

Supplement of Atmos. Chem. Phys., 17, 9435–9449, 2017  
<https://doi.org/10.5194/acp-17-9435-2017-supplement>  
© Author(s) 2017. This work is distributed under  
the Creative Commons Attribution 3.0 License.



*Supplement of*

## **Sources of particulate matter components in the Athabasca oil sands region: investigation through a comparison of trace element measurement methodologies**

**Catherine Phillips-Smith et al.**

*Correspondence to:* Greg Evans ([greg.evans@utoronto.ca](mailto:greg.evans@utoronto.ca))

The copyright of individual parts of the supplement might differ from the CC BY 3.0 License.

## **S.1 Quality Control and Analysis**

The Xact incorporates several Quality Assurance and Quality Control measures. With each measurement the instrument takes, it simultaneously measures the concentration of a Pd rod that is located within the instrument to ensure measurement stability (Batelle, 2012). Additionally, every day at midnight, the instrument completes several tests. In one of these tests, the Xact measured the concentrations of metals located within an upscale rod made of Pd, Pb, Cr, and Cd. The three metals, Pb, Cr, and Cd, in the upscale rod represent each energy level the instrument (Batelle, 2012). When the instrument is operating under normal conditions, these measurements are constant with each test. This feature was invaluable during the August, 2013 campaign when there was a drop in the internal Pd measurement values between August 25 and September 2, 2013. As this had the potential to alter the measured metal concentrations, the changing Pd upscale value was linearly regressed against the upscale values of Cr, Pb, and Cd, which were found to have slopes of 0.63, 8, and 3.3, respectively (Figure S1). These relationships were assumed to be the same for all metals within that energy level, and measurements made August 25 to September 2 were then adjusted assuming a constant ratio between the upscale metal concentration and the various metals within its energy level. To validate this assumption, a linear comparison of the sulphur (S) data before, as well as both the raw and corrected S data during the incident was compared to the collection-efficiency corrected PM<sub>1.0</sub> SO<sub>4</sub> data measured by a soot particle aerosol mass spectrometer SP-AMS (Willis et al., 2014); the AMS sulphate was divided by three to determine the equivalent sulphur mass. Prior to August 25 the slope of the line was 2.75 (Figure S2a); a slope greater than 1.0 was expected given that PM<sub>2.5</sub> and PM<sub>1.0</sub> mass values were being compared. However, the slope of 2.75 was greater than that expected due to the difference in size cutpoints alone. For example, comparison of the ambient ion monitor ion chromatograph's (AIM-IC) PM<sub>2.5</sub> to the AMS's PM<sub>1.0</sub> data yielded a slope of 1.66 (Figure S3a), which suggested that there is 66% more sulphate in PM<sub>2.5</sub> than in PM<sub>1.0</sub>. The difference in the slopes of 2.75 vs. 1.66 implied that the Xact might be measuring additional sulphur that was not in the form of sulphate. However, comparison of the Xact sulphur with the AIM-IC or PM<sub>2.5</sub> filter data, as described below, indicated that the Xact sulphur values were on average only 40% too high. The additional 17% divergence with the AMS data could not be resolved. As described below, the accuracy of most other metal(oid)s determined by the Xact was much better than that for sulphur.

Correcting the Xact S data for Aug 25-Sept 2 based on the concentration-dependent equations seen in Figure S1 raised the  $r^2$  value from 0.77 to 0.96, and changed the slope of the line from 1.60 to

3.57 (Figures S2c and S2d). Overall, there was a large change in the slope of the line from 2.75 before August 25 to 3.57. However, the corrected slope was comparable to the slope in a similar concentration range (Figure S2b) before August 25 (3.26 vs. 3.57), suggesting that the Xact data correction was reasonable. It is more likely that the data after August 25, 2013 may have been over corrected by up to 20%. This represented 35% of the total Xact data used in the PMF analysis.

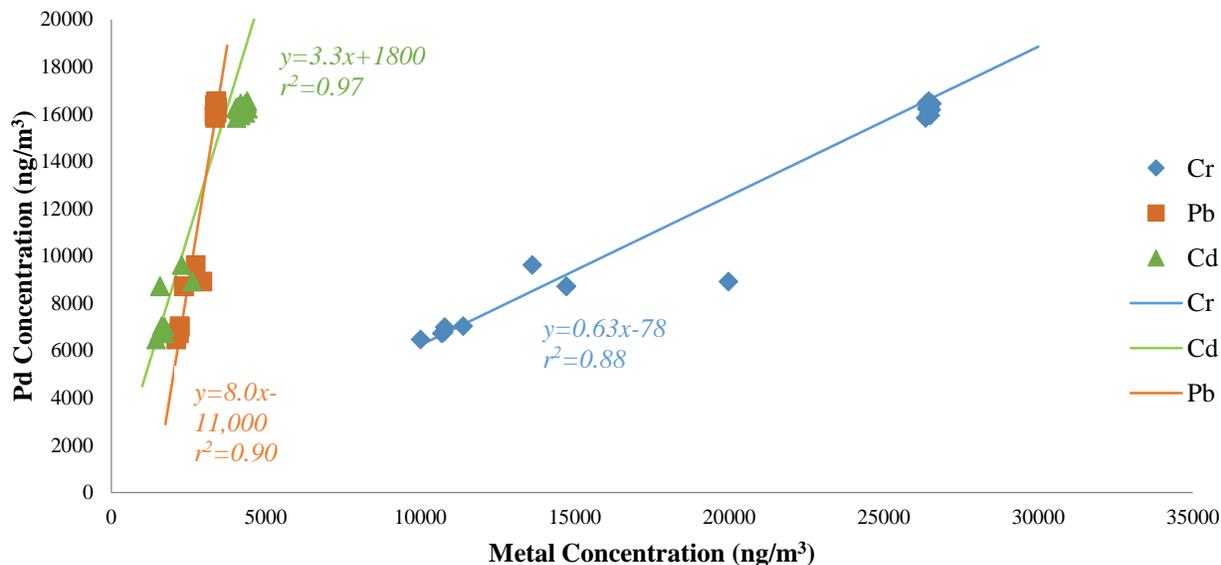


Figure S1. Comparison of Pd rod concentration to the three measured upscale metals: Cr, Cd, and Pb that were measured throughout the intensive campaign.

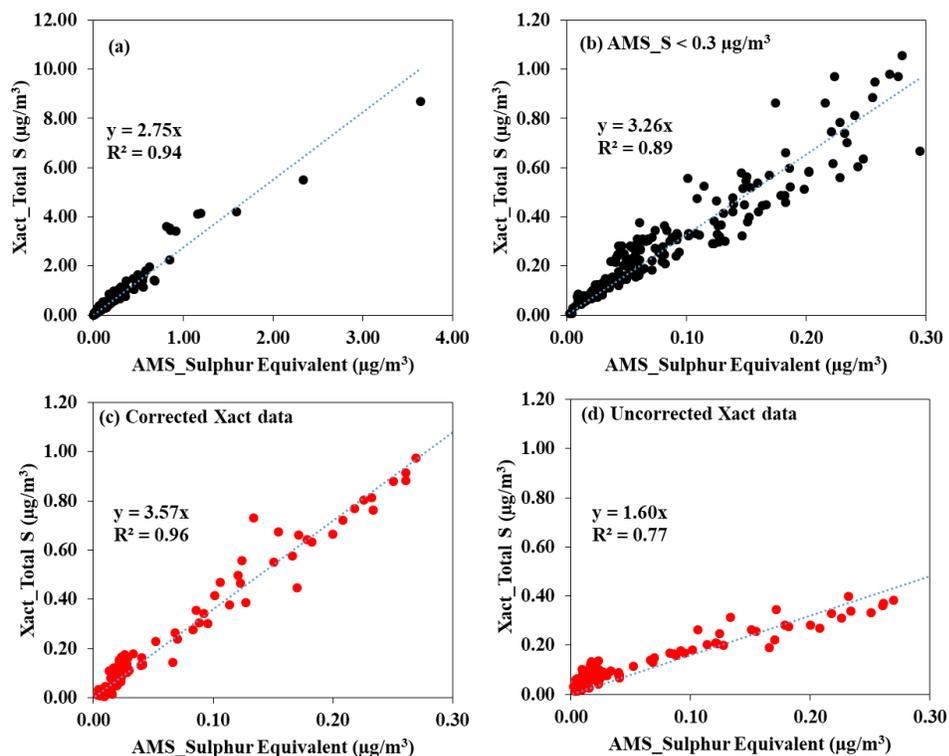
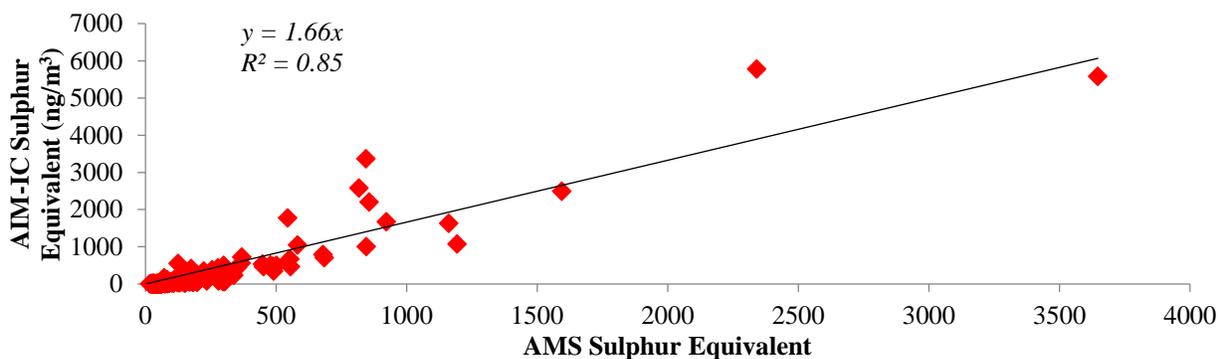


Figure S2. a) Comparison of SP-AMS sulphur equivalent to Xact S before Aug. 25, 2013, b) comparison of SP-AMS sulphur equivalent lower than 300 ng/m<sup>3</sup> to Xact S before Aug. 25, 2013, c) comparison of SP-AMS sulphur equivalent to corrected Xact S data after Aug. 25, 2013, and d) comparison of SP-AMS sulphur equivalent to raw (uncorrected) Xact S data after Aug. 25, 2013. All the SP-AMS sulphate values were divided by three to determine equivalent sulphur mass values.

A comparison of the 1-hr instrumentation data against the coincident 23-hr filter data was conducted to ensure the data measured with the Xact was equivalent to that measured with the filters. The 1-hr data was averaged between 8:30 am and 7:30 am to correspond with the period during which the filter sample was taken. Any averaged values that were below the detection limit (DL), or calculated using data more than 50% of which were below the DL, were removed. The data was then divided into three groups: low, medium, and high concentrations. Low-concentration metals, those with average values  $<10 \text{ ng/m}^3$  (Figure S3b), exhibited excellent agreement, with a linear slope of 1.03 and an  $r^2$  value of 0.95 (Xact to Filter data). This represented 63% of the metals measured by the Xact used in the PMF analysis. The medium-concentration metals, those with averages between  $10 \text{ ng/m}^3$  and  $50 \text{ ng/m}^3$  (Figure S3c), had a slope of 0.77 and an  $r^2$  of 0.99, while the high-concentration elements, with averages above  $50 \text{ ng/m}^3$ , such as sulphur, had a slope of 1.43 and an  $r^2$  of 0.99 (Figure 3d). The sulphur data was also linearly regressed against the  $\text{SO}_4$  data measured with an ambient ion monitor ion chromatograph (AIM-IC), which was divided by three to estimate the S equivalent concentration (Markovic et al., 2012). The result was a line with a slope of 1.42 and an  $r^2$  of 0.84 (Figure 3), this further indicated the likely presence of a 40% bias in the Xact sulphur measurements.



**Figure S3a.  $\text{PM}_{2.5}$  concentrations measured by the AIM-IC to  $\text{PM}_{1.0}$  concentrations measured by the AMS comparison S before Aug. 25, 2013.**

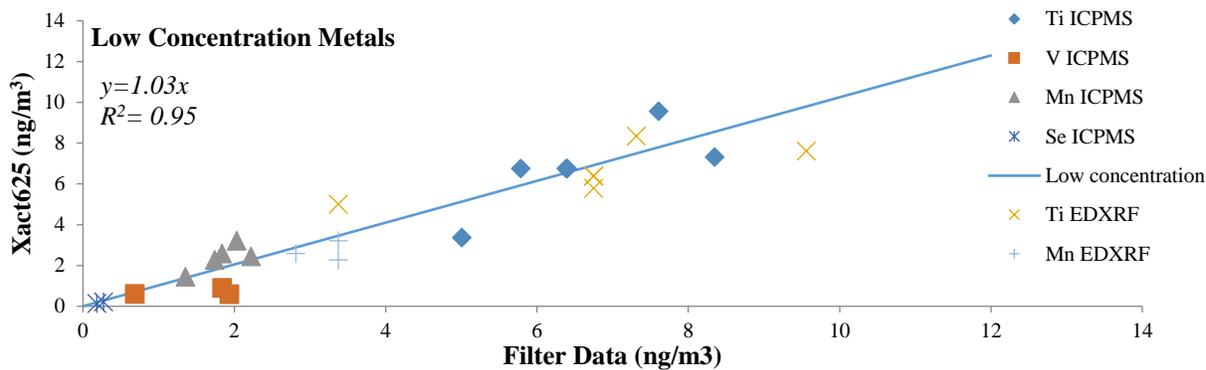


Figure S3b. Filter-Xact comparison for metals with average concentrations < 10 ng/m<sup>3</sup>.

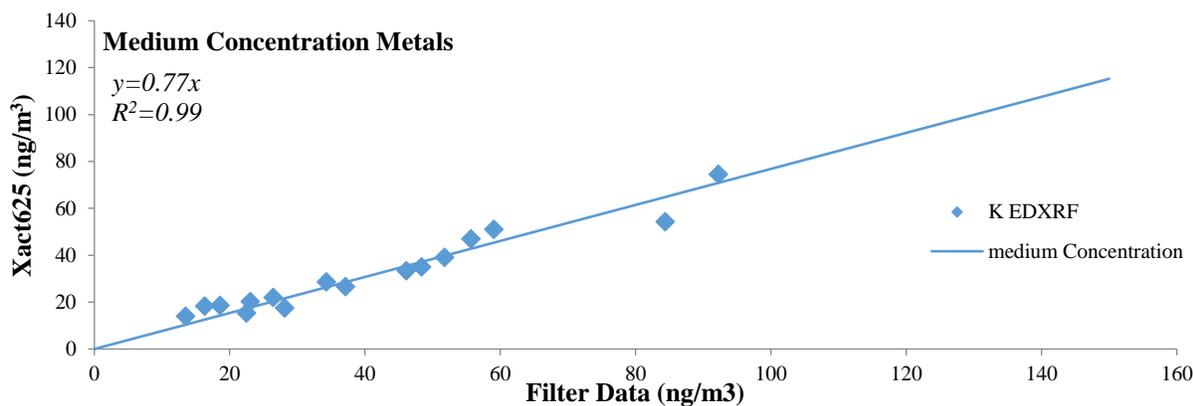


Figure S3c. Filter-Xact comparison for metals with average concentrations >10 ng/m<sup>3</sup> and <50 ng/m<sup>3</sup>.

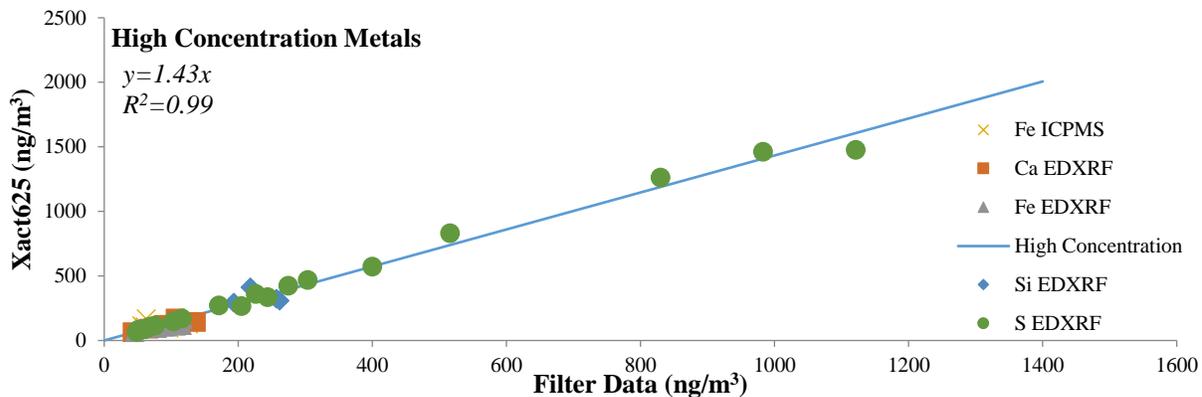


Figure S3d. Filter-Xact comparison for metals with average concentrations > 50 ng/m<sup>3</sup>.

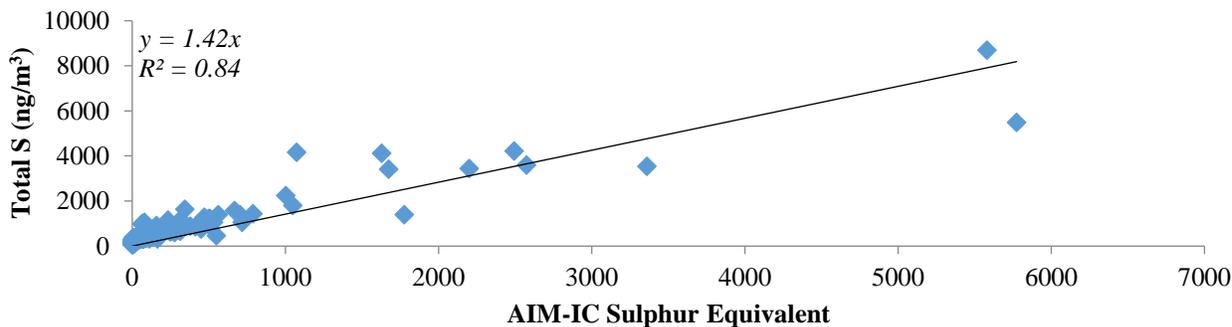


Figure S3e. AIM-IC/Xact comparison for sulphur before Aug. 25, 2013.

After the instrument was returned to the laboratory, all metal standards were run to assess the accuracy, precision, and uncertainty of each metal. High-concentration metal standards, between 7650 and 40852 ng/cm<sup>2</sup> in concentration, were initially run for each metal between 3 and 7 times, and each metal was found to have an accuracy and precision in the range of 98-113% and 0.3-17%, respectively (Table S1). The metal-specific analytic uncertainty was then calculated based on the sum of the average ratio of the difference between the target (T) and measured (X) values of each run ( $\alpha$ ) divided by the target value multiplied by the total number of runs (A), the uncertainty of the flow rate accuracy, set to be 10%, and additional metal-specific uncertainties (M), as seen in Equation 1. If the resulting uncertainty of any metal was less than 10%, it was raised to 10%; if there was no metal standard available for a measure metal, it was assigned an uncertainty based on the average uncertainties of the rest of the metals in the corresponding energy level. The results of this can be seen in Table S1.

$$\left\{ \left| \frac{\sum_{a=1}^A (X-T)}{A \cdot T} \right|^2 + 0.1^2 \right\}^{1/2} + M \quad (1)$$

Concerns that the high concentration metal standards were unrepresentative of the metal concentrations witnessed throughout the campaign led to a secondary test of 6 medium-concentration metal standards; S, V, Ba, Fe, Zn, Ni (Table S1). These metal standards, chosen as they represented all 3 energy levels of the Xact, ranged in value from 490 ng/cm<sup>2</sup> to 2010 ng/cm<sup>2</sup>, were much closer to the instrument's measurements during the campaign (between 0.1 and 1300). Based on these medium standards the metal analysis by the Xact was estimated to have accuracies in the range of 93-113% and precisions in the range of 0.2%-9.5% (Table S1), which was similar to those seen in the high concentration metal standards.

**Table S1. Uncertainties, energy levels, accuracies, and precisions of the metal analysis by the Xact used for the PMF analysis of the intensive monitoring campaign.**

| Element   | Energy Level | <i>High Concentration Element Standards</i> |                                      | <i>Medium Concentration Element Standards</i> |                                      | Uncertainty (%) |
|-----------|--------------|---|--------------------------------------|---|--------------------------------------|-----------------|
|           |              | Accuracy (measured value/target value*100%) | Precision (95% CI/target value*100%) | Accuracy (measured value/target value*100%)   | Precision (95% CI/target value*100%) |                 |
| <i>Si</i> | 1            |   |                                      |   |                                      | 14              |
| <i>S</i>  | 1            | 98  | 12                                   | 108   | 9.5                                  | 12              |
| <i>K</i>  | 1            |   |                                      |   |                                      | 12              |
| <i>Ni</i> | 2            | 113   | 17                                   | 102   | 0.3                                  | 18              |
| <i>Ca</i> | 1            | 108   | 1                                    |   |                                      | 16              |
| <i>Cd</i> | 3            | 99  | 7                                    |   |                                      | 10              |
| <i>Se</i> | 2            | 99  | 1                                    |   |                                      | 10              |
| <i>Mn</i> | 1            | 101   | 4                                    |   |                                      | 10              |
| <i>Ti</i> | 1            | 102   | 4                                    |   |                                      | 10              |
| <i>V</i>  | 1            | 104   | 4                                    | 93  | 3.7                                  | 11              |
| <i>Cr</i> | 1            | 104   | 0.5                                  |   |                                      | 11              |
| <i>Fe</i> | 2            | 100   | 2                                    | 106   | 0.2                                  | 10              |
| <i>Cu</i> | 2            | 99  | 2                                    |   |                                      | 10              |
| <i>Zn</i> | 2            | 102   | 0.3                                  | 113   | 1.9                                  | 10              |
| <i>Br</i> | 2            |   |                                      |   |                                      | 11              |
| <i>Sr</i> | 2            |   |                                      |   |                                      | 10              |

Despite the variations in the slopes that the high, medium, and low concentration metals exhibited when compared to the co-measured filter samples, the  $r^2$  values were high (over 0.95) (Figure 3). This, in addition to the results from the high and medium metal standards, led to the conclusion that the concentrations measured by the Xact were precise and largely accurate for most metals. Unresolved divergence remained among the Xact, AIM-IC, SP-AMS, and filter measurements for some elements such as sulphur. The agreement between the medium and high concentration standards indicated that this was not due to non-linearity in the calibration.

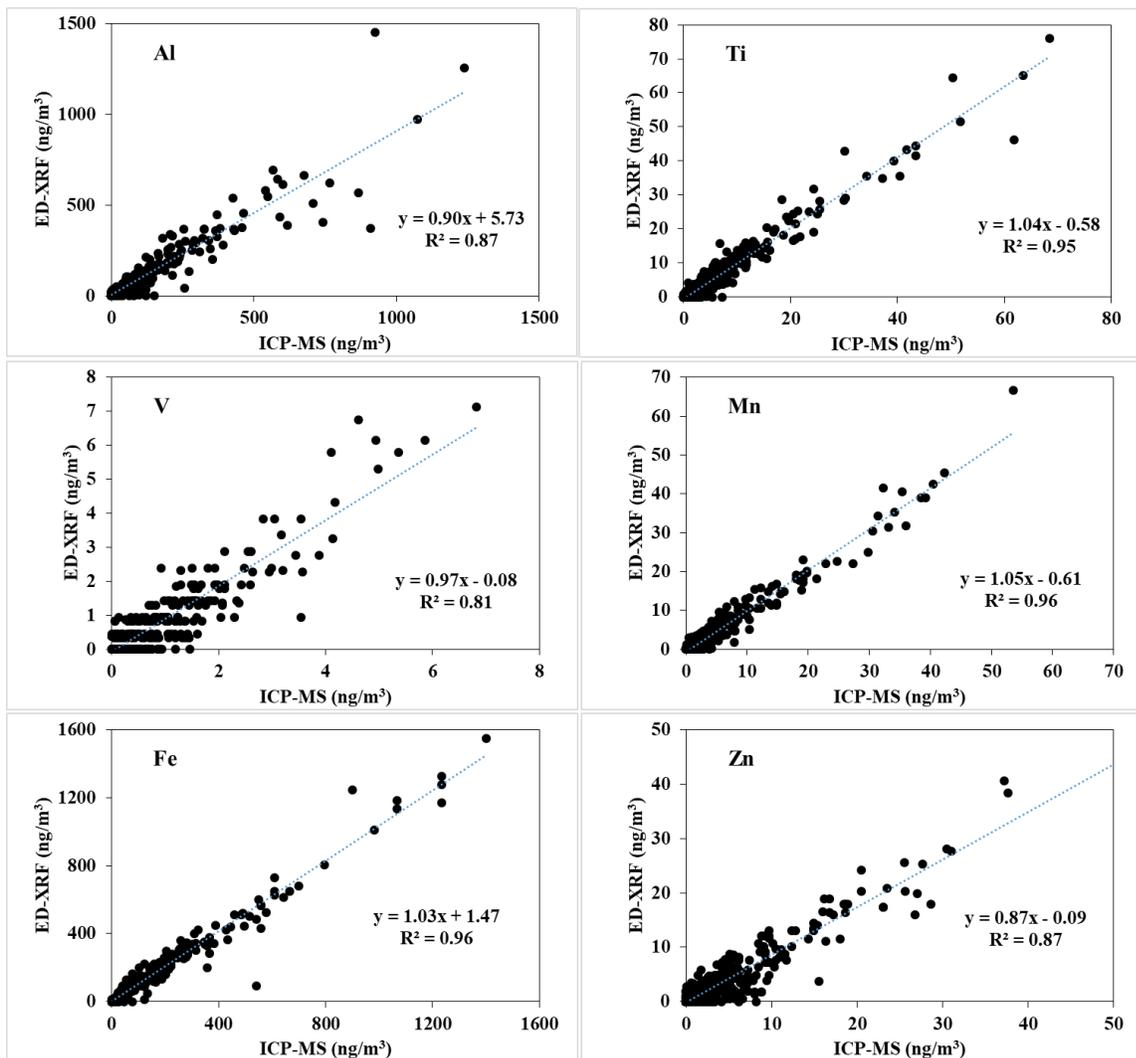


Figure S3f. Comparison of ED-XRF and ICP-MS measurements at AMS 5, 11, and 13 from Dec. 16, 2010 to Nov. 29, 2012. Nine elements with more than 80% of data below the minimum detection limit were excluded from the comparison.

## **S.2 PMF**

Prior to analysis with PMF, the data for both the filter and intensive monitoring campaigns were pre-treated for Quality Control. In addition to only allowing metals which had >10% above DL data, or data that exhibited strong, “plume-like” behaviour when it was above detection limit, into the PMF matrices, the data matrices were also treated for biases. As the intensive monitoring campaign measured one blank value every 24 to 48 -hr, these blank values were averaged to create a single ‘baseline bias’ for each metal from which every datum was subtracted. This was possible as the blank values did not vary from measurement to measurement. This was not the case in the filter campaign. As the blank measurements were spaced a large temporal distance apart, and the values had wide variances, the data measured in this campaign were subtracted from the blank value which was the closest temporally. In order to quantitatively compare the contributions of PMF-resolved sources for the filter and Xact campaigns, the concentrations of major elements (i.e., Si, S, Ca, Fe) measured using Xact were corrected based on the intercomparison of filter and Xact data as described in the previous section.

Special consideration was given to the input data ( $X$ ) and uncertainty ( $\sigma$ ) matrices by sorting the data into three categories: above the DL, below the DL, and missing values (Xie et al., 1999). Data above the DL was input directly, with an uncertainty of  $\sigma_{ij} + DL/3$ . Data below the DL was replaced with  $DL/2$ , and given an uncertainty of  $5(DL)/6$ . Missing values were replaced with the geometric mean of the measured concentrations ( $v$ ), and had the highest uncertainty of the three categories ( $4v$ ) (Xie et al., 1999). Additionally, each remaining metal species was classified as good, weak, or bad depending on the S/N ratio (Equation 2). Metal species with an S/N ratio above 2 were classified as good, while weak data had an S/N ratio between 0.2 and 2, and were down-weighted by a factor of 3 (Paatero and Hopke, 2003, Norris and Duvall, 2014) (Table S2). Data with an S/N ratio below 0.2 were classified as “bad” and given a weight of zero (Paatero and Hopke, 2003). The only exception to this was Cu, which had an S/N ratio of 2.76, but was classified as ‘weak’. This was done as all measurements that were above the DL were very close to the DL, raising their uncertainties. Finally, the data set were all given an additional 10% “Extra Modelling Uncertainty” to further reduce the effects of noise within the data (Norris and Duvall, 2014).

$$S/N = \sqrt{\sum_{i=1}^n x_i^2 / \sum_{i=1}^n \sigma_i^2} \quad (2)$$

**Table S2. Minimum detection limits (DL), S/N ratios, and average values of elements measured by the Xact and analyzed using PMF.**

| Element | DL (ng/m <sup>3</sup> ) | S/N Ratio | Average Value (ng/m <sup>3</sup> ) | Percentage of Missing Data | Percentage of Data Below Detection Limit |
|---------|-------------------------|-----------|------------------------------------|----------------------------|--|
| Si      | 107.0                   | 3.72      | 143                                | 21.13%                     | 45.16%                                   |
| S       | 26.7                    | 8.10      | 468                                | 21.13%                     | 0.97%                                    |
| K       | 11.1                    | 4.29      | 31                                 | 21.13%                     | 16.94%                                   |
| Ca      | 26.2                    | 3.89      | 54                                 | 21.13%                     | 46.29%                                   |
| Ti      | 1.9                     | 4.79      | 3.4                                | 21.13%                     | 39.35%                                   |
| V       | 0.3                     | 3.22      | 0.21                               | 21.13%                     | 60.81%                                   |
| Cr      | 5.1                     | 1.82      | 0.04                               | 21.13%                     | 69.35%                                   |
| Mn      | 0.4                     | 2.40      | 1.12                               | 21.13%                     | 53.71%                                   |
| Fe      | 21.8                    | 4.78      | 60                                 | 21.13%                     | 41.29%                                   |
| Ni      | 0.3                     | 0.80      | 0.08                               | 21.13%                     | 73.55%                                   |
| Cu      | 1.8                     | 2.76      | 2.04                               | 21.13%                     | 59.52%                                   |
| Zn      | 1.4                     | 0.75      | 0.88                               | 21.13%                     | 77.74%                                   |
| Se      | 0.2                     | 1.50      | 0.07                               | 21.13%                     | 59.68%                                   |
| Br      | 0.2                     | 4.79      | 0.54                               | 21.13%                     | 16.7%                                    |
| Sr      | 0.9                     | 0.82      | 0.23                               | 21.13%                     | 71.77%                                   |

### **S.3 Evaluation of the PMF solution**

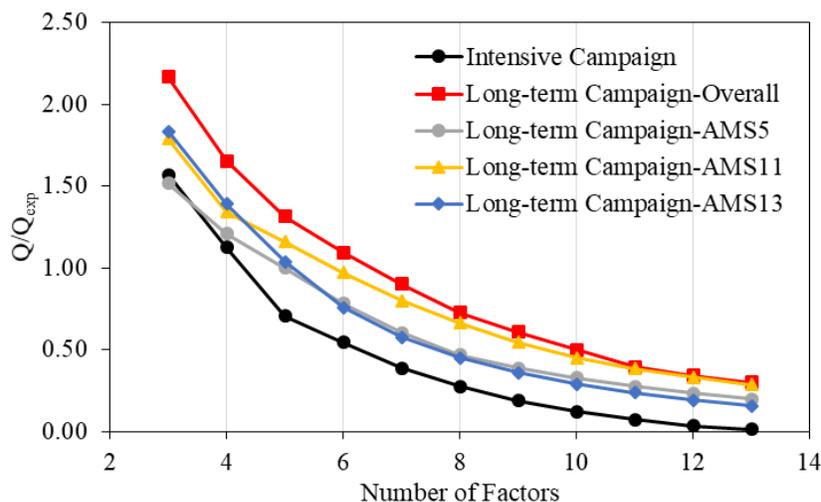
The Q value (Equation 3) calculated by PMF algorithm was used to determine the number of resolvable factors affecting the receptor site, through its stability and inflection point. Because different pseudorandom numbers are selected by the algorithm at the initiation of each run, a stable solution with a preset number of factors will not experience a change in the Q value through a series of runs (Xie et al., 1999). Also, since the ratio of Q to the expected Q ( $Q_{exp}$ , Equation 4) of the solution decreases with an increase in the preset number of output factors, an inflection point in the rate of decrease will indicate the most central, or optimal, solution (Xie et al., 1999) (Figure S4). The optimal number of factors was also determined by examination of the G-space plot, a direct comparison of the time series (G matrix) of one factor against that of another factor, and the scaled residuals of the results. Factor solutions that contain metals with a non-normal distribution in the scaled residual could signify that the uncertainty of the metal is too low, that there is noise in the data, or that there is another pollutant source of the metal not separated in the factor solution

(Paatero, 1996). It should also be noted that a linear relationship between two factors in the G-space plot could signify either a co-aligned or identical source (Paatero et al., 2005).

$$Q = \sum_{i=1}^n \sum_{j=1}^m e_{ij}^2 / \sigma_{ij}^2 \quad (3)$$

$$Q_{exp} = nm - p(n+m) \quad (4)$$

where  $n$  is the number of samples and  $m$  is the number of species,  $e_{ij}$  is the residual concentration of the  $j$ th species in the  $i$ th sample;  $\sigma_{ij}$  is the uncertainty of the  $j$ th species in the  $i$ th sample, and  $p$  is the number of factors.



**Figure S4.** Q analysis of the intensive and long-term campaigns.

The final parameter examined in the selection of the optimal solution for the receptor site was the user-specified rotational parameter,  $\Phi$ , chosen for the FPEAK analysis. When the PMF algorithm is run normally, i.e. with  $\Phi=0$ , it finds the optimal solution the farthest away from any zeros in both the G and the F matrices (Paatero et al., 2002). When the value of  $\Phi$  is changed to lie between -2 and 2, in the FPEAK analysis, the rotation of the algorithm is altered to allow for zeros or negative numbers in either the G or the F matrix. A stable solution will not change radically when the rotation,  $\Phi$ , is changed, but the change in rotation may “clean” the solution by allowing the minor elements of the factor to go to zero. Throughout this study, the value of  $\Phi$  was selected to be 0.

Both the filter and the Xact data were analyzed separately using PMF due to their difference in the sampling intervals. Filter data from the monitoring and the filter campaign were combined to

produce a single data matrix. Additionally, the filter data from all three sites were combined into one data set prior to running the PMF algorithm. The goal of this was to take advantage of the close proximity and remote nature of the three sampling locations, assuming most, if not all aerosol sources are common among the sites, at the expense of independent PMF solutions. The spatial variability of the sites may also add extra factor-discriminating power, as each site is in a different direction to the various sources. As shown in Table S3, this technique succeeded in producing a stable 5-factor solution, which was similar to the 4- or 5-factor solutions produced when the filter data from each site was run independently (AMS 5, AMS 11, and AMS 13). Not only did the 5-factor solution exhibit the most central and stable Q factor, but also proved to be the most physically meaningful when compared to the solutions with 6 and 7 factors.

**Table S3. Standard deviations of the Q-factor across 150 PMF runs for the intensive monitoring and long-term campaigns.**

| Number of Factors | Intensive Monitoring Campaign | Long-term Campaign-Combined Filters | Long-term Campaign AMS13 | Long-term Campaign-AMS5 | Long-term Campaign-AMS11 |
|-------------------|-------------------------------|-------------------------------------|--------------------------|-------------------------|--------------------------|
| 3                 | 0.01                          | 396.97                              | 130.76                   | 65.19                   | 0.06                     |
| 4                 | 14.47                         | 139.66                              | 62.06                    | 0.03                    | 0.01                     |
| 5                 | 0.00                          | 21.94                               | 22.70                    | 15.55                   | 1.75                     |
| 6                 | 1.27                          | 27.92                               | 27.94                    | 18.57                   | 5.35                     |
| 7                 | 0.01                          | 6.91                                | 17.17                    | 14.12                   | 22.27                    |
| 8                 | 0.01                          | 84.40                               | 12.39                    | 0.84                    | 11.66                    |
| 9                 | 0.05                          | 17.01                               | 12.63                    | 6.16                    | 11.97                    |
| 10                | 0.12                          | 16.50                               | 4.02                     | 3.27                    | 1.60                     |
| 11                | 0.01                          | 2.42                                | 7.03                     | 6.85                    | 1.13                     |
| 12                | 0.08                          | 5.56                                | 8.60                     | 3.70                    | 1.68                     |
| 13                | 0.02                          | 6.79                                | 5.62                     | 2.64                    | 1.34                     |

EPA PMF 5.0 includes useful tools to estimate uncertainties and evaluate the robustness and rotational ambiguity of PMF modeling results. The bootstrap (BS) analysis was conducted to evaluate the uncertainties (i.e. random error in data values) of the source profiles and the reproducibility of factors in every bootstrap (Paatero et al., 2014; Brown et al., 2015). In the BS analysis, the BS factors are compared with the base run factors and then mapped to the base factor if the correlation is higher than a threshold ( $r^2=0.8$  in this study). Tables S4 and S5 summarize the diagnostics of the error estimation for three PMF solutions (i.e. 4-, 5-, 6-factor solution) for the intensive Xact data and the long-term filter data, respectively. In the 5-factor solution of both Xact and filter data, we found most bootstrap factors were well assigned to base

factors in >96% of every bootstrap. Overall reproducibility (i.e. average BS mapping percentages) for each factor in the 5-factor solution was higher than other solutions, suggesting the 5-factor solution was very reproducible and the optimal solution. The displacement (DISP) analysis was conducted to evaluate rotational ambiguity in the PMF solution as well. Multiple solutions may be generated with the same value of the object function Q due to rotational ambiguity. In DISP, each fitted element (only good species) in a source profile is displaced in turn from its fitted value until Q increases by a predetermined maximum change in Q. An uncertainty estimate for each element in each factor profile is thereby yielded and factor swaps may occur if factors change too much. A comprehensive error estimate method, bootstrap enhanced by displacement (BS-DISP) combine the strengths of BS and DISP, which evaluate both the robustness to data errors and rotational uncertainty. Overall, no change in DISP Q (%dQ) was found for the 5-factor solutions. Furthermore, no swapped factor was found in DISP BS-DISP runs, indicating the 5-factor solution was a global minimum and well-defined PMF solution.

The source profiles of the 4- and 6-factor solution for the intensive Xact PMF analysis are shown in Figures S5 and S6. In the 4-factor solution, the Soil and Haul Road Dust factors can be combined, but the reproducibility of the solution was poor and there were factor swaps in the BS-DISP runs. In the 6-factor solution, the Soil factor from the 5-factor solution split into two similar soil factors which have poor BS mapping reproducibility and very high factor swaps in the BS-DISP analysis. Another solution could be possible in the 6-factor solution which was characterized by additional resolution of the Mixed sources in the 5-factor solution. As shown in Figure S6b, two Mixed sources were characterized by high loadings of Cu (Mixed Sources I) and Br and Se (Mixed Sources II). Thus, due to the robustness of the solution and the physically meaningfulness of the factor profiles, the 5-factor solution was clearly acceptable for the intensive campaign.

In the long-term campaign the Haul Road Dust and Soil factors in the 5-factor solution can be combined into one factor in the 4-factor solution (Figure S7a). However, there was an alternative solution including a combined factor of Mixed and Upgrader Emissions (Figure S7b). Due to the instability, there were factor swaps in BS-DISP and the reproducibility of the 4-factor solution was poor. With 6 factors, an additional Soil factor characterized by high loadings of rare earth elements and vanadium was found. However, this second Soil factor is only found 56% of the

BS resamples and 85% of the BS-DISP runs were accepted with high factor swaps. In the 7-factor solution, additional factor containing high Pb and Br can be isolated from the Upgrader Emissions factor, but it's stability was very poor and there was no reasonable source for Pb and Br only. The BS resamples and BS-DISP runs of the 5-factor solution was better than the 4- and 6-factor solutions for the long-term filter data (Table S5). These results indicate that the 4- and 6-factor solution are much less certain than the 5-factor solution. As a result, the 5-factor solution was chosen as the most reasonable and stable solution for the filter data.

In order to determine the relative weights of the different factors, a multiple linear regression of the time series of each factor for both campaigns was run against both the summed metal concentrations as well as the PM<sub>2.5</sub> concentrations (obtained from WBEA). As trace metals only account for a small percentage of the overall PM<sub>2.5</sub> mass, the results of the PM<sub>2.5</sub> regression proved to be a poor fit both statistically ( $r^2 < 0.8$ ) and physically, as it resulted in negative relative weights, to the metal speciation factor solutions. Because of this, the total metals concentration was used to determine the relative weights of the different factors, which resulted in a much better fit ( $r^2 > 0.99$ ).

**Table S4. Summary of error estimation diagnostics for intensive Xact data.**

|                                | 4-Factor Solution                          | 5-Factor Solution           | 6-Factor Solution            |
|--------------------------------|--|-----------------------------|------------------------------|
| Robust Mode                    | Yes  |                             |                              |
| Seed Value                     | Random                                     |                             |                              |
| # of Bootstraps in BS          | 100  |                             |                              |
| R <sup>2</sup> in BS           | 0.8  |                             |                              |
| DISP active species            | Si, S, K, Ca, Ti, V, Mn, Fe                |                             |                              |
| BS-DISP active species         | S, K, Ti, V, Fe                            |                             |                              |
| Factors with BS mapping < 100% | Upgrader Emissions II(51%),<br>Mixed (99%) | Upgrader Emissions II (98%) | Soil (53%),<br>Soil II (93%) |
| DISP %dQ                       | 0  | 0                           | 0                            |
| DISP # of swaps                | 0  | 0                           | 0                            |
| BS-DISP % of Cases Accepted    | 95   | 97                          | 27                           |
| BS-DISP # of swaps             | 4  | 0                           | 198                          |

**Table S5. Summary of error estimation diagnostics for long-term combined filter data.**

|                                | 4-Factor Solution   | 5-Factor Solution          | 6-Factor Solution                                   |
|--------------------------------|---|----------------------------|---|
| Robust Mode                    | Yes   |                            |   |
| Seed Value                     | Random  |                            |   |
| # of Bootstraps in BS          | 100   |                            |   |
| R <sup>2</sup> in BS           | 0.8   |                            |   |
| DISP active species            | Si, S, K, Ca, Ti, Fe, Cu, Sr, Al, Cd, Ce, La, Pr, Nd, Sm, Gd, Pb, U |                            |   |
| BS-DISP active species         | Si, S, K, Ca, Fe, Cu, La  |                            |   |
| Factors with BS mapping < 100% | Mixed (29%),<br>Biomass Burning (98%)                               | Mixed (96%),<br>Soil (99%) | Soil (56%),<br>Haul Road Dust (97%),<br>Mixed (99%) |
| DISP %dQ                       | 9.3E-5  | 0                          | 1.8E-5  |
| DISP # of swaps                | 0   | 0                          | 0   |
| BS-DISP % of Cases Accepted    | 87  | 97                         | 85  |
| BS-DISP # of swaps             | 10  | 0                          | 12  |

## S.4 PMF Factor Solutions

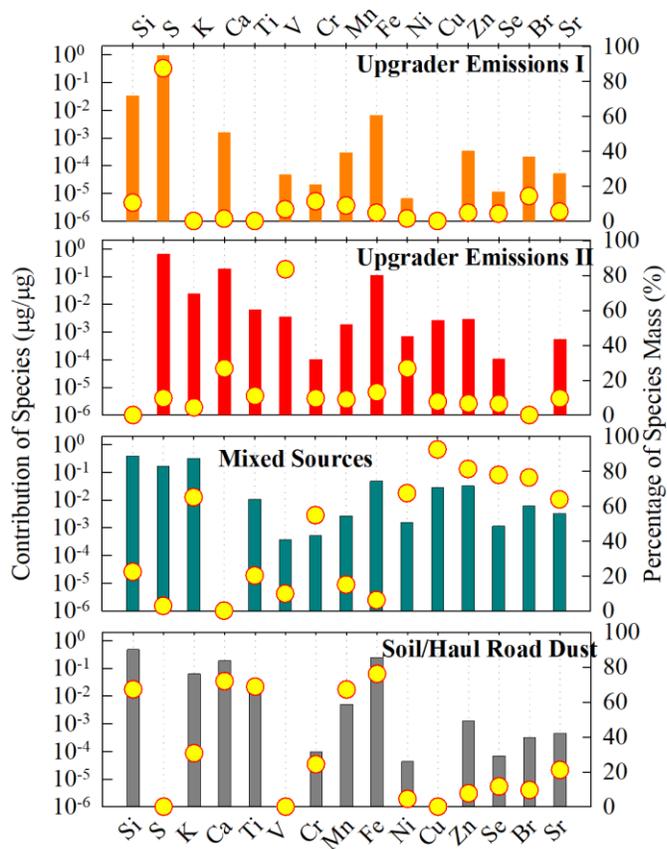


Figure S5. Alternative 4-factor solution calculated using PMF for the intensive campaign (Xact metal data) Factor Concentrations depicted as bars, percentages depicted as circles.

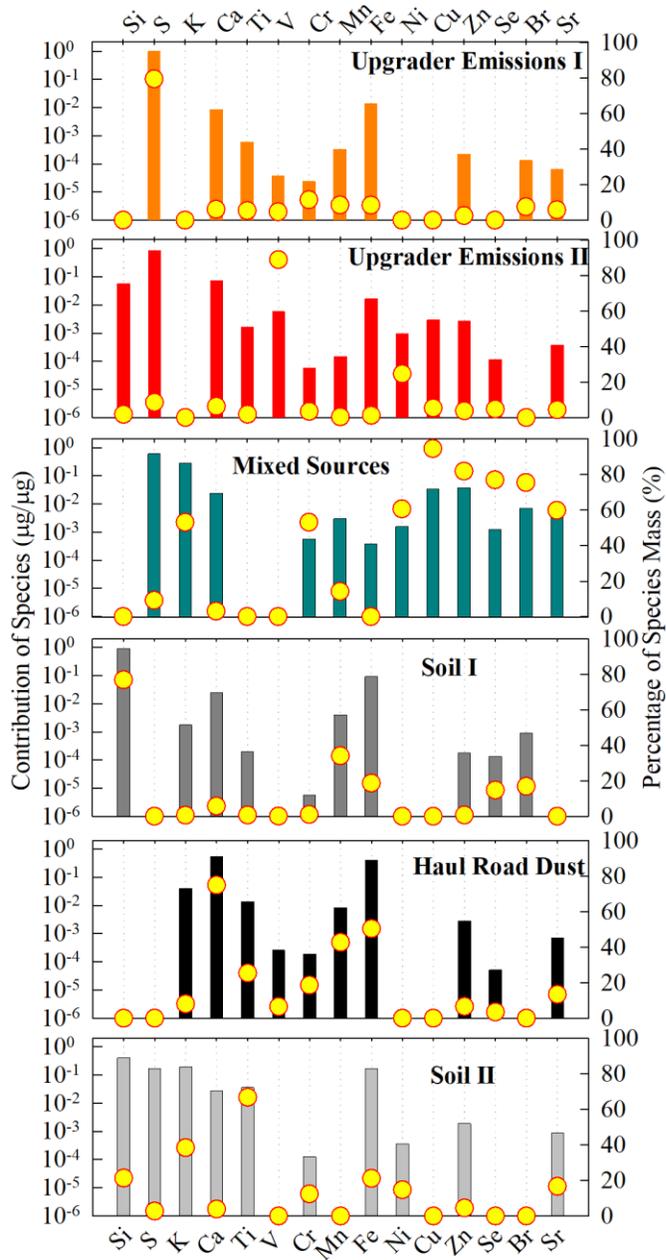


Figure S6a. Alternative 6-factor solution calculated using PMF for the intensive campaign (Xact metal data). Factor Concentrations depicted as bars, percentages depicted as circles.

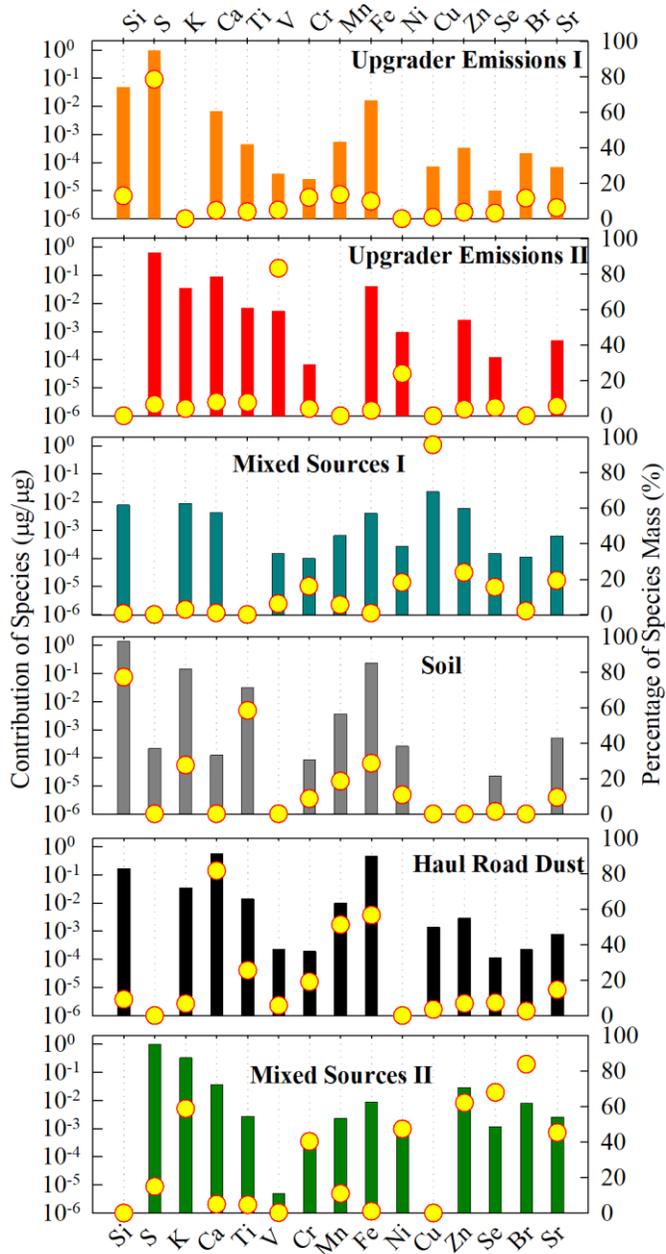
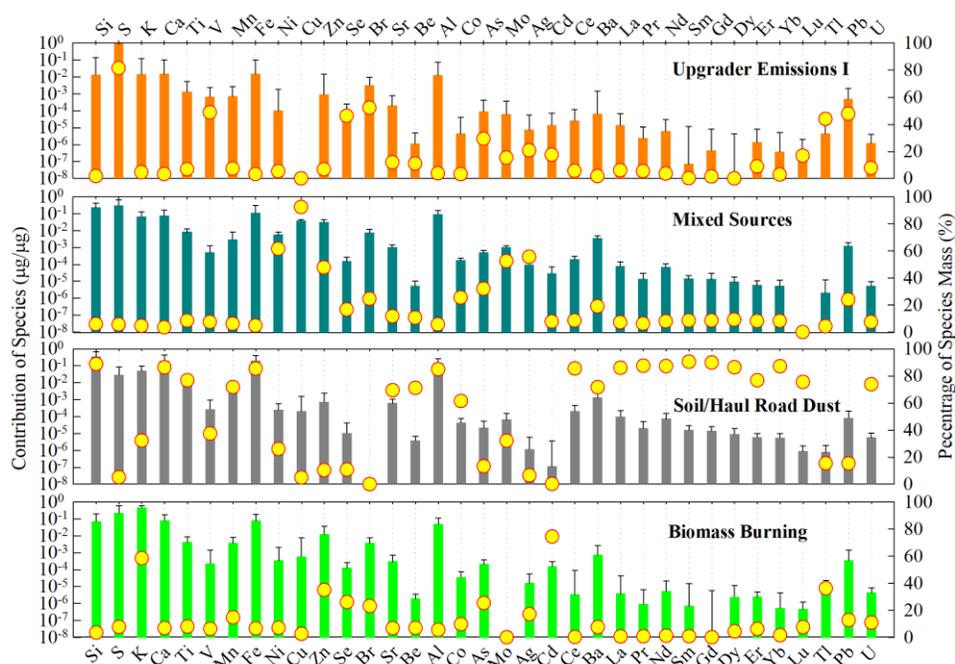
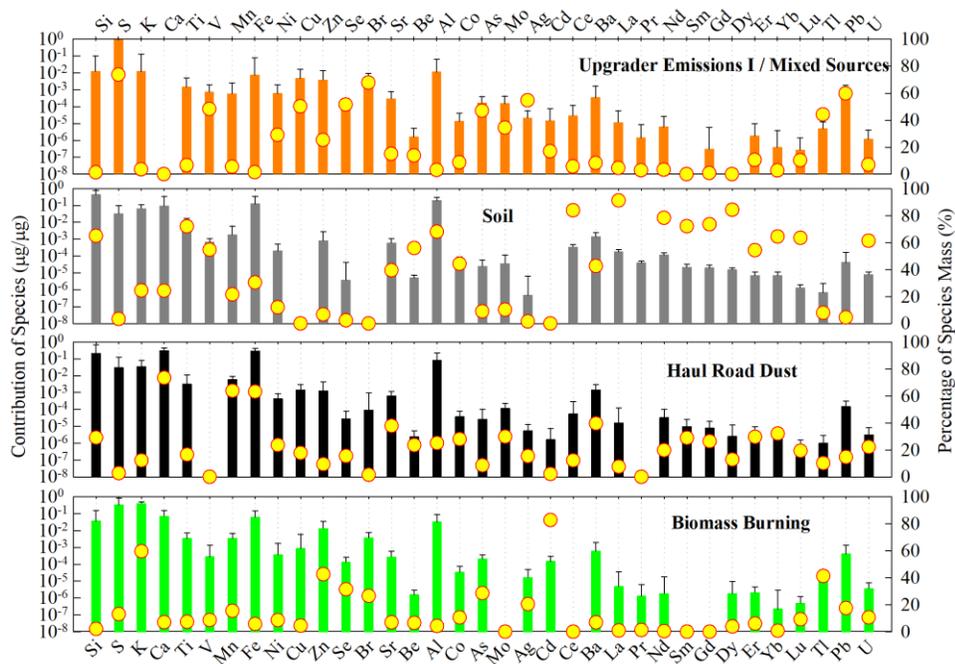


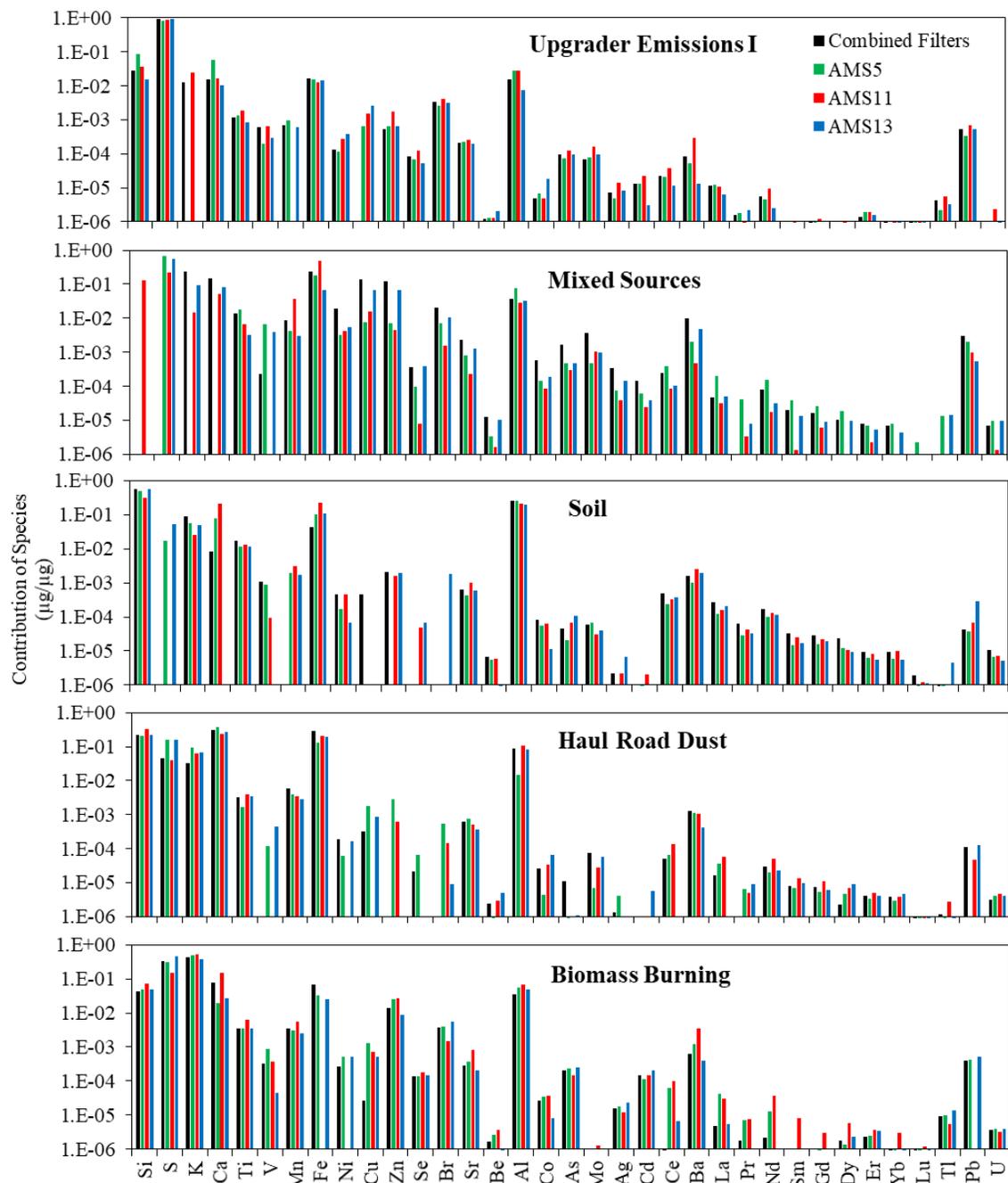
Figure S6b. Alternative 6-factor solution calculated using PMF for the intensive campaign (Xact metal data). Factor Concentrations depicted as bars, percentages depicted as circles.



**Figure S7a. Alternative 4-factor solution calculated using PMF for the long-term campaign (Filter data). Error bars represent standard deviations estimated by 100 bootstrap runs. Factor Concentrations depicted as bars, percentages depicted as circles.**



**Figure S7b. Alternative 4-factor solution calculated using PMF for the long-term campaign (Filter data). Error bars represent standard deviations estimated by 100 bootstrap runs. Factor Concentrations depicted as bars, percentages depicted as circles.**



**Figure S8. 5-factor solutions for 4 independently run PMF analyses of the long-term campaign data (Integrated filter data). Black is the overall solution run after combining the data from all three sites; green is the independently run AMS5 data; red is the independently run AMS11 data; and blue is the independently run AMS13 data.**

## S.5 Overall Species comparison

Table S6. Average concentrations of elements used in the PMF analysis of the long-term campaign average and 90<sup>th</sup> percentile (Dec. 2010-Nov. 2012, Aug. 2013) for all three sites compared to average metal concentrations in Halifax, St John, Montreal, Windsor, Toronto, Edmonton, and Vancouver (Environment Canada, 2015). All metal(oid)s are measured either by acid digested ICP-MS <sup>(1)</sup> or ED-XRF <sup>(2)</sup>.

| Element (ng/m <sup>3</sup> ) | Oil Sands - Average | Oil Sands-90 <sup>th</sup> Percentile | Halifax, NS | St John, NB | Montreal, QC | Windsor, ON | Toronto, ON | Edmonton, AB | Vancouver, BC |
|------------------------------|---------------------|---------------------------------------|-------------|-------------|--------------|-------------|-------------|--------------|---------------|
| <i>Ag</i> <sup>1</sup>       | 0.02                | 0.02                                  | 0.02        | 0.02        | 0.04         | 0.03        | 0.03        | 0.03         | 0.03          |
| <i>Al</i> <sup>2</sup>       | <b>121</b>          | <b>290</b>                            | 36          | 31          | 43           | 29          | 31          | 47           | 29            |
| <i>As</i> <sup>1</sup>       | 0.12                | 0.25                                  | 0.25        | 0.25        | 0.83         | 1.0         | 0.48        | 0.41         | 0.51          |
| <i>Ba</i> <sup>1</sup>       | 1.4                 | <b>3.5</b>                            | 1.2         | 0.88        | 1.79         | 2.26        | 2.45        | 2.87         | 2.42          |
| <i>Be</i> <sup>1</sup>       | N/A                 | N/A                                   | 0.005       | 0.004       | 0.005        | 0.005       | 0.004       | 0.005        | 0.004         |
| <i>Br</i> <sup>2</sup>       | 2.1                 | <b>4.3</b>                            | 2.0         | 1.8         | 2.5          | 2.7         | 2.0         | 2.6          | 2.0           |
| <i>Ca</i> <sup>2</sup>       | <b>154</b>          | <b>390</b>                            | 23          | 14          | 58           | 69          | 45          | 60           | 19            |
| <i>Cd</i> <sup>1</sup>       | 0.04                | 0.08                                  | 0.03        | 0.04        | 0.19         | 0.16        | 0.08        | 0.13         | 0.07          |
| <i>Ce</i> <sup>1</sup>       | 0.16                | 0.35                                  |             |             |              |             |             |              |               |
| <i>Co</i> <sup>1</sup>       | 0.05                | 0.11                                  | 0.12        | 0.02        | 0.03         | 0.02        | 0.03        | 0.15         | 0.03          |
| <i>Cr</i> <sup>1</sup>       | N/A                 | N/A                                   | 0.23        | 0.21        | 0.54         | 0.43        | 0.37        | 1.6          | 0.55          |
| <i>Cu</i> <sup>1</sup>       | 2.7                 | <b>6.0</b>                            | 2.0         | 1.2         | 3.2          | 3.3         | 2.8         | 3.6          | 2.7           |
| <i>Dy</i> <sup>1</sup>       | 0.01                | 0.02                                  |             |             |              |             |             |              |               |
| <i>Er</i> <sup>1</sup>       | 0.01                | 0.02                                  |             |             |              |             |             |              |               |
| <i>Fe</i> <sup>2</sup>       | <b>151</b>          | <b>350</b>                            | 32          | 20          | 59           | 123         | 51          | 120          | 52            |
| <i>Gd</i> <sup>1</sup>       | 0.01                | 0.03                                  |             |             |              |             |             |              |               |
| <i>K</i> <sup>2</sup>        | <b>97</b>           | <b>200</b>                            | 51          | 42          | 75           | 70          | 48          | 66           | 56            |
| <i>La</i> <sup>1</sup>       | 0.08                | 0.17                                  |             |             |              |             |             |              |               |
| <i>Lu</i> <sup>1</sup>       | N/A                 | L/A                                   |             |             |              |             |             |              |               |
| <i>Mn</i> <sup>1</sup>       | 5.1                 | <b>13</b>                             | 0.91        | 0.60        | 2.2          | 4.5         | 1.6         | 7.7          | 2.6           |
| <i>Mo</i> <sup>1</sup>       | 0.21                | 0.39                                  | 0.11        | 0.09        | 0.31         | 0.33        | 0.18        | 0.40         | 0.22          |
| <i>Ni</i> <sup>1</sup>       | 0.94                | 1.4                                   | <b>2.3</b>  | 0.76        | 0.73         | 0.75        | 0.46        | <b>1.5</b>   | <b>1.4</b>    |
| <i>Nd</i> <sup>1</sup>       | 0.06                | 0.14                                  |             |             |              |             |             |              |               |
| <i>Pb</i> <sup>1</sup>       | 0.42                | 0.99                                  | 1.5         | 0.85        | 3.9          | 4.3         | 2.0         | 1.4          | 2.5           |
| <i>Pr</i> <sup>1</sup>       | 0.02                | 0.04                                  |             |             |              |             |             |              |               |
| <i>S</i> <sup>2</sup>        | 373                 | <b>760</b>                            | 334         | 357         | 415          | <b>720</b>  | 455         | 331          | 248           |
| <i>Se</i> <sup>1</sup>       | 0.09                | 0.15                                  | 0.19        | 0.20        | 0.63         | 1.1         | 0.57        | 0.15         | 0.15          |
| <i>Si</i> <sup>2</sup>       | <b>265</b>          | <b>670</b>                            | 46          | 29          | 54           | 53          | 40          | 14           | 26            |
| <i>Sm</i> <sup>1</sup>       | 0.01                | 0.03                                  |             |             |              |             |             |              |               |
| <i>Sr</i> <sup>1</sup>       | N/A                 | N/A                                   | 0.50        | 0.48        | 0.69         | 0.78        | 0.58        | 0.50         | 0.62          |
| <i>Ti</i> <sup>1</sup>       | <b>7.3</b>          | <b>15</b>                             | 0.85        | 0.32        | 1.6          | 0.52        | 0.35        | 0.61         | 0.38          |
| <i>Tl</i> <sup>1</sup>       | 0.01                | 0.03                                  | 0.02        | 0.02        | 0.02         | 0.03        | 0.02        | 0.01         | 0.01          |
| <i>U</i> <sup>1</sup>        | N/A                 | N/A                                   |             |             |              |             |             |              |               |
| <i>V</i> <sup>1</sup>        | 0.84                | 1.9                                   | <b>2.9</b>  | 1.2         | 0.81         | 0.71        | 0.23        | 0.35         | <b>2.4</b>    |
| <i>Yb</i> <sup>1</sup>       | N/A                 | N/A                                   |             |             |              |             |             |              |               |
| <i>Zn</i> <sup>1</sup>       | 6.4                 | 14                                    | 6.5         | 8.6         | 16           | <b>22</b>   | 10          | 12           | 8.9           |

## **S.6 Shannon Entropy**

Shannon entropy is a theoretical measure of the uncertainty associated with a random variable (Healy et al., 2014). Shannon entropy was used to quantify how each species was partitioned across various factors. The bulk population diversity ( $D_i$ ) for each species ( $i$ ) was determined based on the mass fractions of each species ( $p^a$ ) for each factor ( $a$ ) across the total number of factors ( $A$ ). A species equally distributed across all five factors would have a diversity of 5: a species with a value of 1 is considered to be entirely distributed to a single factor. Overall, Shannon Entropy is an information-theoretic measure that has been previously used to indicate biodiversity (Whittaker, 1965), genetic diversity (Rosenburg et al., 2002), and economic diversity (Attaran, 1986). More recently, Shannon entropy has been used to analyze how chemical species present in particles are distributed in an aerosol population (Healy et al., 2014; Reimer and West, 2013). In this study, Shannon entropy was used to analyze how the metals species were distributed across the various factors. A metal species with a diversity value above 3.5 was considered to be equally distributed, and, therefore, that species could not be used as a characteristic species for factor identification.

$$D_i = e^{\sum_{a=1}^A -p^a \ln p^a} \quad (5)$$

**Table S7. Diversity values for the elements analyzed in the intensive monitoring and long-term campaigns. Species in bold had diversity values greater than 3.5 and thus were too diverse to be used to distinguish between factors.**

| Elements  | Intensive Campaign | Long-term Campaign |
|-----------|--------------------|--------------------|
| <i>Si</i> | 2.0                | 2.4                |
| <i>S</i>  | 2.1                | 1.8                |
| <i>K</i>  | 2.4                | 3.3                |
| <i>Ca</i> | 2.3                | 2.2                |
| <i>Ti</i> | 2.6                | 3.3                |
| <i>V</i>  | 1.6                | 2.4                |
| <i>Cr</i> | <b>3.9</b>         |                    |
| <i>Mn</i> | 3.5                | 2.3                |
| <i>Fe</i> | 3.1                | 2.1                |
| <i>Ni</i> | 2.6                | <b>3.7</b>         |
| <i>Cu</i> | 1.4                | 1.9                |
| <i>Zn</i> | 2.0                | 2.8                |
| <i>Se</i> | 2.2                | 3.6                |
| <i>Br</i> | 1.5                | 2.6                |
| <i>Sr</i> | 3.5                | <b>3.6</b>         |
| <i>Be</i> |                    | <b>3.5</b>         |
| <i>Al</i> |                    | 2.6                |
| <i>Co</i> |                    | <b>3.8</b>         |
| <i>As</i> |                    | <b>4.2</b>         |
| <i>Mo</i> |                    | 2.9                |
| <i>Ag</i> |                    | 3.4                |
| <i>Cd</i> |                    | 2.2                |
| <i>Ce</i> |                    | 2.6                |
| <i>Ba</i> |                    | 2.8                |
| <i>La</i> |                    | 2.6                |
| <i>Pr</i> |                    | 2.4                |
| <i>Nd</i> |                    | 2.5                |
| <i>Sm</i> |                    | 2.2                |
| <i>Gd</i> |                    | 2.3                |
| <i>Dy</i> |                    | 2.4                |
| <i>Er</i> |                    | 3.2                |
| <i>Yb</i> |                    | 2.4                |
| <i>Lu</i> |                    | 3.2                |
| <i>Tl</i> |                    | 3.0                |
| <i>Pb</i> |                    | 3.4                |
| <i>U</i>  |                    | 3.4                |

## S.7 Temporal trends

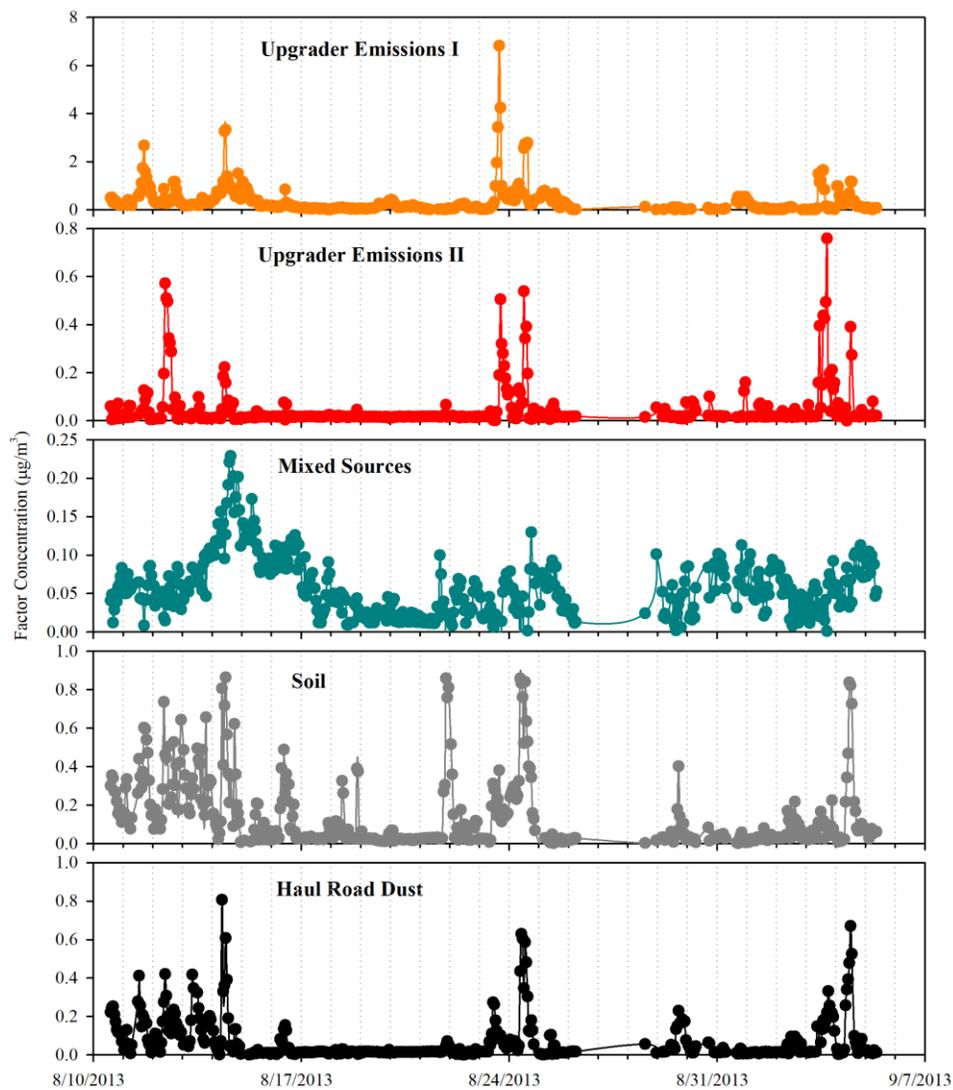
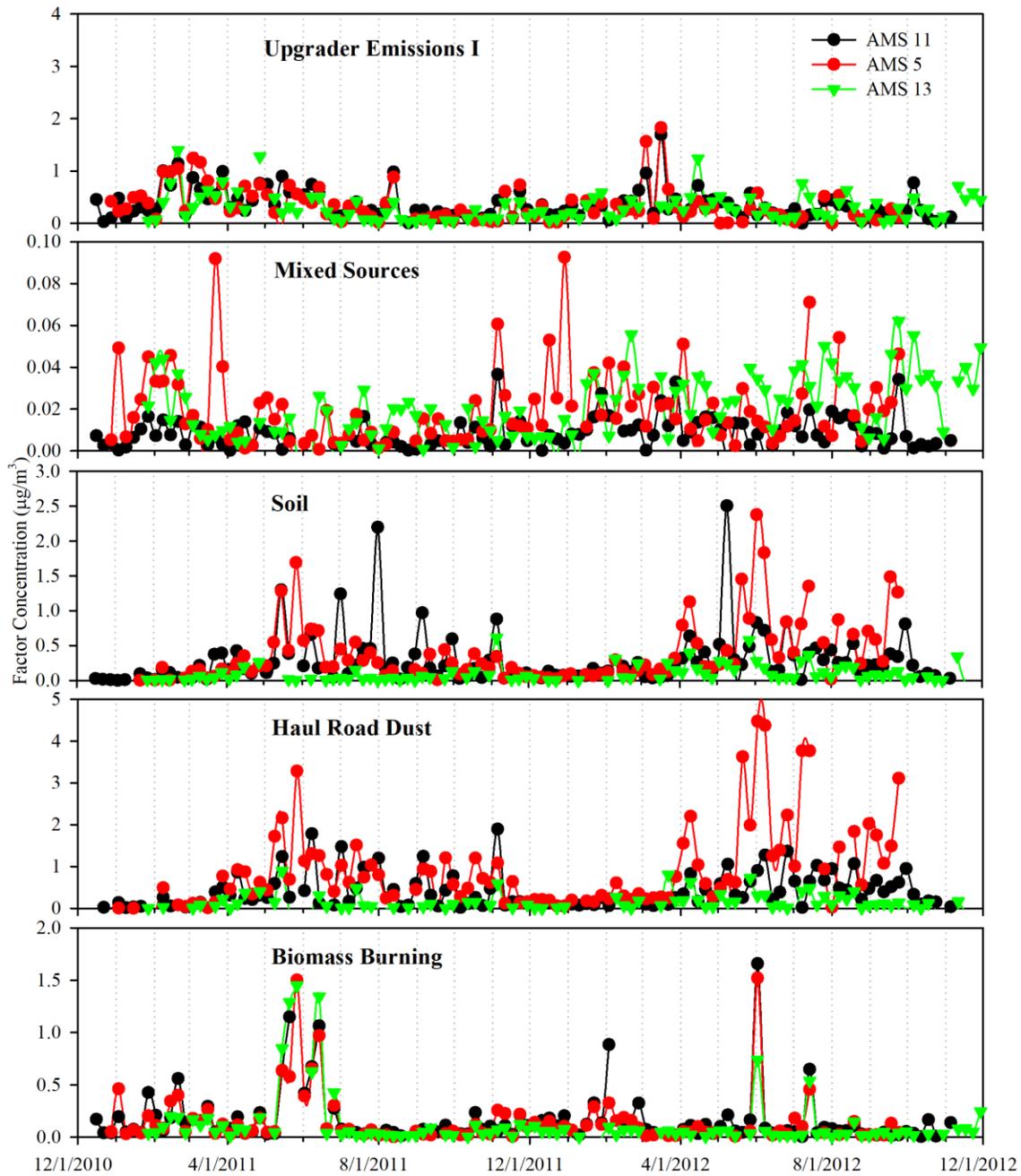


Figure S9. Concentration time series of PMF-resolved 5 factors during the intensive campaign.



**Figure S10.** Concentration time series of PMF-resolved 5 factors during the long-term campaign from December, 2010 to November, 2012 for AMS13, AMS5, and AMS11.

## S.8 Seasonal and Geographical Wind Directions

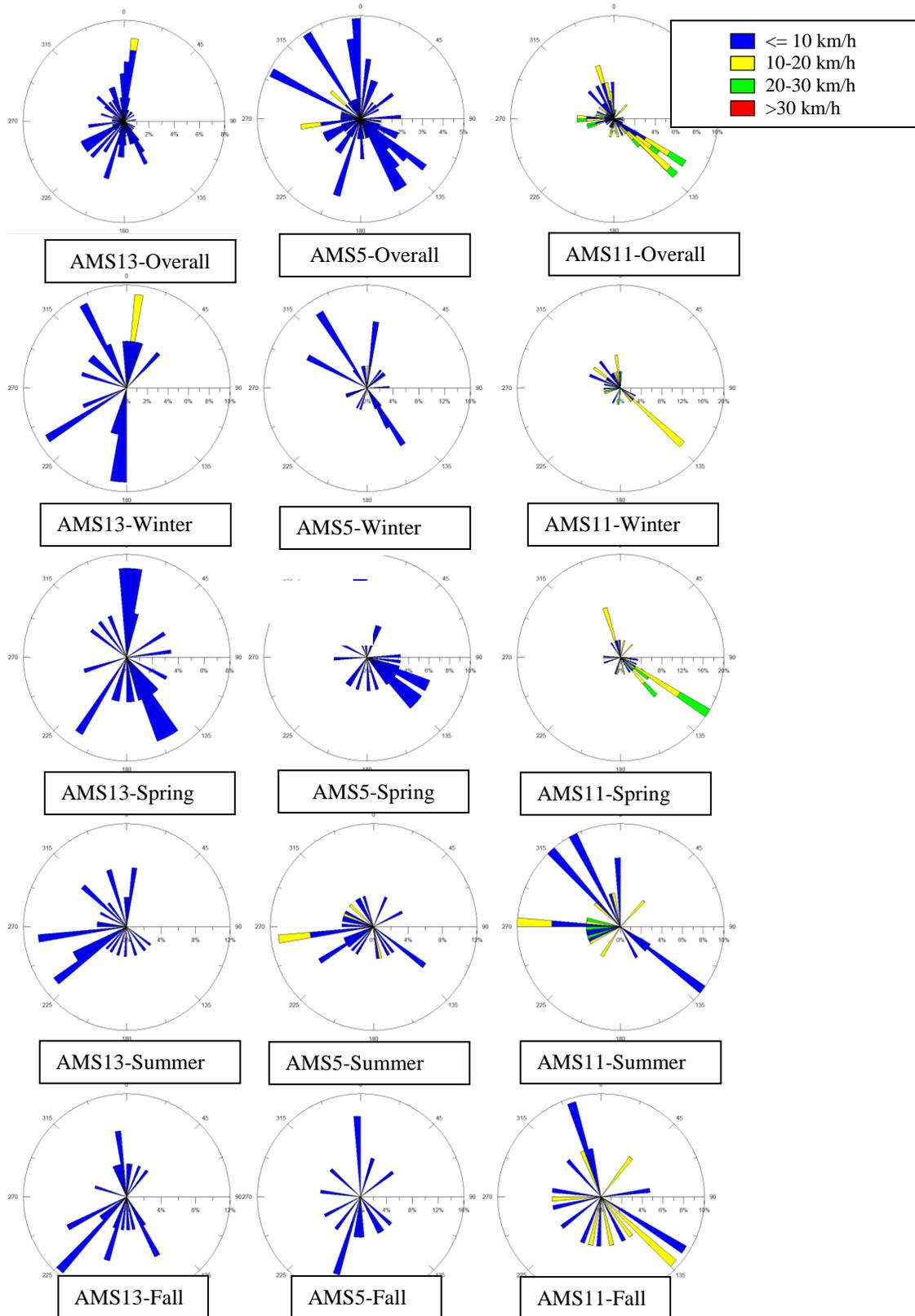


Figure S11. Overall and seasonal wind roses of the three sites analyzed in the long-term campaign.

## S.9 Seasonal Temperatures

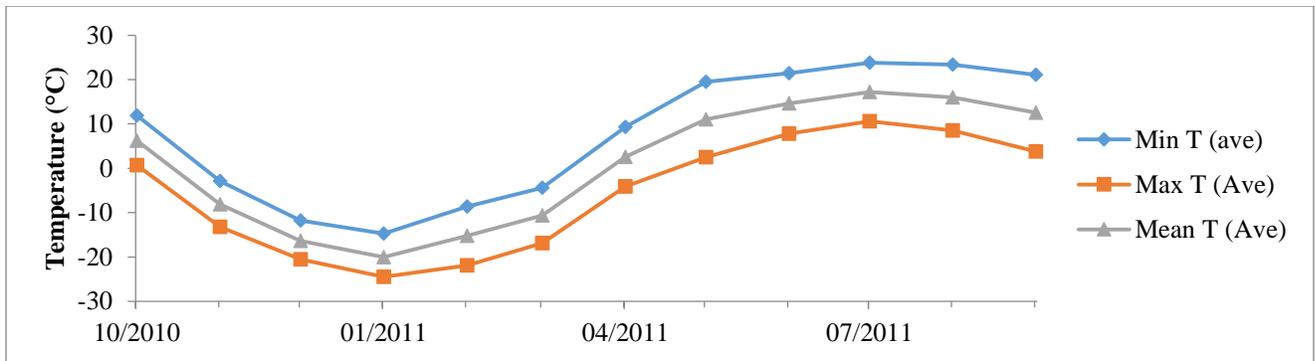


Figure S12. Typical average monthly temperatures for Fort McMurray, AB for the year 2010-2011.

## S.10 Long-Term Campaign CPF Plots

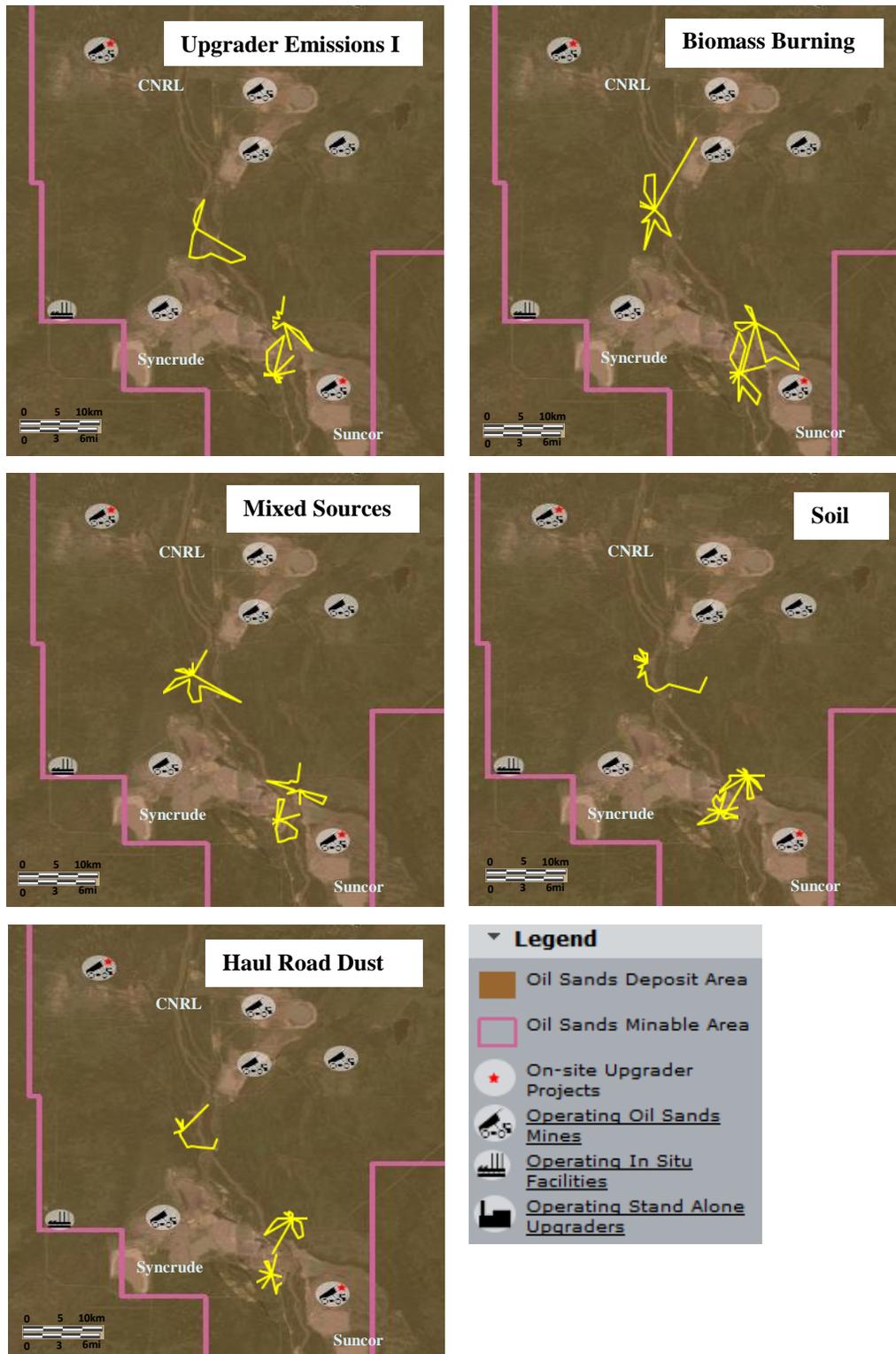
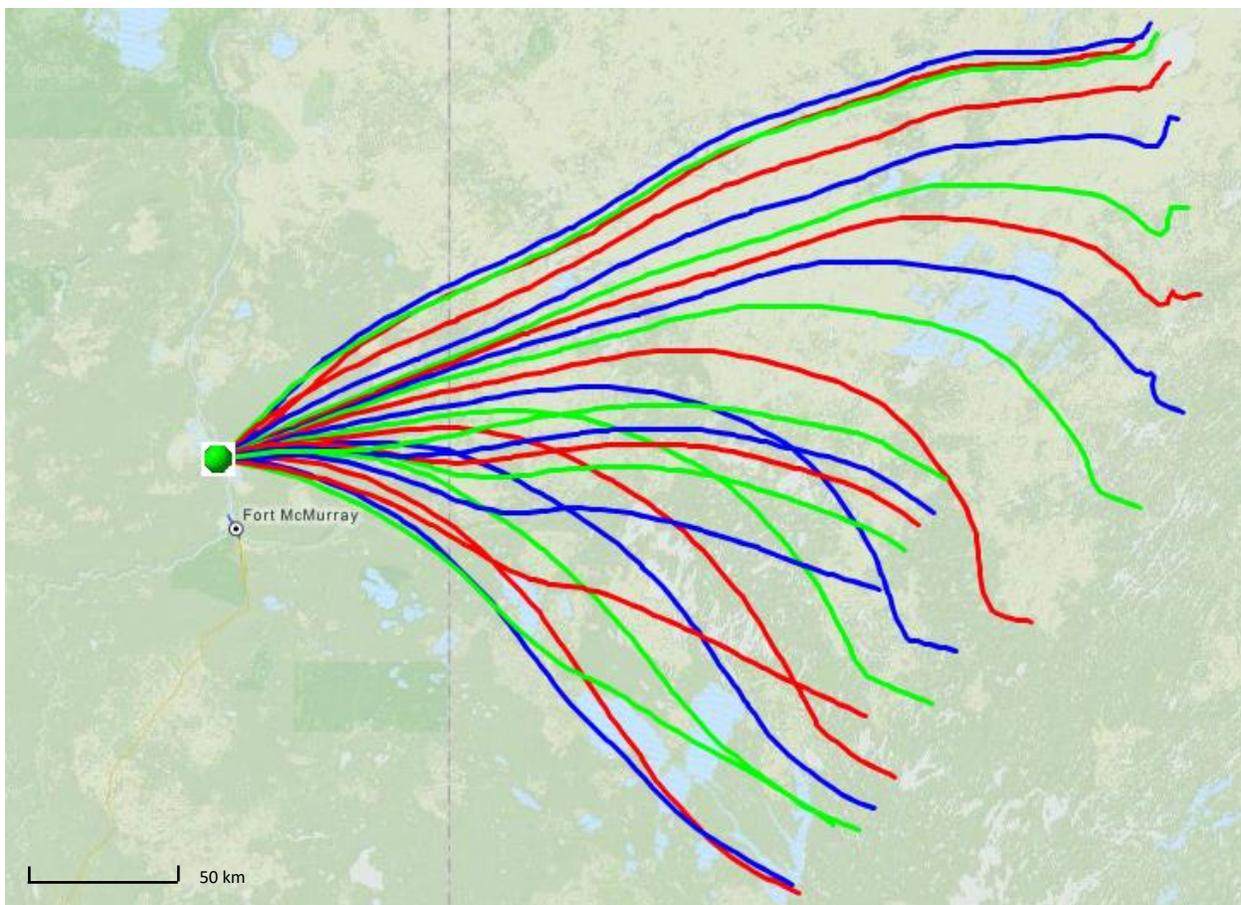
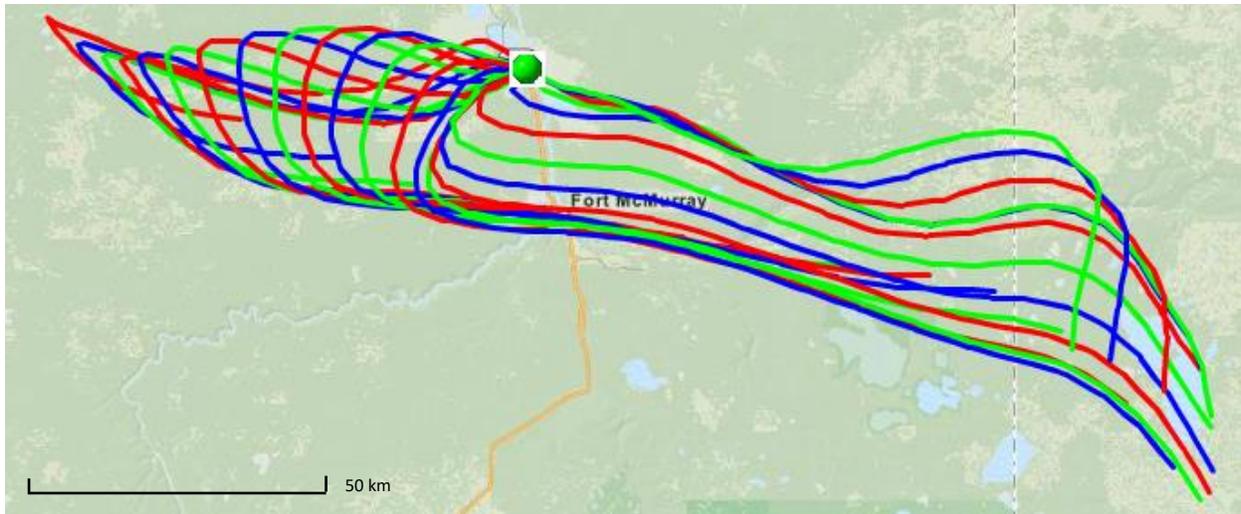
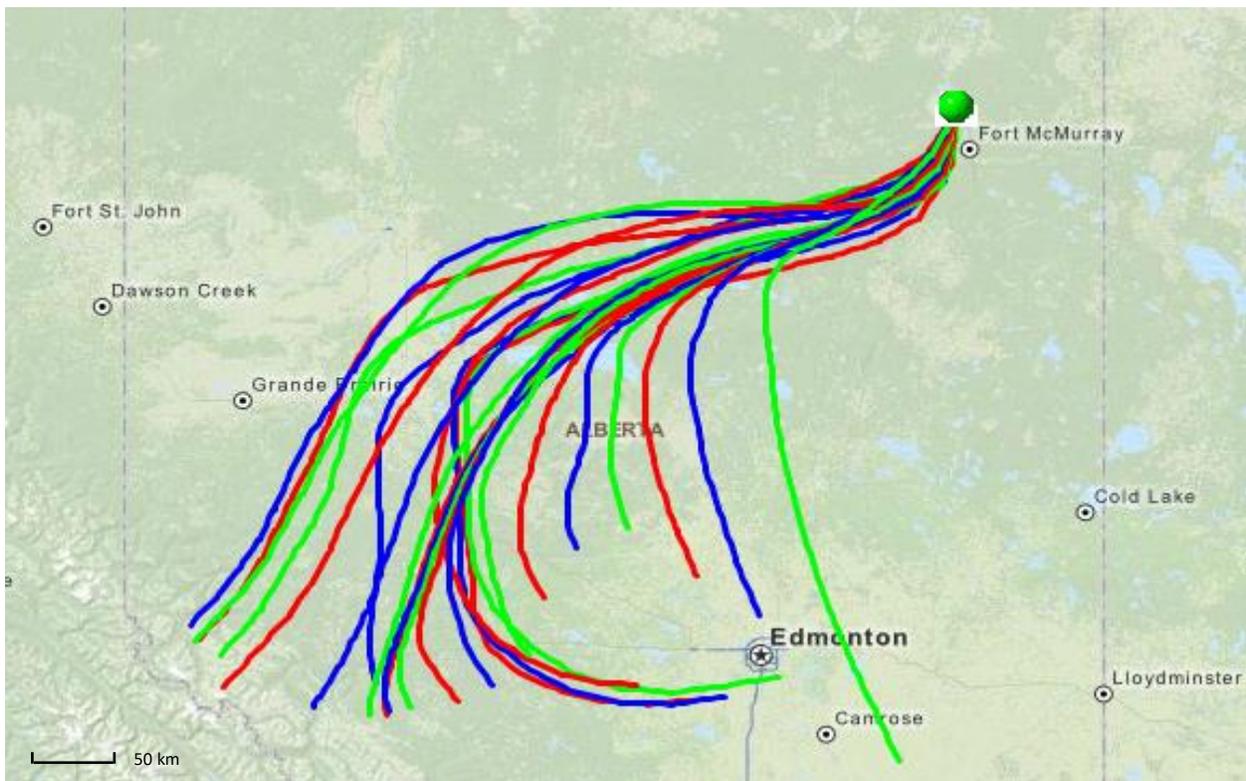
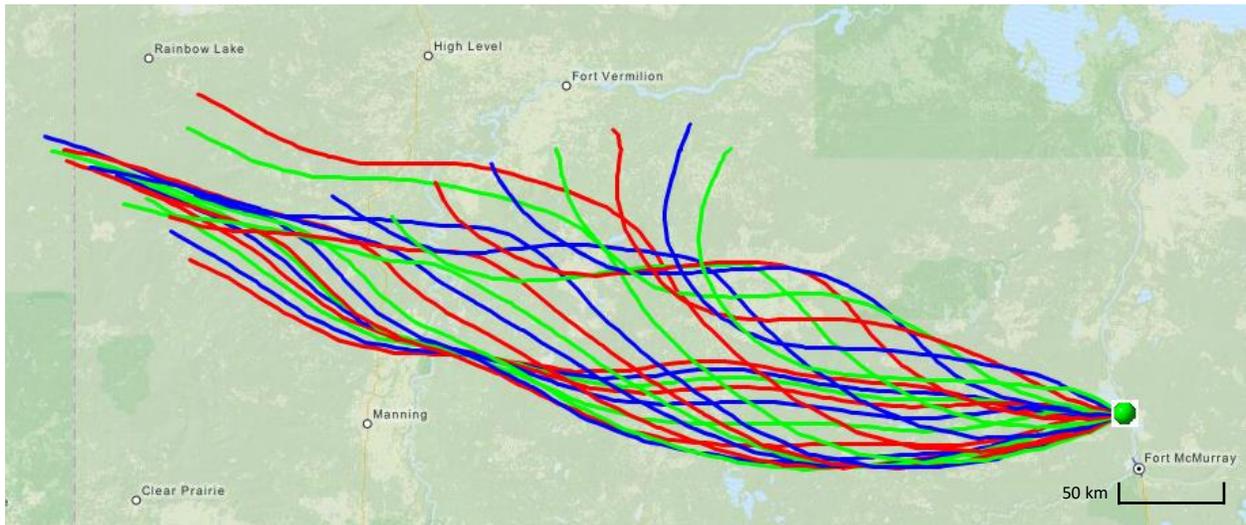


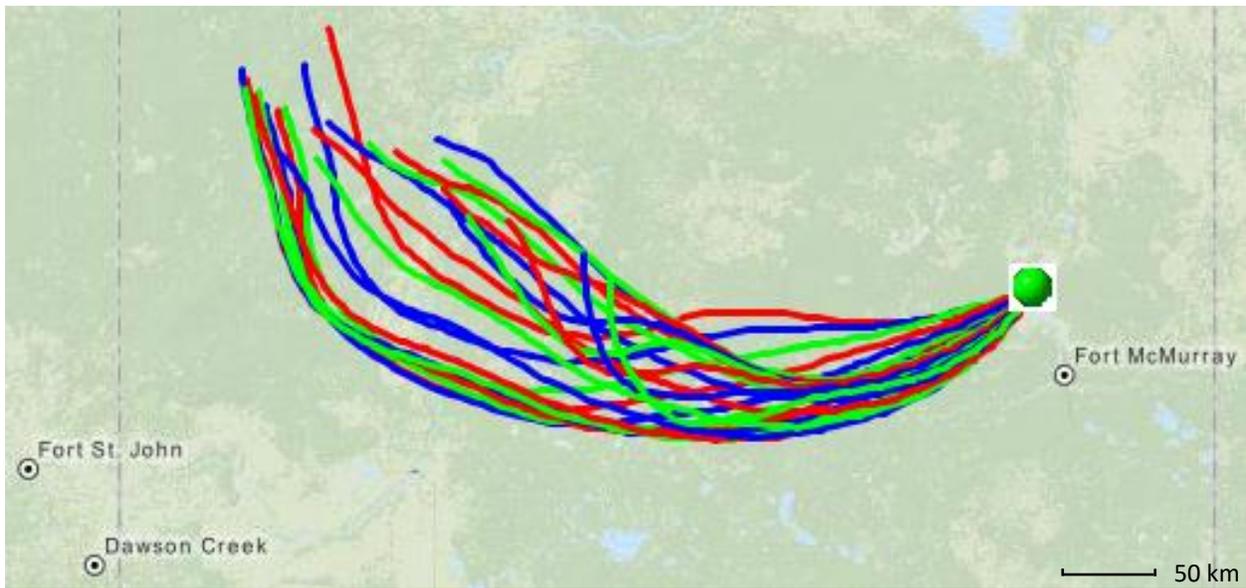
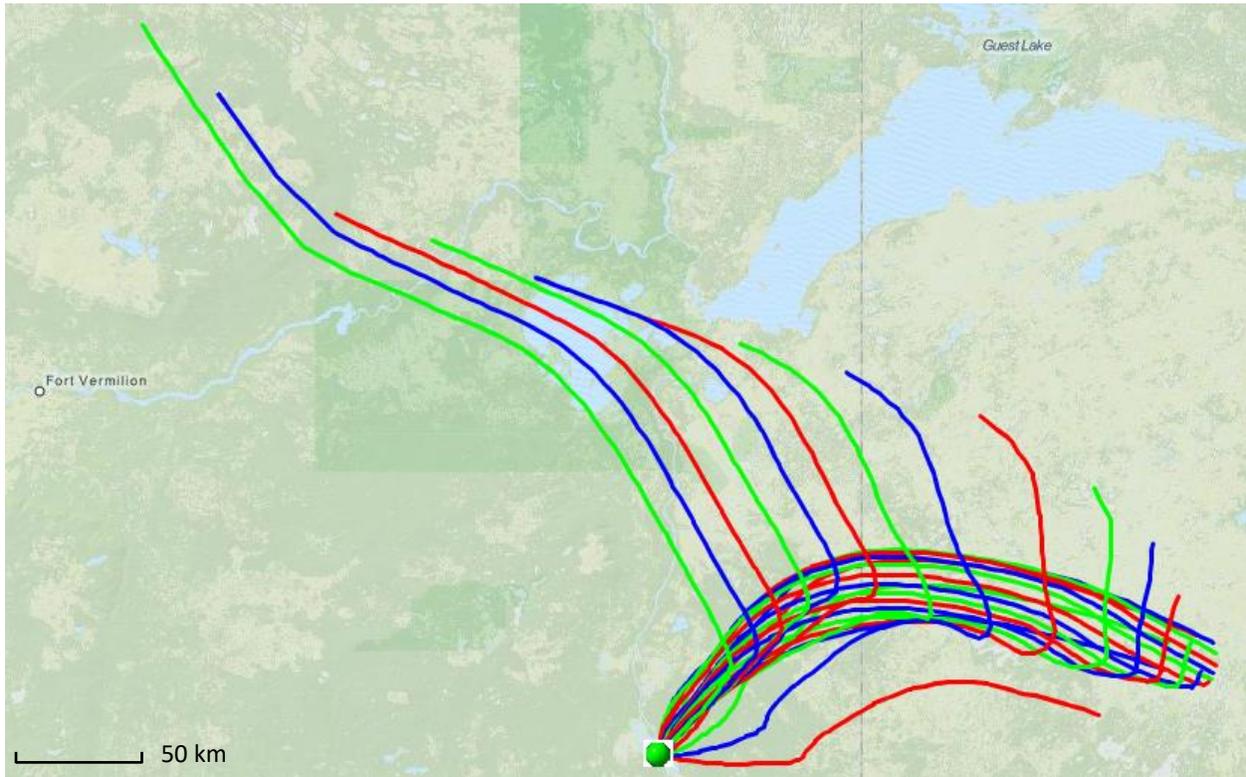
Figure S13. CPF plots from all three sites; AMS13, AMS5, and AMS11 for each factor identified using PMF on the long-term campaign's data. Blue Circles indicate the location of potential sources. 'Gaps' in the yellow wind rose are due to a lack of wind data coming from certain directions. Map courtesy of Alberta: Environmental and Sustainable Resource Development. Available: <http://osip.alberta.ca/map>

## S.11 HYSPLIT Analysis





**Figure S14. HYSPLIT Analysis diagrams of the 4 days with the overall highest contributions of the biomass burning factor: June 2, 2011; June 14, 2011; May 27, 2011; May 21, 2011.**



**Figure S15. HYSPLIT analysis diagrams of the 2 days with the highest overall soil contributions: June 8, 2011; March 22, 2012.**

## References

- Attaran, M.: Industrial diversity and economic performance in U.S. areas, *Ann. Regional Sci.*, 20, 44-54, 1986.
- Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, *Sci. Total Environ.*, 518, 626-635, 2015.
- Environmental Technology Verification Report: Cooper Environmental Services LLC Xact 625 Particulate Metals Monitor, Portland, OR, USEPA, 2012
- Environment Canada, NAPS Data Products, <http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx>, last access Jan. 2015.
- Government of Canada, 2010-2011 Climate, [http://climate.weather.gc.ca/climateData/dailydata\\_e.html?timeframe=2&Prov=AB&StationID=31288&dlyRange=2008-07-03|2011-10-20&cmdB2=Go&cmdB1=Go&Year=2010&Month=11&Day=5](http://climate.weather.gc.ca/climateData/dailydata_e.html?timeframe=2&Prov=AB&StationID=31288&dlyRange=2008-07-03|2011-10-20&cmdB2=Go&cmdB1=Go&Year=2010&Month=11&Day=5), last access Jan. 3, 2015.
- Healy, R. M., Riemer, N., Wenger, J. C., Murphy, M., West, M., Poulain, L., Wiedensohler, A., O'Connor, I. P., McGillicuddy, E., Sodeau, J. R., and Evans, G. J.: Single particle diversity and mixing state measurements, *Atmos. Chem. Phys.*, 14, 6289-6299, doi:10.5194/acp-14-6289-2014, 2014.
- Markovic, M. Z., VandenBoer, T. C., and Murphy, J. G.: Characterization and optimization of an online system for the simultaneous measurement of atmospheric water-soluble constituents in the gas and particle phases, *J. Environ. Monitor.*, 14, 1872-1884, 2012.
- Norris, G. and Duvall, R.: EPA Positive Matrix Factorization (PMF) 5.0: Fundamentals and User Guide, Petaluma, CA, United States of America Environmental Protection Agency, 2014.
- Paatero, P.: User's Guide for Positive Matrix Factorization Programs, PMF2.EXE and PMF3.EXE. Helsinki: University of Helsinki, 1996.
- Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, *Atmos. Meas. Tech.*, 7, 781-797, doi:10.5194/amt-7-781-2014, 2014.
- Paatero, P. and Hopke, P.: Utilizing wind direction and wind speed as independent variables in multilinear receptor modeling studies, *Chemometr. Intell. Lab.*, 60, 25-41, 2002.
- Paatero, P. and Hopke, P.: Discarding of downweighting high-noise variables in factor analytic models, *Anal. Chim. Acta*, 490, 277-289, 2003.
- Paatero, P., Hopke, P., Begum, B., and Biswas, S.: A graphical diagnostic method for assessing the rotation in factor analytical models of atmospheric pollution, *Atmos. Environ.*, 39, 193-201, 2005.
- Paatero, P., Hopke, P., Song, X.-H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, *Chemometr. Intell. Lab.*, 60, 253-264, 2002.

Reimer, N. and West, M.: Quantifying aerosol mixing state with entropy and diversity measures, *Atmos. Chem. Phys.*, 13, 11423-11439, 2013.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., and Feldman, M. W.: Genetic structure of human populations, *Science*, 298, 2381-2385, 2002.

Whittaker, R.: Dominance and diversity in land plant communities: numerical relations of species express the importance of competition in community function and evolution, *Science*, 137, 250-260, 1965.

Willis, M. D., Lee, A. K. Y., Onasch, T. B., Fortner, E. C., Williams, L. R., Lambe, A. T., Worsnop, D. R., and Abbatt, J. P. D.: Collection efficiency of the soot-particle aerosol mass spectrometer (SP-AMS) for internally mixed particulate black carbon, *Atmos. Meas. Tech.*, 7, 4507-4516, doi:10.5194/amt-7-4507-2014, 2014.

Xie, Y.-L., Hopke, P. K., Paatero, P., Barrie, L. A., and Li, S.-M.: Identification of source nature and seasonal variations of arctic aerosol by the Multilinear Engine, *Atmos. Environ.*, 33, 2549-2562, 1999.