Response to Referee #2

We greatly appreciate the reviewer for proving the constructive comments, which have helped us improve the paper quality. We have revised the paper accordingly as detailed in our point-by-point responses below.

RC- Reviewer's Comments; AR – Authors' Responses

RC: General comments
This paper asserts that statistical trends analysis of the linkage between emissions changes and measured wet deposition is obscured by multiple factors including climate anomalies. The target analytes of wet deposition measurements (SO4-2, NO3-, NH4+) undergo complex atmospheric transformations from their emitted precursors and washout or entrainment in precipitation are dependent on the climate. The climate anomalies are not specifically identified, but evidence exists that they influence relationships between wet deposition and emission trends and are occurring more frequently. Thus, a need exists for a statistical analysis technique to reduce the impact of the climate anomalies and increase the time interval of comparisons as emission changes, especially those due to regulations, are phased in incrementally and are not linear. The authors propose a statistical method based on the development of an input dataset termed "climatology" (an average of 12 ranked 24-month wet deposition flux measurements) and trends analyses to produce regression slopes for each of the 12 individual 2-year periods considered and the averaged "climatology" dataset. The regressions used are based off the Mann-Kendall (M-K) method, linear regression (LR), and piecewise linear regression (PLR). The authors propose that the time-series of the derived regression slopes better reflects the trends in reported emissions of precursor gases, than the time-series of the annual wet deposition flux data itself.

The method is novel and them-value time-series relates better to emissions time-series than wet deposition flux (Fwet) time-series at Site 1 for SO4-2 and to a lesser degree for NO3. The m-value time-series appears to reflect inflection points in the emissions time-series that are not as easily observable in the wet deposition flux time-series. However, the method does not improve the relationship of m-values of NH4+ to NH3 emissions at site1. Furthermore, the method does not seem to show improved m-value correlation with emissions over the annual Fwet data any other location (at Sites 2, 3, and 4) or species. There is no direct comparison metrics to gauge the improvement of the m-values over the annual Fwet other than visual interpretation of plots. The m value time-series will obviously be visually "cleaner" since a) the m-value has outliers removed and b) the m-value represents 24 data points and the annual Fwet represents 12 data points.

AR: We have added the comparison at Sites 3 and 4 in the revised Supporting Information, which reads: "Using the m-values over the annual F_{wet} of SO_4^{2-} improves the r value from 0.73 to 0.87 at Site 3 and from 0.91 to 0.93 at Site 4. Using the m-values over annual F_{wet} of NO_3^- improves the r value from 0.81 to 0.87 at Site 3 and from 0.78 to 0.89 at Site 4.” No significant correlation of m-value and F_{wet} with the corresponding emissions existed at Site 2 and the comparison is thereby not presented.

RC: The largest problem with the study is that that technique is not demonstrated to be robust. The method hinges on the stability of the m-values, but they are very susceptible to the large-value outliers (e.g. example described in text for ‘90-‘91, causes a 0.2 change in m-value; shown in Fig 1). Moreover, for Site 1, the authors acknowledge that 8 of 12 (67%) of datasets needed to have an outlier removed, which from my interpretation greatly compromises the robustness of this technique and its applicability to different datasets.

AR: We originally only explained the method from a statistical analysis consideration, which may hinder the real advantage of the method. In the revised paper, we have made substantial revisions in several sections to clarify this point. For example, we have changed this sentence “The actual observed maximum value of 532 mg m^{-2} $month^{-1}$ was much larger than the upper range of the predicted value and was thus believed to be caused by monthly scale climate anomalies” to this: “The actual observed maximum value of 532 mg m^{-2} $month^{-1}$ was much larger than the upper range of the predicted value and was thus believed to be caused by monthly scale climate anomalies, i.e., the occurrence of extreme amount of precipitation. The maximum monthly deposition flux in 1990-1991 occurred in September 1990 when the monthly precipitation depth reached 294 mm, which was much higher than those in the same month of other years, e.g., 169, 68, 95 and 127 mm in 1988, 1989, 1991 and 1992, respectively. The maximum daily precipitation depth in September was also higher in 1990 (91 mm) than in other years (43.6, 12.2, 13.6 and 26.8 mm in 1988, 1989, 1991 and 1992, respectively). However, the monthly geometric average concentration of SO_4^{2-} in precipitation (1.8 mg L^{-1}) in September 1990 was close to the mean value (1.7±0.3 mg L^{-1}) in September 1988-1992 and was even smaller than that (2.9 mg L^{-1}) in August 1990.” There are several other similar changes which can be found from the track change version of the paper.

In our approach, only the maximum value in 24 months severely deviated from the general trend was removed to calculate m-values. Thus, we have 95%-100%, i.e., 23/24 -24/24 monthly values, data to calculate m-value with high R^2 values (e.g., 0.92-0.98 at Site 1 for SO_4^{2-}). The calculated m-value would fully reflect the contribution from emissions of air pollutants since only 5% data are sometimes removed. When the data size is even larger, e.g., the group of Sites 1, 3 and 4, 100% data are used to calculate m-value. Using m-values calculated from Approach C is applicable for different datasets.

Compared with the calculated m-values from Approach B using 12 month data, the use of 24 month data in each array in Approach C largely increased R^2 value and decreased uncertainties of the calculated m-values. However, a linear regression analysis result, assuming zero interception and using the m-values calculated from Approach B against the annual F_{wet} data, showed the value of R^2 as high as 0.99 (Fig. S4 added in the revision). This means that the trend analysis results would be the same regardless of using annual F_{wet} data or the m-values as input if Approach B is used. The extracted trends would include larger perturbations from climate anomalies in Approach B. Thus, it can be concluded that the trend analysis results derived from m-values calculated from Approach C would be more robust than those derived from m-values calculated from Approach B. It is also safe to say that the trend results derived from m-values calculated from Approach C are more robust than those derived from annual F_{wet} data.

From Comment 4 listed below, we realize that the original text may mislead the reviewer, i.e., 8/12 datasets needed to have an outlier removed because of the maximum F_{wet} being positively deviated from the general trend. This is of course impossible and may make the reviewer doubt the robustness of Approach C. We have therefore clarified this in the revised paper, which reads: "Three out of the 12 data sets showed the maximum F_{wet} being positively deviated from the general trend, five negatively deviated from the general trend, and four consistent with the general trend."

RC: There appears to be a m-value error analysis conducted with three different approaches in the Supplemental section (Approaches A through C), but no summary or assessment of m-value stability or uncertainty is given. This needs to be developed much more. Sources of uncertainty in the m-values are not discussed.

AR: We have added such analysis, and details can be found in section 1 of the revised Supporting Information.

RC: A reader will likely ask why are large Fwet values so frequently (the 8 of 12 datasets mentioned above) in disagreement with the rest of the monthly values? This question is not answered. What causes the large flux (other than climate anomaly?) Is this a high or low rain event month? Is this rain after a stagnation event?

AR: The large F_{wet} value was mainly caused by extreme precipitation depth in monthly scale. We have clarified the frequency of the large values (see the response to the comment above). The distribution result is quite normal. Although the maximum monthly value of F_{wet} positively deviated from the general trend was clearly attributed to extreme precipitation, the cause was yet to be identified for the maximum monthly value of F_{wet} negatively deviated from the general trend. This latter case has also been stated in the revised paper.

RC: There is too much assigning uncertainty to vague "Climateanomalies" and "interannual climate variability". These concepts are neither adequately defined nor is any impact that they might have on monthly wet deposition values identified. The section on "interannual climate variability" could be strengthened with local ambient concentrations which are possibly available.

AR: We have revised discussions where appropriate throughout the paper. For example, the secondary paragraph of Section 3.2.2 has been revised substantially, which now reads: “In addition to decadal anomalies of wind fields, the interannual climate variability such as precipitation depth, annual anomalies of wind fields in 2007, etc., (Fig. 3b) also affected the trends in m-values and annual F_{wet} of NO_3^-. The annual precipitation depth largely varied from 601 mm to 1054 mm in the two decades. The perturbations from interannual variability of precipitation depth cannot be completely removed by the new approach. For example, the calculated m-values in 1992-1993 and 1994-1995 were evidently lower than the m-values in 1990-2001. However, the annual geometric average concentrations of NO_3^- in 1992-1995 varied around 0.77±0.11 mg L^{-1} and were even larger than the values of 0.66±0.08 mg L^{-1} in 1990-2001 (excluding 1992-1995). The lower m-values were mainly attributed to the lower precipitation depth in 1992-1994 (Fig 3b) rather than lower emissions of NOx. Interannual climate variability including precipitation depth and annul anomalies of wind fields may complicate the relationship between the F_{wet} of NO_3^- and the emissions of NO_x in British Columbia.” Also in the Conclusion section, the revised version on this point reads: “At this location, the decreasing trends in F_{wet} of SO_4^{2-} and NO_3^- were caused by the decadal anomalies of wind fields, as well as being affected by interannual climate variability including precipitation depth and annul anomalies of wind fields, etc., which overwhelmed the impact of the emission changes of the gaseous precursors in this province. This is the first study that has identified that decadal anomalies of wind fields can dominate trends in F_{wet} of SO_4^{2-} and NO_3^-.”

RC: At the very least, some more detail and explanation describing the meaning of Fig 4 and how it was derived and its effect on sulfur could be provided.

AR: In Fig. 4, the re-analysis data are used. The re-analysis data have been constrained by observational data and the reference has been cited. We have also added more detailed discussion, which reads: “The wind vector and wind speed from the North American Regional Reanalysis (NARR), with a spatial resolution of 32 km by 32 km (Mesinger et al., 2006), were thereby analyzed to study the decadal changes in wind fields and associated potential impacts on the long-range transport of air pollutants over the western coastal Canada and U.S. The average wind fields including mean wind vector and speed (shading in Fig 4a-d) in 1990-2011 at 925 hPa showed air masses over the western coastal Canada and U.S. were primarily originated from the Pacific Ocean (Fig. 4a). However, the anomalies of wind fields in 1990-2001 relative to 1990-2009 clearly showed a counterclockwise pattern in the corresponding coastal area, including Site 2., while a clockwise pattern existed in 2002-2011 relative to 1990-2009 (Fig. 4b, c). The anomalies shown in Fig. 4c indicated the northwesterly wind being enhanced in 2002-2011 over the western coastal Canada and U.S., possibly reducing air pollutants being transported from the continent to Site 2. In contrast, the anomalies in Fig. 4b indicated that the northwesterly wind was reduced in 1990-2001. Consequently, more air pollutants might have been transported from the continent to Site 2, resulting in a distinct demarcation in 2002. This hypothesis was also supported by a large rebound of the m-value in 2006-2007, due to the increase in F_{wet} of SO_4^{2-} in 2007. The climate anomalies of wind fields in 2007 relative to 1990-2009 showed a counterclockwise pattern in the north, while the clockwise pattern was pushed to the south (Fig. 4d). With the northwesterly wind being reduced, a greater contribution of air pollutants from the coast of Canada and U.S. to Site 2 might have led to the large increase in F_{wet} of SO_4^{2-} during a few month-long periods in 2007."

RC: The reader will also pause as to why so much network-validated data is omitted. Most of the rationale appears to be statistically based (i.e. 'because it doesn't fit the trend'; see the ±3σ criteria presented online 173) which is insufficient without some scientific support (see the discussion on uncertainty of m-values above). More worrisome is the omission of the m-values (i.e. omission of 24 network-validated datapoints) in 1999 on the basis that they don't fit the expected emission trend and are "probably caused by a large perturbation in climate anomalies", but no real evidence is presented.

AR: The evidence has been added in the revision in a few places, (1) "The abnormally increased F_{wet} of SO_4^{2-} in 1999 was mainly because of the increased precipitation depth (1312 mm), which was the largest during 1998-2011 (the annual average precipitation depth excluding 1999 was 1067±86 mm). However, the geometric average concentration of SO_4^{2-} in precipitation in 1999 (1.0 mg L^{-1}) was close to those in the other years, e.g., 0.9 mg L^{-1} in 1997 and 1998 and 1.0 mg L^{-1} in 2000." (2) "The sharp increase in F_{wet} of NO_3^- occurred mainly in 1999, which was probably due to largely increased annual precipitation depth as mentioned in Section 2.4. The analysis was also supported by the geometric average concentration of NO_3^- in precipitation, which was 1.1 mg L^{-1} in 1999, 5% lower than that in 1988 and only 5-10% higher than those in 1990-1991, 1993 and 2002."

RC: Specific comments (Individual Science Q) On page 8, the authors state (line 172) which in turn increase the relative contribution of the air pollutants' emissions to the calculated value. I assume that the authors are presuming that a monthly change in emissions would not impact the Fwet as much as a large monthly change in precipitation depth or concentration in precipitations. This point should be stressed more in the discussion.

AR: Yes, monthly change in emissions should not impact the F_{wet} as much as large monthly changes in precipitation depth or concentration in precipitations. In the revision, we have added this statement: "Note that monthly changes in emissions may not impact the F_{wet} as much as does a large monthly change in precipitation depth or concentration in precipitation. For example, the monthly average concentrations of SO_2 were almost the same in May, September and October of 1990 (~0.7 μg m^{-3}) while the monthly F_{wet} of SO_4^{2-} varied significantly, e.g., 113, 179 and 532 mg m^{-2} $month^{-1}$, respectively in the same months. The monthly average concentration of SO_2 in February

(4.8 µg m^{-3}) was the largest among the twelve months of 1990, but the corresponding monthly F_{wet} of SO_4^{2-} was the smallest (34 mg m^{-2} $month^{-1}$).”

As show in the revised Fig S2c, the geometric average concentrations of SO_4^{2-} at Site 1 in six months of 1996, including February, April, May, June, September and November, narrowly varied around 0.63±0.05 mg L^{-1} (Fig. S2c). The six months were almost evenly distributed in 12 months of 1996. This also suggests that monthly change in emissions would not impact the monthly geometric average concentrations of SO_4^{2-}. However, the geometric averages largely oscillated from 0.27 mg L^{-1} to 1.77 mg L^{-1} in the other six months of 1996 at the site. Based on the narrow variation in the former six months, it can be inferred that the large oscillation in the latter six months were less likely due to monthly changes in emissions. For example, the value of 1.77 mg L^{-1} in March of 1996 was the largest and approximately two and half times of 0.68 mg L^{-1} in February of 1996. The monthly average concentrations of SO_2 in ambient air were close to each other, i.e., 2.6 µg m^{-3} in March and 2.4 µg m^{-3} in February of 1996. Thus, the large oscillation in the latter six months were very likely due to the effects of climate anomalies imposing on atmospheric processes. However, we cannot quantify what types of climate anomalies caused this. We have added clarification in the revised manuscript and Supporting Information on this point.

RC: The text does not adequately describe Fig 2 (lines 208 to 213). What is shown and why? I assume the objective of this plot is 1) to show the improvement of the fitted trend of the top row (m-value time-series) to the middle row (Fwet time-series). A metric (correlations with emissions?) is needed to demonstrate the advantage of the m-value over the annual Fwet. The secondary objective of Fig 2 is to show the incremental trends or “phases”. The plots do not currently accomplish this as Phase 1, 2, or 3 are not shown. Also, the PLR segments for Phase 1, 2, and 3 identified in the text are not described. The overall fits shown (e.g. R2 = 0.81 in 2a and R2 = 0.62 in 2b) are not significant in the analysis, but are shown on the plots. The PLR segments should be shown for the emissions as well (or at least compared with the 2a PLR segments). Considering both of these objectives, the strength in this technique appears to be that the PLR segments for the top-row more closely resemble the PLR segments for the bottom row and that the PLR segments for the middle row do not reflect this. Please reorganize the discussion and analysis to support this. For example, the lines from 283-293 describing the improvement of the m-values over the annual Fwet data should be elaborated on and moved up in the discussion.

AR: We have made a substantial revision by reorganizing the discussion and analysis in Section 3.1. Three phases have been labeled in Fig 2. The objectives of Fig. 2 have also been added in the context. We agree that the added objectives makes the part more readable.

The overall fits (e.g. $R^2 = 0.81$ in 2a and $R^2 = 0.62$ in 2b with $P<0.01$) shown here are significant, i.e., “SO_4^{2-} and NO_3^- showed decreasing trends from a LR analysis, with R^2

values of 0.81 and 0.71, respectively, and P values <0.01 (Fig. 2a and 2d)."

RC: I assume the phase year classification sproposed by the authors(Phase 1,2 nd 3) are derived from the emissions data patterns, but the logic behind the years of the phases is not specifically discussed (i.e. why 1988 to 1993 and not 1995?) Do the phases align with emissions regulation implementation? The PLR segments are often derived from a set of points as low as N=5 (e.g. Phase 1 from 1988 to 1993). Comparisons should state that this is a low N for comparison.

AR: The three phases of SO_4^{2-} and two phases of NO_3^- were firstly visibly identified by simple screening. We then confirmed the phase results by t-test statistically. This is the simplest way to do PRL analysis if the data size is not too large. The phases were supported by emissions of SO_2 and NOx to some extent, but a few inconsistences still existed, e.g., the almost constant m-value of NO_3^- in Phase 1 against the bell-shape distribution of NOx emission in the same Phase.

We don't think that emission data alone can allow classifying these phases of SO_4^{2-} and NO_3^-. It is well known that real emissions of air pollutants may not always align with emission regulation schedules. Emissions regulation implementation always needs to be examined by using long-term field measurements. However, emission data can facilitate the analysis of phase changes in m-values, since inflection points of different phases of m-values and emissions should be close to each other.

The m-values in 1988-1993 oscillated approximately 1.38±0.08 while the m-value in 1994-1995 largely decreased down to 0.91, the latter period was clearly related to Phase 2 (1994 to 2005) with m-values around 1.02±0.08. The statistical results confirmed the classification.

The sentence has been revised as: "The m-values of SO_4^{2-} and NO_3^- also allowed for the visible and statistical identification of trends in different phases in support by annual variations in emissions of SO_2 and NO_x (Fig. 2c and 2f) to some extent."

In the revision, we have also added: "The three phases generally aligned with the three-phase regulated SO_2 emissions in Ontario. It should be stated that Phase 1 and Phase 3 each covered only six years (N=6), respectively. Cautions should be taken to explain the trend result in each phase in relation to precursors' emissions."

RC: On lines 348 -354; the m-value time-series for Site 2 NO3- (Fig 3d) is interpreted to support the decadal shift hypothesis. However, strictly observing the data, without the hypothesis in mind, it is clear that the fourm-values from 1990 and '96-'00 are elevated, while the values from '92 and '94 are similar to values observed after the decadal shift has taken place. This is acknowledged in the text, but no support given other than it is attributable to climate anomalies.

AR: In the revision, we have added this statement: “For example, the calculated m-values in 1992-1993 and 1994-1995 were evidently lower than the m-values in 1990-2001. However, the annual geometric average concentrations of NO_3^- in 1992-1995 varied around 0.77±0.11 mg L^{-1} and were even larger than the values of 0.66±0.08 mg L^{-1} in 1990-2001 (excluding 1992-1995). The lower m-values were mainly attributed to the lower precipitation depth in 1992-1994 (Fig 3b) rather than lower emissions of NOx.”

*RC: TECHNICAL COMMENTS 1. Figures need descriptive captions and local explanations. 2. Labels on Fig 2 (title incorrect) 3. line 223: “in contrast”. Suggest removal, not really in contrast. 4. line 237: Vlaue should read value 5. For Figure 1, distinguish the outlier point removed for each plot (as done in Fig 2) also specify which fit (R2 and p-value applies to the modified fit (I believe it is *, but it is not labeled).*

AR: Figure captions and labels have all been revised as suggested.