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 SO2 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 SO2 emissions and other ground-based aerosol data sets, such as reconstructed aerosol extinction and PM2.5 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^{-1} ($p < 0.001$), 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^{-1} ($p < 0.001$), 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^{-1}, $p<0.001$) also exceeds that for Southern China (0.2% yr^{-1}, $p=0.04$)."

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. Reponses 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_2 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 50th and 75th 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 SO2 emission inventories include all sources?

RESPONSE: Yes. This sentence is revised as "We apply bottom-up total anthropogenic SO_2 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 bext 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 ~12 Mm^{-1} corresponds to Reyleigh 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, line7), 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 $PM_{2.5}$. 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: SO2 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_2 emissions are much lower than in the Eastern US, organic aerosols dominate in $PM_{2.5}$ and forest fires are more prevalent". Thanks for helping complete the reasoning.

Page 18, line 19: Are SO2 trends calculated over the same regions as 1/vis?

RESPONSE: We present annual time series of SO_2 emission over each region in this figure. As stated in the caption of Figure 12, "The SO_2 emission in Lu et al. (2011) in orange and the EDGAR SO_2 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, Line13-15) 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 (Leibensperger 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_2 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 bext 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.

The paper by Li et al. analyzes the multi-decadal 1/Vis changes over the US, Europe and Eastern Asia and their correlations with SO2 emissions. Building on previous works, the paper makes important efforts to ensure data consistency over the long term, which leads to important findings on the 1/vis trends that are largely consistent with historical SO2 emissions. The paper is well written and could be published with minor revisions.

RESPONSE: Thanks very much for the positive review of the manuscript. The comments and suggestions below improve the quality of the manuscript. We have taken them into full consideration during the revision. Reponses to these comments are provided below. All the page and line numbers refer to the revised manuscript.

It would be interesting to compare the 1/Vis trends with satellite AOD data. For example, is the 1/Vis trend in the recent decade consistent with MODIS AOD trend? The comparison would be interesting especially for regions with fewer ground-based aerosol (optical) measurements.

RESPONSE: Good suggestion. We have added a summary of discussion of AOD trends in Section 5.4 (Page 20, Line 20-31):

"Reliable AOD data over land are limited to the recent two decades, but exhibit even greater consistency with 1/Vis trends. The recent decrease in 1/Vis after late-1990s over the US and Western Europe in this study is consistent with previous studies on AOD trends based on both ground based (e.g. Li et al., 2014; Yoon et al., 2012) and satellite (e.g. Chin et al., 2014; Hsu et al., 2012; Pozzer et al., 2015) observations. Over China, several studies on AOD trends in the 2000s showed notable increasing tendency (e.g. Hsu et al., 2012; Pozzer et al., 2015; Yoon et al., 2012), while some recent studies also discovered that separating AOD time series could reflect the plateauing and reversal of trends in recent years due to emission control strategies (Che et al., 2015; He et al., 2016; Lu et al., 2011). $PM_{2.5}$ trends derived from satellite AOD over 1998-2012 have decreasing tendencies over North America and Europe, and increasing tendencies over Eastern Asia (Boys et al., 2014; Van Donkelaar et al., 2015), similar to the 1/Vis trends found here."

In the abstract, please clarify 'change point detection'.

RESPONSE: We have expressed this term as in the following text "to identify and separate methodological discontinuities such as the introduction of instrumentation".

The nighttime and daytime visibility measurement methods and meanings are very different. Does the monthly 1/vis calculation take into account this difference?

RESPONSE: Thanks very much to point out this issue. We made a sensitivity test by applying the same screening procedure to the daytime and nighttime data based on local time. The change points detected using the combined data were applied since these should be consistent in both daytime and nighttime datasets. Relative trends of 1/Vis over all remained stations and the 8 time periods were calculated and compared to the trends calculated in the paper (from the combined data) as presented in the figure below. It could be observed that after representing the data into a monthly resolution and normalizing the changes in 1/Vis into relative trends, the difference between daytime and nighttime data is negligible. In other words, the daytime and nighttime data show consistent trends in haze level with the combined data.

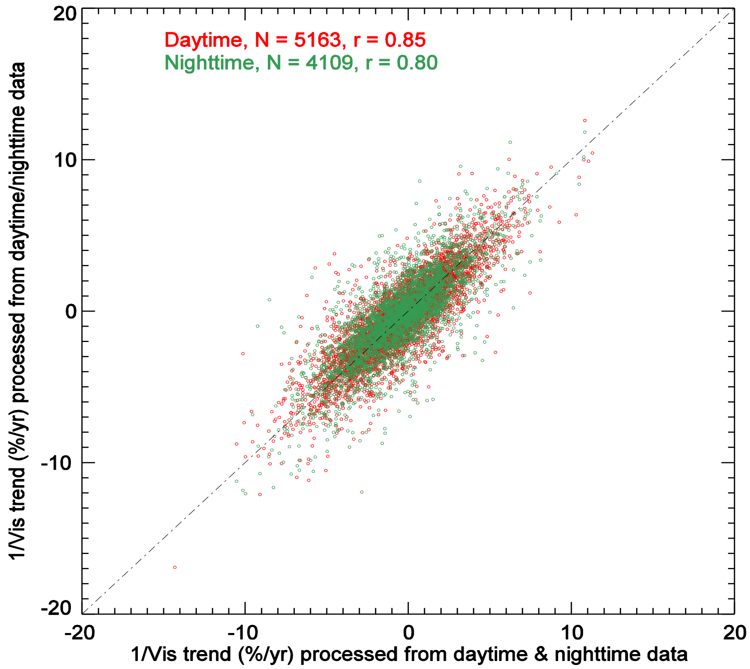

Figure 1. Scatter plot of relative trends calculated for the 8 periods using monthly 1/Vis derived from only daytime or nighttime data V.S. those calculated from monthly 1/Vis derived from both day time and nighttime data.

We added additional discussion in the text (Page 11, Line 21-32) to address this issue:

"The meaning and observing methods of daytime and nighttime data differ. According to WMO (2008), Vis at night, as determined using illuminated objects, also depends on the light source intensity, the adaptation of the observer's eyes to darkness and the observer's illuminance threshold. We compare the relative trends calculated using daytime and nighttime data to the combined trends adopted in this paper, over all remaining sites and the 8 periods. The 5183 daytime trends have a correlation of 0.85 with the combined trends, in which 84% of the differences between significant trends ($p < 0.1$) are within 50%. For the comparison between 4109 nighttime and combined trends, the correlation is 0.80 and 78% of the differences between significant trends are within 50%. Therefore, after representing the data into a monthly resolution and normalizing the changes in 1/Vis into relative trends, the daytime and nighttime data show generally consistent trends in haze level compared to the combined data, and do not meaningfully alter our results and conclusions."

Page 7, Line 20-27: How exactly to determine the changes manually?

RESPONSE: As discussed in the text, we figured the automatically detected change points are too fragmentary, thus decided to visually check the time series superimposed with the candidate change points (as in Figure 1), and only select those points that are obvious structural breaking points in the time series.

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 8, Line 11-13: Is it possible to test the importance of change point detection for the long-term analysis, since this is a major contribution of the present study?

RESPONSE: Thanks very much for this suggestion. We similarly tested our trend calculation on the data before change point detection and separation. Interestingly, we found negligible differences in the calculated regional time series and trends compared to the results in the paper, thus similar results and conclusions could be drawn without the change point detection and separation. After re-examining the data archive, we discovered that in the remained 3930 stations after all the screening, 856 stations (22%) are actually reported with change points and separated after the last step. This small fraction of stations with change points does not alter the results greatly. Therefore, we suggest that the whole data screening processes, including the threshold filtering, the requirements on the temporal and spatial representativeness, all contribute to the reliability of the results. But the discontinuities are still a major issue if studies were to be conducted on independent sites or small regions, and data after separation are favored. Significant differences are expected due to the structural discontinuities in the original time series.

We added in the text (Page 8, Line 15-21) to address this issue:

"A total of 3,930 stations (5,320 time series) remain after this processing step, in which 856 sites (22%) are diagnosed as containing change points and thus separated. This small fraction of structural discontinuities generally has minor impacts on the large-scale trend features and regional trends in Section 5 according to our sensitivity test using data without separation. But the separated data reduce spatial incoherency in the derived trend maps, and are more reliable for studies over small areas or independent stations, as shown in Fig. 1."

Page 17, last paragraph: the discussion on Chinese pollution transport is relatively subjective. Could you provide more analyses or references on how Chinese pollution transport would change/reverse the pollution trend in Korea, since this has important implications for regional pollution control and collaboration?

RESPONSE: We have added discussion of two articles about air quality trends in Korea (Page 18, Line 25-30):

"Lee et al. (2015) also discovered insignificant improvement of Vis over urban areas of Korea after late 1990s despite the national emission reduction policy launched in early 2000s, which was attribued to the regional transport from upwind continental areas. Long-term aerosol measurement over Gosan Island, Korea showed rapid increase of sulfate and nitrate concentrations from early 2000s to ~2006, which were closely related with the trends of China's emission (Kim et al., 2011)."

We also reemphasize the importance of regional collaboration (Page 19, Line 11-13). Thanks for this good point.

" This analysis highlights the sensitivity of 1/Vis to long range transport, and the value of international collaboration for air quality improvement over Eastern Asia."

Especially in recent decades, the importance of other aerosols or precursors has increased relative to SO_2, such as NO_x, NH_3, SOA, and BC. The paper mainly compares the 1/vis trend with SO_2 trend. It will be interesting to discuss if other species have affected the 1/vis trends in recent years. For example, Lin et al. (2010) and Zhao et al. (2009) both showed the offset of SO_2 reduction over China by rapidly increasing NO_x.

Lin, J.-T. et al. Recent Changes in Particulate Air Pollution over China Observed from Space and the Ground: Effectiveness of Emission Control. Environmental Science & Technology 44, 7771- 7776, (2010).

Zhao, Y. et al. Soil Acidification in China: Is Controlling SO2 Emissions Enough? Environmental Science & Technology 43, 8021-8026, (2009).

RESPONSE: Yes. As discussed in the last paragraph, "SO_2 emission inventories cannot fully explain the trends in ambient haze due to the influence of other emissions and meteorological factors." We extended this discussion by adding the following text in the last paragraph (Page 22, Line 19-26). Thanks for this recommendation.

"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 (Leibensperger 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_2 on air quality trends in recent years."