure 1 for space. The information is conveyed in the text, and the number of remaining stations after each step can be included in the text.

*RESPONSE: Good suggestion. We have removed Figure 1, and added several sentences in the text to describe the remaining stations:*

*Page 5, Line 30: “After this screening step, 21,703 stations remain from the 30,895 original ISD sites.”*

*Page 6, Line 15: “This data screening step further reduces the number of qualified station to 10,446.”*

*Page 7, Line 5: “A total of 6,466 stations comply with these standards and remain in the data archive.”*

*Page 8, Line 15-16: “A total of 3930 stations (5320 time series) remain after this processing step...”*

Page 5, line 24: Do each of these stations have collocated RH measurements?

*RESPONSE: RH data are not directly available in the ISD data files. We estimate RH from air temperature and dew point. As added in Page 5, Line 24: “estimated from temperature and dew point”.*

Page 7, line 22: Clarification: Were change point detections required to occur at the same time in both the 50th and 75th percentiles to determine an actual change point? Is a minimum amount of data required after a detection point to determine if it was an actual change (e.g., 201304 50th percentile detection)?

*RESPONSE: The actual change points are manually selected from the candidate change points reported by the RHtest software. We include time series of monthly 50<sup>th</sup> and 75<sup>th</sup> percentiles of 1/Vis to provide possible candidates for decision. It is not exactly required that change points always occur at the same time in the two time series, but the selection is rather subjective to report visually obvious structural breaks in any of the two time series. The 201304 in 50th percentiles time series was not selected as a change point, because it is not judged as visually significant structural breaks. We added several sentences (Page 7, Line 22-24) to modify the description: “By visually inspecting each remaining station from Section 3.1.3, we retain only obvious structural discontinuities in the time series of 50th or 75th monthly percentiles from the candidate change points provided by the RHtest results.”*

Page 9, line 6: Which IMPROVE algorithm was applied? Malm et. al (1994) or Pitchford et al., (2007)?

*RESPONSE: The  $b_{ext}$  data from the revised algorithm in Pitchford et al., (2007) was used. We added the reference of Pitchford et al. (2007) in Page 9, Line 18.*

Page 9, line 10: Was site-specific Rayleigh used?

*RESPONSE: Yes. We revise this sentence as (Page 9, Line 20) “including aerosol extinction and site-specific Rayleigh scattering”.*

Page 9, line 14: How many IMPROVE sites were used and over what time periods?

*RESPONSE: We added “for 56 IMPROVE sites over 1993-2013” to the caption of Figure A2.*

Page 9, line 21: Do these SO<sub>2</sub> emission inventories include all sources?

*RESPONSE: Yes. This sentence is revised as “We apply bottom-up total anthropogenic SO<sub>2</sub> emission inventories to interpret historical 1/Vis trends.” in Page 10, Line 2.*

Page 10, line 21: Does ‘multi-year’ correspond only to only the time period of the trend?

*RESPONSE: Yes. This sentence is revised as “All monthly data are deseasonalized by removing multi-year monthly means of each period before trend estimation.” in Page 11, Line 3-4.*

Page 10, line 30: Should “MK-Sen trends” actually be “MK-Sen slopes”?

*RESPONSE: Yes. Revised.*

Page 10, line 30: Clarification: The monthly 1/vis values are 75th percentile (pg 6, line 28), but they are normalized by the monthly mean? Or a mean of the monthly 75th percentile values over the trend time period?

*RESPONSE: Yes, the relative trends are calculated by normalizing the absolute slopes to the mean of monthly 75 percentile values of 1/Vis over this period. Since as indicated in Page 6, Line 28-29, these 75th percentiles of 1/Vis are referred as “monthly 1/Vis” throughout the text, we have revised this sentence as “Relative trends are calculated by normalizing the absolute MK-Sen slopes to the multi-year mean of*

monthly 1/Vis in the corresponding period to facilitate the comparison and interpretation with other in situ data.” in Page 11, Line 13-14.

Page 11, line 9-10: Does this mean that the data were actually interpolated? If so, using what technique?

*RESPONSE: The data was only gridded to avoid biased averaging, while blank grids were not interpolated. That is why the grid number (i.e. red dots in Figure 8, 10 and 12) of each monthly 1/Vis is different. The gridding is done by averaging all values located into one grid of 1 degree resolution.*

Page 11, line 12: Could the authors clarify how the maximum number of grids is determined?

*RESPONSE: Each monthly 1/Vis corresponds to several grids for averaging, and this “maximum number of grids” is the number of unique grids covered by all these monthly data for each time period.*

*This sentence (Page 12, Line 7) was revised as “only monthly data derived from at least 75% of the total grids (i.e. number of unique grids covered by all the monthly data) for each study period are used in the composite trend estimation”.*

Page 11, line 20: Are the comparisons between monthly mean best and 75th percentile 1/vis?

*RESPONSE: Yes. Section 3.1.2 and 3.2 specified what the monthly statistics are.*

Page 11, line 24: Over what years?

*RESPONSE: We have revised this sentence as (Page 12, Line 19-20) “Collocations are considered between IMPROVE and ISD time series over 1988-2013...”*

Page 11, line 25: The offset in the comparison shown in Figure A2 is likely due to Rayleigh scattering.

*RESPONSE: Yes. We added “The intercept of  $\sim 12 \text{ Mm}^{-1}$  corresponds to Rayleigh scattering” to the caption of Figure A2.*

Page 13, lines 21- Page 14, line 23. Consider reordering this section and figures to streamline the discussion: Discuss/present Figure 5, followed by Figure 7, with the discussion in the text starting with the initial paragraph (pg 13, line 14-20), followed by (pg 13, line 30- pg 14, line 7), followed by (pg 14, line 8-10), then (pg 13, line 21-29), followed by (pg 14, line 11-23).

*RESPONSE: Thanks for this good suggestion. We revised the order of discussion, and the order of the three figures as well.*

Page 14, line 13: The opposite seasonality between Figure 6a and 6b could be associated with urban/rural differences, assuming most of the European sites are in cities?

*RESPONSE: The seasonality of aerosols over the US was well documented, while to our best knowledge there is no existing studies that investigate in detail about aerosol seasonal variation over Europe. According to Hand et al. (2012), the seasonality of sulfate and nitrate aerosols are similar over urban and rural areas (summer maxima for sulfate and winter maxima for nitrate), while organic matter shows a winter maxima in urban areas over the Western US, opposite to rural sites. Most urban sites over the Eastern US still show a weak summer maxima in total  $\text{PM}_{2.5}$ . According to Tørseth et al. (2012), the EMEP sites are located in rural and background areas to minimize local influences and to represent a larger region, similar to the IMPROVE design. Therefore it is uncertain whether the urban/rural difference could explain the different seasonality in Figure 6. That is why we used the term “could be attributed to” in the text to indicate this unresolved issue, which is out of the scope of this paper.*

Hand, J., Schichtel, B., Pitchford, M., Malm, W., and Frank, N.: Seasonal composition of remote and urban fine particulate matter in the United States, *J. Geophys. Res.*, 117, D05209, doi:10.1029/2011JD017122, 2012.

Tørseth, K., Aas, W., Breivik, K., Fjæraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg, S., and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972 – 2009, *Atmos. Chem. Phys.*, 12, 5447-5481, doi:10.5194/acp-12-5447-2012, 2012.

Page 15, line 4: It would help to remind the reader here that the 89-96 map was shown in Figure 5.

*RESPONSE: This is followed by “(Fig. 4 contains 1/Vis trends over 1989-2013)” in the revised manuscript (Page 16, Line 4).*

Page 15, line 7: Point out that the range in trends represents more than one time period.

*RESPONSE: We added “during the following 2 periods” in the revised manuscript (Page 16, Line 8-9).*

Page 15, line 25: SO<sub>2</sub> emissions in the western US are also much lower than in the East.

*RESPONSE: This sentence (Page 16, Line 25-26) is revised as “where SO<sub>2</sub> emissions are much lower than in the Eastern US, organic aerosols dominate in PM<sub>2.5</sub> and forest fires are more prevalent”. Thanks for helping complete the reasoning.*

Page 18, line 19: Are SO<sub>2</sub> trends calculated over the same regions as 1/vis?

*RESPONSE: We present annual time series of SO<sub>2</sub> emission over each region in this figure. As stated in the caption of Figure 12, “The SO<sub>2</sub> emission in Lu et al. (2011) in orange and the EDGAR SO<sub>2</sub> emission in purple are summed from all pixels inside the defined region.”*

Page 18, line 26: While this section/paragraph is interesting and useful to include, I am not sure it warrants another section. Can it be blended into the discussion?

*RESPONSE: Since Reviewer #2 suggested comparing AOD trends, we have extended this section as “Connections to AOD and SSR trends” in the revised manuscript (Page 20, Line 5-31).*

Page 18, line 32: Can the authors provide some quantification of “agrees well”?

*RESPONSE: We revised this sentence (Page 20, Line 13-15) as “Despite these uncertainties, the observed reversals of SSR from “dimming” to “brightening” in 1980-1990 over the US and Europe (Streets et al., 2006; Turnock et al., 2015; Wild, 2012) generally agree with the reversals around the 1980s of 1/Vis trends in this study.”*

Page 20, line 27: Other species could also be contributing to these differences.

*RESPONSE: Yes. As also part of the concern of Reviewer #2, we have added several discussion on other emissions (Page 22, Line 19-26) to complete this discussion in the revised paper.*

*“Notable reductions in emissions of nitrogen oxides and black carbon have been reported over North America and Western Europe (Bond et al., 2007; Lu et al., 2015; US EPA, 2012; Vestreng et al., 2009), while steady increase in emissions of nitrogen oxides, organic carbon and black carbon were identified over China (Lu et al., 2011; Zhao et al., 2013). Observed (Lebensperger et al., 2012; Murphy et al., 2011) and simulated (Lin et al., 2010; Wang et al., 2013) changes in various aerosol chemical species suggest increasing importance of emissions other than SO<sub>2</sub> on air quality trends in recent years.”*

Page 21, line 11: I believe IMPROVE has an acknowledgement statement recommendation.

*RESPONSE: We corrected this acknowledgement. We now cite and acknowledge all data sources following available recommendations in the revised manuscript.*

**Technical corrections:**

Page 1, line 21: Define bext at first usage (see line 23).

*RESPONSE: Revised.*

Page 1, line 27: Define EDGAR

*RESPONSE: Revised. We also define EDGAR at first usage in the main text (Page 10, Line 9).*

Page 2, line 3: What is meant by “inferred” in this context?

*RESPONSE: It simply means “calculated”. We deleted this term in the revised version.*

Page 2, line 4: What is meant by “reconstructed” 1/vis?

*RESPONSE: This term is used initially to emphasize the processing efforts introduced in Section 3.1. We revised this term to “quality assured” which is better at expressing this purpose.*

Page 5, line 12: State that these are global observations and provide years.

*RESPONSE: This sentence is revised as “We begin with raw Vis data from global synoptic observations over 1929-2013” in Page 5, Line 13.*

Page 7, line 16: Please provide manufacturer information for the RHtest software.

*RESPONSE: The algorithm and code is maintained by Dr. Xiaolan Wang and Dr. Yang Feng at Climate Research Division, Environment Canada as open access resources. We have specified the original website of the code in the text (Page 7, Line 19-20), and acknowledged the PIs in the acknowledgement.*

Page 17, line 28: Define EANET

*RESPONSE: Revised.*

Page 17, line 32: Define MODIS, AOD

*RESPONSE: Revised.*

**Comments on Figures:**

Figure 1: This figure is not necessary (see earlier comment).

*RESPONSE: Deleted.*

Figure 2: What do the “0”s represent on the figure (under the dates)? Would these be dates for ‘undetermined change’ and ‘>95 insig change’ if there were any?

*RESPONSE: Yes. We noticed that too much text in the Figure makes the caption and interpretation cumbersome. Since the 3 different statistical types have already been introduced in the main text, we deleted the legends of the other two kinds of change points, to make the Figure clearer for readers. Accordingly, the caption of the new Figure 1 has been revised as “An example of change point detection and determination based on the time series of 50th (red) and 75th (black) percentiles of monthly 1/Vis.*

*Automatically detected change points are represented by vertical lines. Text in the inset lists the dates of automatically detected points. In this example, 5 significant change points are identified, in which February 1988 is determined as the separation point for further analysis, while other reported breaks are considered as false detections.”*

Figure 4: Consider rearranging these figures to follow the discussion, with the mean text on the top left, followed by the figures in order of the discussion. Include the time period in the caption.

*RESPONSE: Revised in both the figure and text.*

Figure 5: The open-circle symbols are very difficult to see. Perhaps thicken the symbol line?

*RESPONSE: The open-circles correspond to insignificant trends, which are often in small magnitudes and light (blue/red) colors in the figure, and thickening the lines cannot help highlight them significantly. We have changed the symbols in all the trend maps to resolve this issue. In the revised manuscript, larger colored points with black outline indicate trends with at least 95% significance, smaller colored points with black outline represent trends with 90%-95% significance, and colored points without outline indicate insignificant trends. We believe now the symbols are more distinguished.*

Figure 7: Same as figure 5.

*RESPONSE: Revised as previously stated.*

Figure 8: Same as figure 5.

*RESPONSE: Revised as previously stated.*

Figure 9: “Average monthly” still refers to the 75th percentile? Adding symbols to the black line might help to show data gaps.

*RESPONSE: Yes, the “monthly 1/Vis” refers to the nonparametric 75th percentiles throughout the text. We have added black dots in addition to the black lines in Figure 8, 10 and 12.*

Figure 10: Adding the red boxes to all of the plots would help the eye to discern the divisions in the other time periods, especially with the high site density. Insignificant symbols are hard to see.

*RESPONSE: Revised. We added red boxes to show the defined regions in Figure 9 and 11.*

Figure 12: Same as Figure 10.

*RESPONSE: Revised as previously stated.*

Figure A1: Typo on line 3 “visibility”

*RESPONSE: Revised.*

Figure A4: Insignificant symbols are hard to see.

*RESPONSE: Revised as previously stated.*