SOURCE APPORTIONMENT OF SPOKANE FINE FRACTION AIR POLLUTION USING THE SPOKANE HEALTH EFFECTS DATABASE AND POSITIVE MATRIX FACTORIZATION

By

JENNIFER LYNN HEHL SHALTANIS

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Civil and Environmental Engineering

DECEMBER 2006

To the faculty of Washington State University:

The members of the committee appointed to examine the dissertation of JENNIFER LYNN HEHL SHALTANIS find it satisfactory and recommend that it be accepted.

_____
Chair

_____

_____

_____

ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge those who have mentored and supported me during the completion of this thesis.

This work detailed in chapter 2 is based on analysis from XRF trace elements, ionic, and carbon information. I would like to thank Drs. Dennis Finn and Brian Rumburg, and Mr. Mark Hoffman, for their expertise in measurement methods and educating me about the collection and analyses processes. Additionally, I would like to thank Drs. Astrid Schreuder and Eugene Kim for their communication regarding preliminary research using part of the data.

Research for chapter 3 entailed the addition of INAA trace element data, along with the XRF, carbon, and ion data, which led to a previously unavailable, more complete look at the Spokane $PM_{2.5}$ source profile. For his INAA expertise and filter measurements, I would like to thank Mr. Dan Dugan. His contributions were key to understanding neutron activation information and ultimately determining accurate PMF source features with this data.

The work presented in chapter 4 pertains primarily to elaborating on detailed carbon information, one more layer of new Spokane $PM_{2.5}$ source information. For help with this data, I would like to once again thank Dr. Dennis Finn for his tireless work on

the carbon measurements. Not only were his contributions to the database valuable, but his help in resolving instrumentation differences was also much appreciated.

Just as important as those who supported me in particular aspects of this thesis research, I would like to thank those who have supported me throughout my graduate experience. My doctoral committee: Drs. Candis Claiborn, Tim Larson, Hal Westberg, and Brian Lamb, have mentored, motivated, and educated me with their unique strengths and wisdom. My peers – Drs. Irra Sundram, Laura Wendling, and Jorge Jiminez, Ranil Dhammapala, and Maya Place, have been my greatest mentors and many times my life preservers. Most of all I would like to acknowledge my husband, Major Dan Shaltanis, who has given me unfettering support and love, keeping me on course throughout this experience.

ATTRIBUTION

This thesis is comprised of five chapters. Chapter One is an overview of the research topic, foundational dataset used, and the various instruments and methods used to analyze data. Chapters Two-Four are manuscripts to be submitted for publication. Chapter Two is authored by me, Candis Claiborn, Tim Larson, and Dennis Finn; Chapters Three and Four are authored by me, Candis Claiborn, Tim Larson, Brian Lamb, and Hal Westberg. Chapter Five summarizes the conclusions from this research and possible future work on this research.

Chapter Two is a manuscript submitted for publication, entitled, "Determining Sources of Particulate Matter Air Pollution in Spokane, Washington, with composite XRF Data and Positive Matrix Factorization" This research investigated the sources of air pollution identifiable by positive matrix factorization using X-Ray fluorescence data. Two XRF instruments were used collectively to analyze the entire period of study, and the pairing of two machines created opportunities to follow specific elements, based on instrumental sensitivity. Drs. Dennis Finn and Astrid Schreuder assisted me in the processing and understanding how best to use this data, given their instrumental and statistical knowledge relevant to this research.

Chapter Three is a manuscript submitted for publication, entitled, "PMF-derived $PM_{2.5}$ Sources in Spokane, Washington, with emphasis on Trace Metals Analysis" Instrumental Neutron Activation Analysis was added to supplement the previous research

involving XRF. INAA data were modeled with PMF and compared to the results found using XRF. INAA and XRF uniquely identify some elements in PM mass, and also share common elements. Dr. Dennis Finn and Mr. Dan Dugan assisted me in processing the INAA data, Dr. Finn through his experience with the XRF instrumentation, and Mr. Dugan with his experience with the INAA method.

Chapter Four is a manuscript submitted for publication, entitled, “Sources of $PM_{2.5}$ modeled by EPA PMF version 1.1 using Trace Elements and Temperature-resolved Carbon” Original carbon data are divided into organic fractions and elemental carbon for further details about Spokane particulate pollutions sources. These carbon data are united with previous XRF and INAA data to create a finalized analysis of sources, comparing the increased level of information to that of total carbon. Dr. Dennis Finn was especially helpful with his knowledge of both carbon methods, as well as his overall knowledge of instrumental corrections and limitations.

SOURCE APPORTIONMENT OF SPOKANE FINE FRACTION AIR POLLUTION USING THE SPOKANE HEALTH EFFECTS DATABASE AND POSITIVE MATRIX FACTORIZATION

Abstract

by Jennifer Lynn Hehl Shaltanis, Ph.D.
Washington State University
December 2006

Chair: Candis S. Claiborn

The Spokane PM database currently contains one of the largest continuous $PM_{2.5}$ daily mass databases available. Coverage includes, but is not limited to: total carbon, elemental and organic fraction carbon, particulate ions, and trace element information, from 1995-2002, taken from an area of Eastern Washington subjected to agricultural, industrial, urban, and rural pollution. Daily $PM_{2.5}$ mass samples were analyzed with several measurement methods: Kevex and Jordan Valley EDX-771 energy dispersive spectrometers, Instrumental Neutron Activation Analysis, for trace element species; total carbon by thermal manganese oxidation; total carbon, elemental and organic fractions by thermal/optical transmittance; sulfate and nitrate by ion chromatography; and ammonium by colorimetry. The data were modeled to identify pollution sources with the new Positive Matrix Factorization (PMF) model, version 1.1, released by the United States Environmental Protection Agency.

The PMF models revealed seven source features, similar to profiles identified in earlier work using various sets of the $PM_{2.5}$ data. PMF analyses using the XRF data yielded airborne soil, nitrate, biomass burning, chlorine-rich, metal processing, and vehicle combustion features. INAA data yielded the seventh feature, a Cr-rich feature, which was associated with elemental carbon. Bootstrapping and t test methods confirm the features' mass contributions to the overall ambient $PM_{2.5}$ concentrations. Airborne soil was associated with late summer dust events, with little to no carbon. Metal processing contains organic carbon, whereas the Cl-rich source contained elemental and light organic carbon. The Cl-rich source determined using Jordan Valley data is confused with the vehicle exhaust and biomass burning sources, due in part to the lack of a calibration standard for the instrument. These source features were tied to their specific industrial sources, which can vary sporadically throughout the study period. Nitrate is associated with domestic burning and increases in winter, with organic carbon components. Biomass burning occurred in late fall or early spring, when fields were burned for future growing seasons, and shares similar feature species with vehicle exhaust, which was present consistently throughout the entire study period. Both biomass burning and vehicle exhaust contain elemental and organic carbon; the exhaust exhibited heavier organic compounds, and biomass burning, mostly lighter compounds.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

CHAPTER 2: Determining Sources of Particulate Matter Air Pollution in Spokane, Washington, with composite XRF Data and Positive Matrix Factorization

LIST OF ABBREVIATIONS

Ag	silver
Al	aluminum
Br	bromine
C	carbon
Ca	calcium
Cd	cadmium
Ce	cerium
Cl	chlorine
Cr	chromium
Cs	cesium
Cu	copper
Co	cobalt
CMB	chemical mass balance
E	error matrix
E	error
EC	elemental carbon
EDXRF	energy dispersive X-ray fluorescence
EM	Error Model
Eu	europium
F	source profile matrix
Fe	iron
G	source contribution matrix

Hf	hafnium
Hg	mercury
IGERT	Integrated Graduate Education and Research Training
INAA	Instrumental Neutron Activation Analysis
JV	Jordan Valley
K	potassium
LAR	Laboratory for Atmospheric Research
MDL	minimum limit of detection
ME	multilinear engine
Mn	manganese
Mo	molybdenum
Na	sodium
ng/m^3	nanograms per cubic meter
NAAQS	National Ambient Air Quality Standards
NRC	Nuclear Radiation Center
NH_4^+	ammonium ion
NO_3^-	nitrate ion
OC	organic carbon
OC_1	organic carbon, lightest mass group
OC_2	organic carbon, second lightest mass group
OC_3	organic carbon, third lightest mass group
OC_4	organic carbon, fourth lightest mass group
OC_5	organic carbon, heaviest mass group

Pb	lead
Pd	palladium
PM	particulate matter
$PM_{2.5}$	particulate matter with aerodynamic diameter of 2.5 microns or smaller
PM_{8}	particulate matter with aerodynamic diameter of 8 microns or smaller
PM_{10}	particulate matter with aerodynamic diameter of 10 microns or smaller
PMF	positive matrix factorization
PMF2	two-dimensional PMF
PCA	principal component analysis
Q(E)	object function
REU	Research Experience for Undergraduates
Rh	rhodium
R^2	correlation coefficient
S	sulfur
Sb	antimony
Sc	scandium
SCAPCA	Spokane County Air Pollution Control Authority
Se	selenium
SO_2	sulfuric gas
SO_4^{2-}	sulfate ion
Si	silicon
Sm	samarium
Sn	tin

σ	sigma, standard deviation
Ta	tantalum
TC	total carbon
TEOM	tapered element oscillating microbalance
Th	thorium
Ti	titanium
TMO	Thermal Manganese Oxidation
TOT	Thermal Optical Transmittance
TOR	Thermal Optical Reflectance
U	Uranium
USEPA	United States Environmental Protection Agency
UW	University of Washington
$\mu g/m^3$	micrograms per cubic meter
V	vanadium
VAPS	Versatile Air Pollutant Sampler
WSU	Washington State University
$\mathbf{X}$	specie matrix
X	specie
χ^2	chi-square value, measure of normality in a data sample
XRF	X-ray fluorescence
Yb	ytterbium
Zn	zinc
Zr	zirconium

Dedication

This dissertation is dedicated to my husband and parents, who have given me the support to do this research, and to my children, who give me cause to take up environmental work.

CHAPTER 1: INTRODUCTION TO THE SPOKANE DATA PMF MODELS

1.1 Introduction

Air pollution can harm health and welfare of all life on the earth. Minimizing the exposure of such contaminants to living beings is a necessary task, and to do this task, scientists need to understand pollutant origins. The United States Environmental Protection Agency (US EPA) recognizes many pollutants as having significant health effects, six of these are identified as the National Ambient Air Quality Standards (NAAQS) criteria pollutants. One of the criteria pollutants, particulate matter (PM), has been linked to respiratory distress, inflammation of pre-existing respiratory illnesses, and in some cases mortality (Dockery et al., 1993; Norris et al., 2000). The specific exposure threats of PM on human health are not clearly understood because of the complicated relationship between physiology, human activity, and aerosol behavior (Phalen, 1998). For regulatory purposes, the US EPA sets primary standards for PM and other criteria pollutants, limits set to protect human health and welfare for the most sensitive populations. PM can be divided into PM_{10} and $PM_{2.5}$, particles of 10 microns or less and particles of 2.5 microns or less, respectively. PM_{10}, also known as coarse fraction PM, has a 24-hour primary standard of 150 $\mu g/m^3$, whereas $PM_{2.5}$ has both a 24-hour primary standard of 35 $\mu g/m^3$ and annual primary standard of 15 $\mu g/m^3$. PM is the only one of all criteria pollutants which is heterogeneous in nature, and therefore it is difficult to characterize. From this point on, unless otherwise specified, PM refers to both size fractions.

PM_{10} and $PM_{2.5}$ have differing levels of penetration into the respiratory tract. Most particles up to ten microns in diameter can easily enter the human nasal cavity. The larger of these particles can be trapped by hair and mucus membranes in the nose; the particles are effectively immobilized and do not pose a threat to health. The smaller of these particles can penetrate such barriers and travel as far into the respiratory system as the lower lungs; recent studies suggest that these intervals may not be the most representative of the particle sizes that infiltrate the respiratory system and are under review (Schwartz et al., 1999; Norris et al., 2000).

Eastern Washington is susceptible to violating the EPA primary standard for particulate matter and the region has been out of compliance more than once (Villasenor et. al, 2001). Contributing to these violations are weather/climate induced conditions and anthropogenic activities (Claiborn et. al, 2000). The first of these situations is the high wind events following dry weather that Eastern Washington can experience. The lack of precipitation coupled with strong winds can generate windblown dust from agricultural fields. These particles can remain suspended in the air for long periods; dust loading dramatically increases ambient PM concentrations. Another condition is cooler weather stagnation of the air. During fall and winter, the air in this region can become stratified, creating inversions. These inversions, along with light winds, cause the air to become stagnant, causing PM to become locally concentrated. $PM_{2.5}$, or fine fraction PM, concentrations can increase noticeably during these periods, in part because of the increased use of wood burning stoves during the cold months. Another condition leading to high PM levels occurs during post-harvest field burns. Fine PM of an organic carbonaceous nature can increase as a result of the burning of crop stubble and grass.

To address the unique pollution and health concerns of Eastern Washington, a group of air quality and health researchers launched a massive field campaign to collect PM, meteorological, and health data. Larger and more comprehensive than any of it contemporaries, this research was a collaborative effort among Washington State University, University of Washington, Spokane County Air Pollution Control Authority (SCAPCA), the American Lung Association of Washington, the Mickey Leland Nation Urban Air Toxics Research Center, and the Harvard School of Public Health. The purposes of the study were to collect health and air pollution data during an extended period in the area, analyze the data for connections between health and pollution, and to potentially identify specific links between pollutants and human health effects.

The overarching goal of this collaborative research is extensive, both in terms of understanding the nature of air pollution and its detrimental effects on human health. Several tenant goals have grown out this blanket objective, including the unique opportunities to characterize long-term shifts in PM pollution, the myriad effects multiple instruments can have on data measurements, and the vast elemental coverage produced by multiple measurement techniques. From an air quality research view point, this database is a goldmine of information, more complete than any database before it, and can lead to advanced PM source knowledge. The purpose of this thesis is to identify $PM_{2.5}$ sources in Spokane, Washington, for the entire study period, based on composite XRF, carbon, INAA, and ionic measurements. The results of this thesis will serve to further explain the air quality levels due to $PM_{2.5}$ in Eastern Washington.

1.2 Data Collection and Quality

Air quality and meteorological information were taken at a site in north Spokane, from January 1, 1995 – May 15, 2002. A photo of the site and map of its location are found in **Figures 1** and **2**. The choice of parameters measured were EPA criteria pollutants and any environmental or weather conditions which may influence the nature of the criteria pollutants. The site where data collection was performed was chosen because of its downwind proximity to several established sources of air pollution. The instruments were primarily automated; however, technicians visited the site regularly for maintenance and quality assurance checks.

1.2.1 Data Collection and Analysis

Data were collected using a suite of instruments and methods. For this thesis, research is focused on $PM_{2.5}$ species and the instrumentation used for their measurements. The instrument capabilities were intended to synergistically build a composite picture of air pollution in Spokane, Washington, and are listed in **Tables 1** and **2.** Measurements include the ions: SO_4^{2-}, NH_4^+, NO_3^-; and carbon species, including total carbon, elemental carbon, and five organic fractions. **Table 3** shows the coverage of elements for the EPA Kevex and Jordan Valley Energy Dispersive X-Ray Fluorescence (EDXRF) instruments, as well as the Instrumental Neutron Activation Analysis (INAA) method. Instrumental Neutron Activation Analysis (INAA) and XRF overlap in the majority of their elemental coverage, but INAA uniquely identifies eight other elements (*Sm*, *Ce*, *Eu*, *Hf*, *Ta*, *Th*, *U*, and *Yb*). Included with the ion and carbon species, instruments account for an overall coverage of approximately 62 species.

1.2.2 Quality Assurance

Quality assurance practices are essential for ensuring published data are reliable and the results based on that data are optimal. Several checks were performed to highlight and address any problematic data. The main tasks performed to improve the quality of the data were to fill in estimates for missing data and remove or replace questionable data from within the dataset. PMF can handle some problematic data, but the wide scope of analyses being done on this dataset requires the data to be in the best shape possible prior to modeling. Models were run on the different treatments of suspect data, to determine which protocol best serves the model and the results produced.

1.2.2.1 Missing Data

For particulate elemental carbon, organic carbon, nitrate, ammonium, and fine fraction mass sampled by the Versatile Air Pollutant Sampler (VAPS) or Tapered Element Oscillating Microbalance (TEOM), missing data were replaced with the geometric mean values; corresponding uncertainties were calculated as the standard deviation for the species concentration plus one-third of the limit of detection for the species or a root mean square of the analytical uncertainties. XRF and INAA elements reported either valid data or below detection, and their uncertainties were treated accordingly prior to entering the database. The missing value substitutions are as follows:

$$TC = EC + OC \tag{1a}$$

$$NO_3^- = 0.5 * MDL_{NO_3^-} \tag{1b}$$

$$SO_4^{2-} = \sum_{i=1}^{n} \frac{\left(SO_4^{2-}\right)_i}{n} \quad (1c)$$

$$NH_4^+ = 0.5 * MDL_{NH_4^+} \quad (1d)$$

$$VAPS_{2.5} = TEOM_{2.5} \quad (1e)$$

$$UNC_x = \sigma_x + \frac{1}{3} MDL_x \quad (1f)$$

$$UNC_{VAPS} = \sqrt{\left(\sigma^2_{sampling} + \sigma^2_{gravimetry}\right)} \quad (1g)$$

where *MDL* is the limit of detection for the specie, *TC* represents total carbon, *EC* is elemental carbon, *OC* is the organic carbon, NO_3^-, SO_4^{2-}, NH_4^+ are the nitrate, sulfate, and ammonium concentrations, respectively, $VAPS_{2.5}$ is the VAPS fine fraction concentration, $TEOM_{2.5}$ is the TEOM fine fraction concentration, UNC_x is the uncertainty calculation for the ions, and UNC_{VAPS} is the uncertainty for the VAPS mass. The mean of the sampling and gravimetric uncertainties used for VAPS is greater than that of the TEOM, so it was used in lieu of the reported TEOM uncertainty. All of these quantities were used to assess which samples would be kept in the dataset for PMF analysis and which cases were to be modified before use.

1.2.2.2 Below Detection Data

The limits of detection for carbon, ions, and trace elements were determined in several ways. The limits of detection for elemental analyses were determined differently for each method. Discussion of how XRF MDLs were derived is provided in Chapter 2, within the methods for PMF analyses of XRF information. INAA MDLs are detailed in Chapter 3, where PMF analyses for INAA and XRF are covered. Carbon and ion detection limits were provided by previous laboratory analyses; if two MDL values were determined by independent studies, the greater of the two numbers was used, in order to maximize the resulting calculated uncertainty. The MDL values for carbon are: TC_{MDL} is $2.26\ \mu g/m^3$; EC_{MDL} is $0.56\ \mu g/m^3$; and OC_{MDL} is $1.70\ \mu g/m^3$. MDL_{NO3} is set at $0.035\mu g/m^3$ and the MDL_{NH4} value used is $0.15\mu g/m^3$. No previously derived *MDL* value was given for SO_4^{2-}, so forty per cent of the geometric mean value of the dataset was used. The geometric mean was used instead of the average because its calculation for uncertainty produces the largest uncertainty value. These equations and substitutions are a composite approach based on methods suggested by Polissar et al. (1998, 2001) and previous model work on this dataset (private conversations with Dr. Eugene Kim, Clarkson University, and Dr. Astrid Schreuder, University of Washington PM Center).

1.2.2.3 Data Filtering

As Air quality data were collected or analyzed, any equipment or sampling issues were noted in a flag log file, and blanks were frequently tested for any systematic errors. These flags are published with the data so users know to use the information with caution. In addition to tracking filters during handling, the data were scrutinized post-

analysis to pull out any other suspect points. One test performed was a time series plot of each specie. These time series highlighted several points for many species as suspicious. The majority of these of the suspect points had corresponding flags, and were either thrown out, substituted, or their uncertainties increased.

A comprehensive $PM_{2.5}$ reconstruction was performed for the Kevex data period during the collection process, in order to ensure carbon, ionic and trace element information correctly apportioned $PM_{2.5}$ mass into its specie constituents. Details of this reconstruction are given in the Appendix, which show the instrumental methods reasonably accounted for fine fraction mass by species. Questionable samples accounted for approximately 4% of the five year period, and were either subjected to increased uncertainty or elimination. Primary causes for these unreliable samples were filter contamination, instrumental failure, and/or operator error. Several of these samples were taken from 1995; as the project matured, researchers gained working knowledge on how to prevent future problems.

1.3 Source Apportionment Modeling: PMF

Researchers have used several techniques to speciate PM into specific sources, in order to hypothesize about its origins. Each of these methods is individually important in identifying the composition of PM; unfortunately, each also has its limitations of analysis. Given the reality of current analysis techniques, scientists are faced with a sometimes challenging task of translating analysis data into tangible air pollution sources.

Principal component analysis (PCA) is a technique to assign pollution factors to the data. PCA methods rely on the patterns found within the data and do not assign

factors based on resemblance of data to chemical signatures, a method used by the traditional chemical mass balance (CMB) approach, such as the study done on this dataset by Hoffman (2002). Positive Matrix Factorization (PMF) is one successful PCA method of source apportionment. PMF does not require any pre-existing knowledge of the pollution sources, only the data itself and corresponding uncertainty. In essence, it does not recognize the information is air quality data and that the factors are supposed to reflect certain pollution fingerprints. Because a PMF model does not have internal nudging to direct the results, a user must make intelligent decisions about how the results represent pollution sources. PMF has successfully modeled several datasets, ranging from general air quality studies to particulate matter specifically (Kim et al., 2003; Lee et al., 1999; Song et al., 2001; Liu et al., 2003; Hien et al., 2004).

PMF is flexible in that it can handle information across a range of conditions. This model can also accept values very close to zero, which makes it favorable to use with instruments that have extremely high resolutions. The model does have a nonnegative constraint, which is suited to the receptor approach. Because the model concerns itself with pattern recognition and not strictly tied to concentrations, levels below detection for the instrument can still be incorporated into the model. The assigned uncertainty for each data point or the entire set can be altered as the user deems necessary. The preparatory work done on the data and the flexibility of the PMF model maximizes the amount of data one can analyze, while decreasing the emphasis on those artificially supplied values. PMF is beneficial in that the user is allowed to make decisions about data and work around the limitations of the measurement techniques.

Given these liberties, PMF is still constrained by how well the instruments capture pollutants or whether they detect them at all.

1.3.1 Governing Equations

The PMF model follows the general receptor modeling approach, which states that the contribution to pollution by *p* independent sources to all species in a given sample can be written as:

$$x_{ij} = \sum_{k-1}^{p} g_{ik} f_{kj} + e_{ij} \qquad (6)$$

where x_{ij} represents the concentration of the *j*th specie and *i*th sample, g_{ik} is the particle mass concentration for the *i*th sample due to the *k*th source, f_{kj} is mass fraction of the *j*th specie due to the *k*th source, e_{ij} is the residual for the ith sample of the *j*th specie, and *p* is the total number of sources. To evaluate these values with the model, the data should be arranged in matrices, with the following equation describing the overall calculations:

$$\mathbf{X} = \mathbf{GF} + \mathbf{E} \qquad (7)$$

where **X** represents data in a *n x m* matrix, **G** contains *n x p* source contributions, **F** is a *p x m* source profile matrix, and **E** is the *n x m* residual matrix. The array dimensions *n*, *m*, and *p*, are the number of measurements, elements, and sources, respectively. Further discussion of these equations can be found in PMF studies by: Henry (1987); Kim et al. (2003); and Paatero (1997, 2000).

The governing equations as given can produce an infinite number of solutions, and therefore the model results are not definite. This problem is common in receptor modeling, and has been addressed by Henry (1987, 2002); Hopke (1991); Paatero (1997, 2000); and Kim et al. (2003). The normal approach to deal with the non-uniqueness is to run PMF in such a manner as to minimize an object function, Q(E). The object function is derived from the uncertainties for the observations and is defined as:

$$Q(E) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{u_{ij}} \right]^2 \qquad (8)$$

where g_{ik}, f_{kj}, and x_{ij} are previously defined, and u_{ij} is the jth element uncertainty of the ith sample. This minimization, in concordance with the non-negativity constraint, is solved in an iterative manner using a weighted least squares calculation. The minimization scheme is only a guide for the user. Judgment of a best solution should be based on the results that produce the most realistic picture of pollution for the receptor site; prior knowledge of an emission inventory or pollution-producing activities is useful (Hopke, 1991; Paatero, 1997, 2000; and Kim et al., 2003).

A model solution is often not optimized in the first trial and there are several methods to refine to the algorithm. Model output of the datasets used in this study indicated slightly negative biases in the residuals so data were nudged by setting the Fpeak value to non-zero values and the model rerun. The Fpeak option will direct the

data either negatively or positively in response to a non-zero bias in the computed residuals (Paatero, 2000).

1.3.2 EPA PMF version 1.1

In 2005, EPA released an interactive form of the model, PMF version 1.1. The new version is favorable because it reduces the preparation time needed to initiate model runs, can perform several continuous runs, and has bootstrapping programs to assess uncertainties in model results. This version has a graphical user interface, or GUI, which guides the user through the selection of input controls for running the model. It has the capability of running several iterations with random starting points, which verifies the robustness of a model. Description of this version are available in the program manual, by Eberly (2005).

The new PMF version incorporates most of the traditional input parameters, adjusting model runs quickly and generating more conclusive output. Unlike the traditional PMF algorithms, the EPA method is based on the multilinear engine (ME), a multidimensional form of PMF, introduced by Paatero (1999). The premise is similar in concept: the model attempts to fit equations to measured data, by a system of factors with estimates and residuals, reducing error by minimizing the deviation of measurements from actual data, with consideration for uncertainties. The ME approach is more involved because it allows for greater dimensions of equation arrays. The base equations are as follows:

$$x_m = y_m + e_m = \sum_{k=1}^{Km} \prod_{n \in \Im mk} f_n + e_m \qquad (9)$$

where the subscript m indicates the number of equations to be solved, x_m refers to the data to be fitted to the linear regressions, y_m are fitted values, e_m are the residuals. The corresponding f_n values are a vector which represent known and unknown factor elements; this vector essentially corresponds to the scores and loadings in the traditional PMF (**G** and **F** values). Additionally *Km* corresponds to *p* in PMF, the number of factors produced. N_{mk} elements in J_{mk} sets represent n indexes of the *kth* product of the *mth* equation; these subscripts would indicate one product of one equation.

The minimization of Q is also more detailed to allow weights to be assigned to particular *m* equations. The Q values are based on how deviant the modeled values are from the actual masses, with consideration for model weights. A best fit approach to solving this equation is to minimize Q with respect to the vector *f*. *Equations 10*, *11a*, and *11b* detail how ME derives a minimal Q value, according to the following:

$$Q(x, f) = \sum_{m=1}^{M} w_m e_m^2 = \sum_{m=1}^{M} w_m \left(x_m - y_m\right)^2 \tag{10}$$

$$\hat{f} = \left[\hat{f}_1, ..., \hat{f}_n\right] = \arg\min_f Q(x, f) \tag{11a}$$

$$= \arg\min_f \sum_{m=1}^{M} \left(\frac{x_m - \sum_{k=1}^{Km} \prod_{n \in \Im mk} f_n}{\sigma_m} \right) \tag{11b}$$

The term w_m represents weights assigned in the m equations, and σ_m, uncertainties for each equation in the model, which are used to determine w_m.

PMF 1.1 can handle several formats of data, and can incorporate unique uncertainty constraints. Its greater ability to manipulate data decreases time between model runs because adjustments to concentration and uncertainty files can be made quickly. Its default settings are specifically designed to handle environmental data (Eberly, 2005). The model is set to run in robust mode, using the error model 12, with an allowance for a slightly negative concentration, another noteworthy difference from traditional PMF. Environmental data should be run in robust mode because its dynamic nature; the robust setting can adjust for sudden changes in variability (Kim et al., 2003). In running models with the previous version, error model 12 almost unanimously performed the best. EPA version 1.1 allows the user to choose one of several types of uncertainty. As with the original form, the EPA version allows for direct point-to-point uncertainties, which can be useful in treating specific samples. This version also has an equation form of the uncertainty, which assigns values based on concentration, detection limit, and some error (*Equation 12*). The first two inputs are values directly related to the data. The error can be a calculated value or the modeler's subjective measure of how well a specie is captured by instrumental analysis.

$$\sigma = \sqrt{\left(\frac{\%error}{100} \times concentration\right)^2 + MDL^2} \qquad (12)$$

Once the concentration and uncertainty files are read by the engine, a modeler has an additional opportunity to skew the model away from suspect species. The model allows for characterizing the strength of each specie as "strong," "weak," or "bad." "Weak" and "bad" species are downweighted and eliminated, respectively. Adding to all

of this pre-model setup, the modeler has one last chance to increase the uncertainty of the whole file, which may help the model to converge and produce a reasonable χ^2. This version of PMF has one limitation over the previous, script version. The Fpeak function is not included in possible manipulations, so any negative or positive skew in the data cannot be resolved with this traditional function; however, in running the previous version, small Fpeak values were used and were secondary in their influence, relative to other input parameters. Overall, the EPA version 1.1 of PMF is powerful next generation factor analysis tool which increases the efficiency and intelligence of source apportionment.

1.4 Thesis Research

Research for this thesis consists of PMF evaluation of the Spokane $PM_{2.5}$ data, and is separated into three main emphases: XRF, XRF and INAA trace elements, and detailed carbon information. Because this dataset is so large, both in terms of period length and species coverage, source profile details, in terms of species coverage and changes over time will be discussed, as well as instrumental uniquities that can be assessed. The results published in the following chapters details these source features in context of these issues, and provide a complete picture of particulate source features for Eastern Washington.

References

Claiborn, C.S., Finn, D., Larson, T.V., Koenig, J., 2000. Windblown dust contributes to high PM2.5 concentrations. Journal of Air and Waste Management Association 55, 1440-1445.

Dockery, D.W., Pope, C.A., Xu, X.P., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G.,

Speizer, F.E., 1993. An association between air-pollution and mortality in 6 United States cities. New England Journal of Medicine 329, 1753-1759.
Eberly, S. 2005. EPA PMF 1.1 User's Guide. US Environmental Protection Agency.
Henry, R.C., 1987. Current factor analysis models are ill-posed. Atmospheric Environment 21, 1815-1820.
Henry, R.C., 2002. Multivariate receptor models – current practices and future trends. Chemometrics and Intelligent Laboratory Systems 60, 43-48.
Hien, P.D., Bae, V.T., Thinh, N.T.H, 2004. PMF receptor modeling of fine and coarse PM10 in air masses governing monsoon conditions in Hanoi, northern Vietnam. Atmospheric Environment 38, 189-201.
Hoffman, M. (2002). Elemental analysis and receptor modeling of airborne particulate matter in Spokane, Washington, M.S. thesis, Washington State University, Pullman, WA.
Hopke, P.K., 1991. Receptor Modeling for Air Quality Management. Elsevier, Amsterdam, The Netherlands.
Kim, E., Larson, T.V., Hopke, P. K., Slaughter, C., Sheppard, E., Claiborn, C., 2003. Source identification of PM2.5 in an arid northwest U.S. city by positive matrix factorization. Atmospheric Research 66, 291-305.
Liu, W., Hopke, P.K., Han, Y-j, Yi, S-M, Holsen, T.M., Cybart, S., Kozlowski, K., Milligan, M. Application of receptor modeling to atmospheric constituents at Potsdam and Stockton, NY. Atmospheric Environment 37, 4997-5007.
Norris, G., T. Larson, J. Koenig, C. Claiborn, L. Sheppard, D. Finn, 2000. Asthma aggravation, combustion, and stagnant air. Thorax 55, 466-470.
Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and intelligent laboratory systems 37, 23-35.
Paatero, P., 1999. The multilinear engine – a table-driven, least squares program for solving multilinear problems, including the *n*-way parallel factor analysis model. Journal of Computational and Graphical Statistics 8:4, 854-888.
Paatero, P., 2000. User's guide for positive matrix factorizetion programs PMF2 and PMF3, Part I: tutorial.
Phalen, R.,F., 1998. Uncertainties relating to the health effects of particulate air pollution: The US EPA's particle standard. Toxicology Letters 96,97, 263-267.
Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: 1. Elemental composition and sources. Journal of Geophysical Research 103 (D15), 19045– 19057.
Polissar, A.V., Hopke, P.K., Poirot, R.L., 2001. Atmospheric aerosol over Vermont: chemical composition and sources. Environ. Sci. Technol. 35, 4604– 4621.
Schwartz, J.G., G. Norris, T. Larsen, L. Sheppard, C. Claiborn, J. Koenig, 1999. Episodes of high coarse particle concentrations are not associated with increased mortality. Environmental Health Perspectives 107, 339-342.
Song, X-H, Polissar, A. V., Hopke, P.K., 2001. Sources of fine particle composition in the northeastern US. Atmospheric Environment 35, 5277-5286.
Villasenor, R., Claiborn, C., Lamb, B., O'Neill, S., 2001. Mesoscale modeling of wintertime particulate matter episodes in eastern Washington, USA. Atmospheric Environment 35, 6479-6491.

Specie Measured	Instruments Used	Sampling Precision	Analytical Precision
PM_{10}	Tapered Element Oscillating Microbalance (TEOM)	2.80 $\mu g/m^3$	0.5 $\mu g/m^3$
PM_8	Versatile Air Pollutant Sampler (VAPS)	NA	~2 $\mu g/m^3$
$PM_{2.5}$	1. Versatile Air Pollutant Sampler (VAPS) 2. Tapered Element Oscillating Microbalance (TEOM)	1.2.80 $\mu g/m^3$ 2. NA	1.0.5 $\mu g/m^3$ 2.~2 $\mu g/m^3$
Organic Carbon	1. R & P Series 5400 Carbon Analyzer 2. Thermal Optical Transmittance (TOT) analyzer	NA	NA
Elemental Carbon	1. R & P Series 5400 Carbon Analyzer 2. Thermal Optical Transmittance (TOT) analyzer	NA	NA
NO_3^-	1. Dionex Model 4000I 2. Dionex Model DX-120	0.14 $\mu g/m^3$	0.05 $\mu g/m^3$
SO4	1. Dionex Model 4000I 2. Dionex Model DX-120	0.19 $\mu g/m^3$	0.1 $\mu g/m^3$
NH4	Technicon automated colorimeter	0.08 $\mu g/m^3$	
Trace elements	contained in VAPS filter samples	2.80 $\mu g/m^3$	0.5 $\mu g/m^3$

Table 1. Data collection Methods/Instrumentation.

Specie Analyzed	Instruments used	Analytical Precision
PM_8 Mass	Gravimetry: Cahn Microbalance Model C-33	0.5 $\mu g/m^3$
$PM_{2.5}$ Mass	Gravimetry: Cahn Microbalance Model C-33	0.5 $\mu g/m^3$
$PM_{2.5}$ and PM_8 trace elements	EDXRF: Kevex EDX-771 energy dispersive spectrometer	0.1 ng/ m^3
$PM_{2.5}$ and PM_8 trace elements	X-ray fluorescence: Jordan Valley model	0.1 ng/ m^3
$PM_{2.5}$ trace elements	Instrumental Neutron Activation Analysis: WSU Nuclear Radiation Center	0.001 ng/ m^3

Table 2. VAPS Filter Analysis Methods/Instrumentation.

Element	Kevex	JV	INAA	Element	Kevex	JV	INAA	Element	Kevex	JV	INAA
As	X		X	Sb		X	X	V		X	
K	X	X	X	Sc	X		X	Mn	X	X	
La	X		X	Se	X	X	X	Cu	X	X	
Na		X	X	Sr	X	X	X	Ga	X		
Sm			X	Ta			X	Ge	X		
W	X		X	Th			X	Y	X		
Au	X		X	U			X	Mo	X		
Br	X	X	X	Zn	X	X	X	Rh	X		
Ba	X	X	X	Zr	X	X	X	Pd	X		
Ce			X	Yb			X	Ag	X		
Co	X	X	X	Al	X	X		Cd	X	X	
Cr		X	X	Si	X	X		Sn	X		
Cs	X		X	P	X	X		Te	X		
Eu			X	S	X	X		I	X		
Fe	X	X	X	Cl	X	X		Hg	X		
Hf			X	Ca	X	X		Pb	X	X	
Ni	X	X	X	Ti	X	X		Mg	X	X	
Rb	X	X	X								

Table 3. Elemental coverage by the XRF and INAA methods.

Figure 1. Location of receptor site. VAPS are located on top of the building, with controls housed inside the facility. The site is surrounded by a large road, residential neighborhood, and a shopping district.

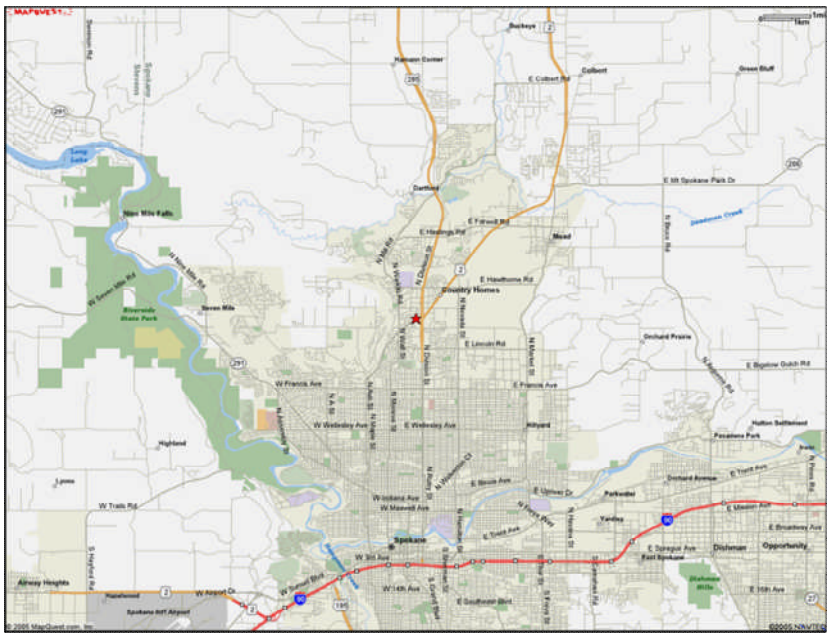


Figure 2. Map of Spokane and its surroundings. The industrial section of Spokane is east of the city center. Farmland is all around the urban section of the city. The interstate traverses east to west on the south side of the city and several major roads run north to south throughout the city. The red star denotes the receptor site location.

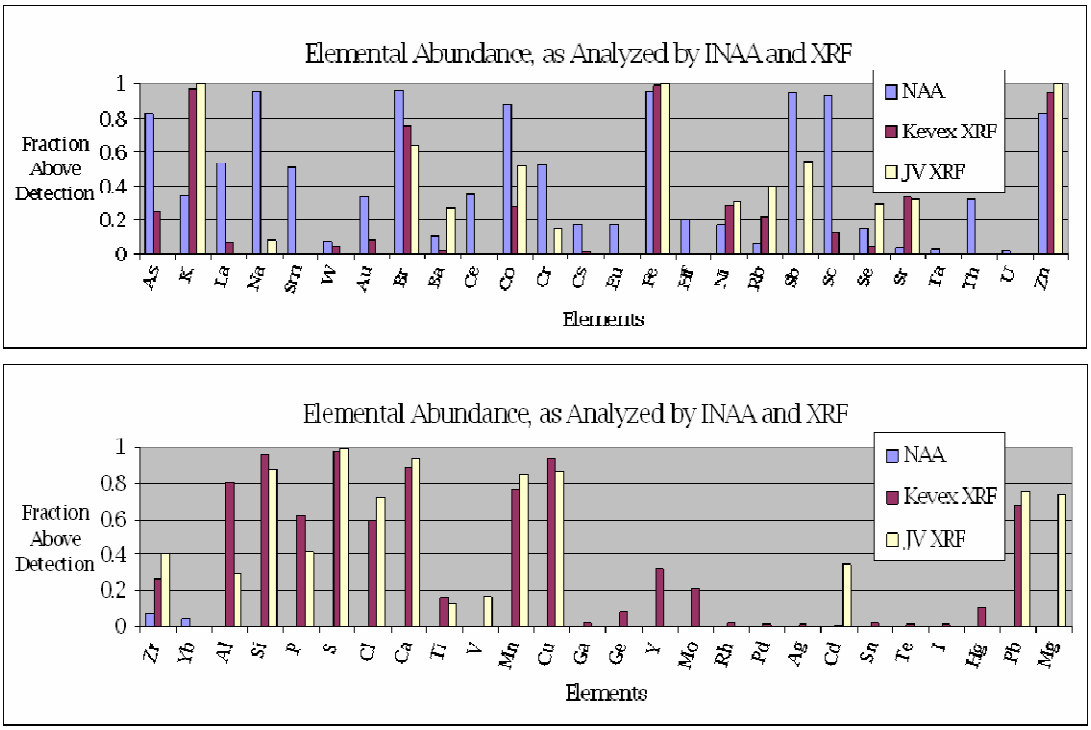


Figure 3. Abundance frequencies for XRF and INAA methods. The graphs indicate what fraction of samples contained mass above the minimum detection limit, per specie.

CHAPTER 2

Determining Sources of Particulate Matter Air Pollution in Spokane, Washington, with composite XRF Data and Positive Matrix Factorization

This chapter is presented as a paper that has been prepared for publication in a peer-reviewed journal.

$PM_{2.5}$ Source Features Analyzed by PMF 1.1, with Emphasis on Changes Due to a Long Study Period and Instrumental Shifts

Jennifer Shaltanis, Candis Claiborn, Dennis Finn
Laboratory for Atmospheric Research, Washington State University, Pullman, WA

Timothy Larson
Department of Civil Engineering, University of Washington, Seattle, WA

Correspondence:
Jennifer Shaltanis
Laboratory for Atmospheric Research
WSU/NSF IGERT Center for Multiphase Environmental Research
Department of Civil and Environmental Engineering
Washington State University
PO Box 642714
Pullman WA 99164-2714
Phone: (509) 335-6248
Fax: (509) 335-7632
Email: jshaltanis@wsu.edu

Abstract

The Spokane Health Effects Study currently contains one of the largest continuous $PM_{2.5}$ daily mass databases available. Coverage includes, among other things: total carbon, particulate ions, and trace element information from 1995-2002, in an area of Eastern Washington subjected to a plethora of agricultural, industrial, urban, and rural pollution. Daily $PM_{2.5}$ mass samples were analyzed with several measurement methods, including but not limited to: the Kevex and Jordan Valley EDX-771 energy dispersive spectrometers, for trace element species; total carbon by thermal manganese oxidation and Thermal/Optical transmittance; sulfate and nitrate by ion chromatography; and ammonium by colorimetry. The data were modeled to identify pollution sources with the new Positive Matrix Factorization (PMF) model, version 1.1, released by the United States Environmental Protection Agency. Kevex and Jordan Valley (JV) data were modeled individually to create both a continuous profile of air pollution as well as highlight the unique features that each instrument uncovered.

The PMF models revealed six sources, similar to a profile identified in earlier work using the first three years of data. Kevex and JV models yielded airborne soil, nitrate, biomass burning, chlorine-rich, metal processing, and vehicle combustion sources. Nitrate, biomass burning, and vehicle exhaust features yielded the largest mass contributions, collectively representing over three-fourths of the total source mass. In general, airborne soil, metal processing, and Cl-rich sources provided less than 10% each. However, short-term peaks produced individually high $PM_{2.5}$ masses. Mass contribution analyses by bootstrapping and statistical t test methods highlight key features in the instrumental transition as well as changes in pollution over time. Airborne soil was

seasonally associated to the post-harvest dry climate. Metal processing and Cl-rich sources were tied to their specific industrial sources, which can vary sporadically throughout the study period; nitrate was associated with domestic burning and generally increases in winter. Biomass burning occurred mostly in late fall or early spring, when fields were burned for future planting, but can share similar feature species with vehicle exhaust, which is present consistently throughout the entire study period.

Keywords: PMF, source apportionment, XRF, trace elements, $PM_{2.5}$

2.1 Introduction

Eastern Washington is susceptible to violating the EPA primary standard for particulate matter and the region has been out of compliance more than once (Villasenor et al., 2001). Contributing to these violations are weather/climate induced conditions and anthropogenic activities. High wind events accompanying dry summer weather can beget atmospheric dust loading and long periods of suspension. During cooler months, stratification can occur, leading to inversions and stagnation in the air; stagnation, coupled with wood-burning domestic heat, will concentrate $PM_{2.5}$ levels. Post-harvest field burns emit organic and mineral particles into the air (Haller et al., 1999; Kim et al., 2003). In addition to these unique Eastern Washington attributes are vehicle traffic and industrial processes in Spokane. Any of these scenarios can pose a potentially unsafe environment for human health.

In 1995, the Spokane Health Effects Study began to address the unique pollution and health concerns of Eastern Washington. This research was a collaborative effort

among the Mickey Leland Nation Urban Air Toxics Research Center, Washington State University, University of Washington, Spokane County Air Pollution Control Authority (SCAPCA), the American Lung Association of Washington, and the Harvard School of Public Health. The purposes of this umbrella study were to collect health and air pollution samples in the area over an extended period, analyze the data for connections between health and pollution, and identify specific links between pollutants and human health effects (Kantamaneni et al., 1996; Norris, 1998; Claiborn et al., 1998; Haller et al., 1999; Gauderman et al., 2000; Norris et al., 2000; Pope, 2000; Villasenor et al., 2001; Hoffman, 2002).

This database represents one of the largest continual $PM_{2.5}$ records available. Its specie coverage complements its length, due to the numerous instrumental methods applied, including carbon, ionic, and trace element species. In several of these cases, more than one instrument was used for the same elements, which increases the likelihood of producing quality measurements and ultimately characterization of the $PM_{2.5}$ compositions. With the added benefits of multiple instruments, the database also provides an opportunity to explore how instrumental differences can affect a source apportionment model. The primary goal of building the database was to create a bank of information with which researchers could figure out many aspects of PM; the added benefits of the anthology would also lead to knowledge regarding instrumental features and long-term shifts over time.

X-Ray fluorescence (XRF) has been used frequently for analyzing trace elements (Chow et. al, 1994; Kim et. al, 2003). Trace elements can exist in very small concentrations in $PM_{2.5}$ but may have significant effects on health. XRF can provide the

composition analysis needed with low detection limits and ample specie coverage. Additionally, samples are unaffected by the XRF process, rendering them reusable for future analysis if necessary. The XRF method is a valuable tool in furthering the knowledge of $PM_{2.5}$ composition.

The purpose of this research was to investigate and determine sources of particulate pollution using positive matrix factorization, based on major $PM_{2.5}$ constituents and the full XRF trace element analysis. Previous studies have used only segments of the database and this work will continue the path of particulate knowledge for the Spokane Health Effects data by completing the XRF study.

2.2 Methods

2.2.1 Collection and Species Measurement

Collection for the Spokane Health Effects study began on January 1st, 1995, and ended on May 15th, 2002, with 2692 sampling dates available for specie analysis by carbon, ion, and trace element instrumentation. Samples were taken over a 24 hour period, at an ideal flow rate of 15 liters per minute (or an ideal total sample volume of 21.6 m^3 per daily sample). **Figure 1** shows the $PM_{2.5}$ collected for the study. Lab and field blanks were used with every filter preparation and exchange to assess instrument drift and other operational issues. Four Versatile Air Pollutant Samplers, or VAPS (Stevens et al., 1993; Sommerville et al, 1994) were used at the receptor site, three for each 24-hour period and a fourth as a field blank, for an average site visitation of every three days; in addition to regular site checks, monthly function and hardware checks were also made. Exposed filters were returned to the laboratory, weighed and stored in a

controlled environment until evaluation by the analysis instruments. A Tapered Element Oscillating Microbalance (TEOM) was also used at the site for gravimetric PM data, to verify VAPS masses (Rupprecht and Patashnick, TEOM series 1400a).

Two XRF instruments were used for trace analysis. The EPA Kevex Energy Dispersive X-Ray Fluorescence (ED-XRF) instrument was chosen as a primary trace element sampler, along with a newer ED-XRF instrument manufactured by Jordan Valley (JV), model 6600AF, which was brought from EPA Region 10 and operated in the WSU Laboratory for Atmospheric Research. The purpose of using a second instrument was to shorten the wait time between sampling and analysis, as the Kevex instrument was backlogged with samples. Using a second instrument would introduce issues regarding instrument function and differing sensitivities; however these issues were considered secondary to delaying the generation of species information for apportionment modeling.

To complete the specie analyses, carbon, SO_4^{2-}, NO_3^-, and NH_4^+, were added to the trace element measurements. The carbon was measured with the thermal manganese oxidation (TMO) and thermal optical transmittance (TOT) instruments, while ion chromatography derived ion concentrations. These species, along the trace elements were used in PMF analysis because they have been shown in previous Spokane Health Effects studies to be major constituents of particulate pollution (Hoffman 2002; Kim et al., 2003). Additionally, PM reconstructions calculated as part of laboratory quality checks conclude these species, along with trace elements, account for nearly all PM mass.

2.2.2 Source Apportionment Modeling

This apportionment research was carried out primarily using PMF, EPA version 1.1. The EPA version was written specifically to treat environmental data; thus many of the settings that modelers have to establish before running the program are set as defaults, decreasing preparation time. This model is based on the PMF model family released in the 1990s and was received with acclaim in the apportionment field (Paatero, 1997; Huang et al., 1999; Lee et al., 1999; Xie et al., 1999; Ramadan et al., 2000; Willis, 2000; Polissar et al., 2001; Song et al., 2001; Henry, 2002; Paatero et al., 2002; Qin et al., 2002) The EPA model uses the multilinear engine (ME) algorithm in lieu of the more traditional PMF2, which relaxes the constraint of nonnegativity in the concentrations. In principle, negative contribution to PM mass is not logical; small negative values in ME allow more freedom for PMF to adhere a source profile to the data. The EPA version is also equipped with more a flexible data format, batch runs using several random points to reach the χ^2 minimum, external controls such as uncertainty down weighting and post-analysis uncertainty bootstrapping, and analysis graphics. Some controls, such as Fpeak, are not yet available in the EPA program, so the original model was used on the data as well, as a quality standard of comparison. The results presented below are all produced by the EPA version, with exception of the model-actual PM reconstructions, based on the original PMF2 script version.

Kevex and JV data were calculated differently, and adjustments were made to create data continuity for the PMF model. Jordan Valley data were not reported with detection limits or uncertainties, but Kevex data contained uncertainty along with

minimum detection limits (MDLs) (Kellogg, 1994). The JV uncertainties were calculated, per specie, as follows:

$$UNC_x = \sigma_x + \frac{1}{3} MDL_x \qquad (1)$$

where UNC_x is the uncertainty per sample, and σ_x and MDL_x are the standard deviations and minimum detection limits for a specie in the JV dataset. *Equation 1* was used to determine uncertainties for SO_4^{2-}, NO_3^-, and NH_4^+, also used in the model. Using this equation for uncertainty brings into the uncertainty values instrument function and sensitivity. Recall that the MDL values are calculated directly from the blank values, an indirect measure of the instrumental sensitivity. When involving so many instruments, it is important to recognize and address the underlying instrumental biases in the data. While the ion uncertainties were calculated in the same manner, their ambient concentrations were generally high enough not to be considered small source contributors. The trace elements by nature, however, were more scarce in mass. As discussed previously, several PMF model settings were tested to ensure a robust source profile was determined, representative of the true nature of all species involved.

In addition to output differences, the JV detector also began to suffer deterioration during the analyses, which decreased agreement between the two XRF data, and further clouded how to build continuity in the data for the PMF model. Because of the detector problem, MDLs were calculated directly from the data in lieu of instrument function, in hopes of capturing the additional uncertainty from the hardware problem. MDLs were taken as the average blank values, per specie; blanks and standards were regularly included in the analysis, to limit instrument drift and treat detector degradation. Kevex and JV compare well for many elements (see **Table 1**). Along with detector consideration

embedded in the MDLs, two additional corrections were made to the JV data. Kevex was used as a benchmark by which JV was standardized, so a step change factor was applied to the JV data. The second correction was applied to correct for the detector shift seen in standards, as the hardware began to fail. Both corrections were made per specie, not as general blanket multipliers, and were meant to treat superficial influences, not alter the source-inspired variability.

As with any dataset, the information used in this study is limited by instrumental imperfections and assumptions. An important consequence of using an equation to determine uncertainties in lieu of instrument-derived uncertainties is that it can decrease variability and makes the job of a model more difficult to identify the origin of the pollution. Instrumental corrections can also inhibit a model's ability to separate variation due to pollution sources from background noise. Use of multiple instruments often introduces abrupt changes in data behavior. All of these design choices impose external influence on what should be raw pollution data. In spite of the inherent issues with using several instruments with varying capabilities, the methods set forth in this study have provided unique apportionment information, because of the uniquely large continuous PM mass and species data history, and particular strengths of more than one instrumental method and how they affect one another.

2.2.3 Post-PMF Analyses

A series of bootstrap PMF programs were run after the main PMF model was selected, to determine mass contributions from each source feature. The bootstrap program provided in the EPA PMF model characterizes the uncertainties associated with

the derived profile solution. Output indicates the confidence intervals for each source mass contribution, as well as the amount of variation explained for each specie in these source features. In lieu of using the given Kevex-JV data group split, data were separated into five smaller groups. Using smaller groups of data would help track any changes in the profiles, which could be due to shifts in the feature contributions over time, or changes due to the instruments. Included in these datasets were the actual $PM_{2.5}$ mass measurements, with mass uncertainty down weighted thirty fold. The strong increase of uncertainty would allow the $PM_{2.5}$ masses to show up in the source feature profiles without affecting their structure. The smaller groups (485-563 samples each) were analyzed with $PM_{2.5}$ mass by the bootstrap feature to characterize source feature mass inputs to the overall $PM_{2.5}$, while taking into account gradual changes in features over time, as well as instrument shifts. The increased mass uncertainty effectively eliminates $PM_{2.5}$ influence on the source profile, and the net outcome signified how PMF separates mass into its constituent features. The 95^{th} percentile $PM_{2.5}$ mass source uncertainties were used as ranges of variability in mass contributions for each source feature and helped qualify the overall Kevex and JV profiles.

In addition to bootstraps, independent t-tests were performed on the five subgroups, to further clarify whether the subsets of data show any temporal or instrumental anomalies beyond natural variations. The base information for these tests originated from variances determined from the daily source masses, as well as variances estimated by the 5 and 95^{th} percentile confidence intervals produced by bootstrapping. Ten tests were conducted for each source feature (all pair combinations), and the primary hypothesis was whether average mass contributions for a given source by one data group

was similar to that of another data group. A 95% confidence interval was the guideline, and the following equations were used to verify or disprove the hypothesis:

$$s_{pool} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{2a}$$

$$V_{(\bar{y}_1 - \bar{y}_2)} = s_{pool}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \tag{2b}$$

$$s_{(\bar{y}_1 - \bar{y}_2)} = s_{pool} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \tag{2c}$$

$$\bar{y}_1 - \bar{y}_2 \pm t_{v,\frac{\alpha}{2}} \left(s_{\bar{y}_1 - \bar{y}_2} \right) \tag{2d}$$

where s_{pool}, $V_{(\bar{y}_1 - \bar{y}_2)}$, and $s_{(\bar{y}_1 - \bar{y}_2)}$ are the pooled standard variance, estimated variance of the difference, and standard error of the difference. If the hypothesis is true, then the range of values calculated should include 0; that is, the expected difference in averages should show similar average masses, after consideration for variability. The t value chosen was 1.96, based on a large sample size and confidence interval of 95%.

2.3 Results and Discussion

Based on previous work on the Spokane Health Effects study, it was hypothesized the model would be optimized with 6-8 factors, using the following species: *TC, NO_3^-, SO_4^{2-}, NH_4^+, Si, Cl, K, Ca, Mn, Fe, Zn, Cu, Br,* and *Pb*. Species were chosen because of their relative abundance and previous research indicating significant particulate contribution. **Figure 2** shows which species exhibit concentrations well above detection level. **Figures 3a** and **3b** show continuous data between the two instruments for the chosen species; the spikes roughly halfway through the data represent the transition

between Kevex and Jordan Valley. If too many factors are chosen, actual χ^2 would be smaller than expected χ^2, residuals may have a positive or negative skew, and one or more of the source features would not be identifiable. If too few factors are chosen, χ^2 may be too large, a significant proportion of the variation would not be explained, and poor correlation between the measured and model generated masses would result.

In spite of the corrections made to the JV data, initial models using the entire combined Kevex-JV dataset did not produce reasonable PMF source features. Both DOS and EPA version models could not effectively derive more than four of the original source features, and much of the variation was left unexplained. Modeling the data as two independent sets provided the clearest and representative source profile for the Spokane data.

2.3.1 Kevex and Jordan Valley PMF Models

PMF was run on Kevex data, analyses spanning 1/1/1995-3/31/1999, and JV data, from 4/1/1999-5/15/2002. The Kevex model relied on 1456 samples, approximately 94% of available sampling dates. Of the 6% eliminated, many were found within the first year, and were due to instrumental, contamination, or other operational errors associated with establishing a new sample collection mission. The Jordan Valley dataset was comprised of 1130 samples approximately 99% of available dates. The PMF models were run with a choice of factors ranging from five to ten, with six factors producing the most physically realistic results. PMF model settings were mostly defaults, because the EPA version is designed to receive environmental data, and the Kevex data did not display any abnormal distributions or instrument-induced problems. Initial work using the original PMF model,

in which error model, uncertainty, Fkey, Fpeak, and factor number settings were modified to test robustness, support these default choices.

The six factors produced yielded reasonable correlation scores between models and actual mass concentrations, with maximum R^2 values of 0.77 and 0.72, for Kevex and Jordan Valley, respectively. **Figures 4**, **5**, and **6** show the PMF generated $PM_{2.5}$ data versus the VAPS-TEOM measured fine fraction mass. Agreement between the model and actual masses decreases with increase in mass. This spread emphasizes the need for a large dataset; the influence of a single comparison between model and actual data would lessen with a greater number of comparisons.

The six source features – vehicle exhaust, a chlorine rich source, airborne soil, biomass burning, metal processing, and nitrate – generally agree with previous research on this dataset (Hoffman, 2002; Kim et al., 2003; Schreuder et. al, 2006). The work by Hoffman was a chemical mass balance study done for a subset of the database and served primarily as a starting point for identifying major pollution sources. The work by Kim et. al was a PMF study on the first three years only of carbon, ion, and Kevex XRF data, and the Schreuder research was an epidemiology study using the entire time period. While Schreuder et. al used all years of the database, simpler instrumental adjustments were made between XRF instruments, other trace element data was used, and the focus was primarily on the epidemiology aspects. The Kim and Schreuder studies concluded additional sources, specifically sulfate, a specie which is tied into other sources for this study. For Schreuder et. al, As-rich and marine sources were also found, based on trace element data from another method. Among the three PMF studies, the corresponding sources agree in structure and generally in mass contribution. The bootstrap uncertainty

program performed on the model run chosen shows much of the variation is effectively explained by the six factors presented here. The original version used to calculate mass reconstruction also optimized the model with a choice of six source features, using the "iterative least squares" error model (EPA default), and a minimal positive Fpeak value.

The nitrate source feature is seen in **Figure 7**, and is comprised of NO_3^-, NH_4^+, SO_4^{2-}, carbon, crustal elements, and Br. Nitrate can be due to a secondary formation in the air; that is, its parent species are the pollutants emitted into the air. One can see similarities in the presence of carbon, Pb, and Zn, with the addition of crustal elements to the nitrate source feature. The nitrate source feature may be a second generation source feature of vehicle exhaust combined with airborne soil. Production of this feature is fairly consistent year to year, but increases during the warmer months, in congruence with the nature of NO_3^- formation from NO_x. Despite this formation process, nitrate is still considered a unique feature, because it appears independently of other source features, over time, and in composition. Nitrate, seen in **Figure 7**, shows the primary markers for this feature: nitrate, sulfate, crustal elements, Pb, and Br, a composite of soil and vehicle reintroduced into the air. Because Br is much higher than in Kevex, and only shows up in two features, it may be tied to Pb which is related to soil and vehicle exhaust.

Metal processing, seen in **Figure 8**, is similar to the chlorine rich source feature in that it is comprised, by relative specie contribution, of one significant specie. Nearly all of the variation in copper in the Kevex and JV datasets is accounted for in this feature. Supporting this conclusion are the patterns in the temporal analysis. By season and day of week, this feature seems to remain relatively constant. Looking at the year, however, one can see a noteworthy rise and fall for 1997. Coupled with that pattern are several peaks in

the time-series plot, both of which agree with notion of increased and decreased production at an industrial plant. The industrial processing shows strong agreement with Kevex, and is seen in **Figure 8**. The primary input is copper, with secondary inputs from carbon, ions, and some crustal elements.

Biomass burning is another common particulate source feature for Eastern Washington (**Figure 9)**. The high levels of carbon and K, both in relative source feature and specie mass, suggest combustion of agricultural materials. Consistent input for every year, but seasonal peaks in Fall and Winter (which corresponds to biomass burning) support in time, what the species suggest in content. Pb and Br may be found in the ground or could be debris from farming equipment. Biomass burning, seen in **Figure 9**, has significant amounts of Br, K, carbon, SO_4^{2-}, NO_3^-, NH_4^+, and crustal elements. The profile agrees with Kevex, except that it appears to include other types of combustion (vehicle and municipal). Seasonal analyses, however, show similar peaks in cool months, and fairly consistent emission by year.

The chlorine rich source feature has a prominent Cl component, with secondary contributions by ions and carbon, indicating a high heat combustion feature. **Figure 10** shows this feature. The municipal incinerator in Spokane emits chlorine, nitrates, sulfates, and carbon. The JV chlorine rich feature does not have as clear a profile as for the Kevex data, containing more geological elements than the earlier data. Both XRF models show that the cooler months appear to have more concentrations by this feature. Increased loading of solid waste during fall and winter could be the reason for increased particulate loading, or meteorological conditions may enhance the loading at the receptor site. Concentrations are fairly consistent from week to weekend days, suggesting a

continual input of fuel. This source feature has a few episodes of extremely high relative contributions to $PM_{2.5}$ which may be of epidemiological concern. The JV data suggest an increase in overall emissions in the later years, which may be due to increase solid waste loads to the incinerator.

Vehicle exhaust is another common source feature of $PM_{2.5}$. The presence of Pb, Zn, carbon, and sulfate confirm this type of combustion. **Figure 11** shows the source feature as well as the temporal analysis. JV leaves out Zn and adds K in this source feature, possibly due to different sensitivities of this trace element for the two instruments or confusion of combustion types. The JV source feature also decreases input of SO_4^{2-}; sulfate standards were performed less frequently at the end of data collection, which may have influenced concentrations. Some crustal elements also occur in this source feature, but are, by relative specie contribution, small. These soil constituents may get tied into exhaust and carried downwind to the receptor by vehicle interaction with road surfaces. By year, traffic volume-induced pollution appears consistent, with a small dip in 1997 and rise in 1999, and small seasonal increases in winter. This particulate feature is slightly more prominent during the work week, suggesting primary emissions may be due to professional commuters in lieu of weekend traffic.

Airborne soil is a common source feature of particulate pollution throughout an entire calendar year and the entire study, with significant seasonal fluctuations. **Figure 12** shows the source profile and temporal analysis for this feature. Crustal elements are prominent, both in relative feature and specie mass, while combustion and ion species are less abundant, relative to their overall specie contribution. One can see the increase of this feature in the summer months, corresponding to the dry, hot conditions Eastern

Washington experiences during these months, which may aggravate $PM_{2.5}$ levels. Looking at the overall $PM_{2.5}$ concentrations from 1995-2002, of which airborne soil is a major component, the levels appear consistent throughout the study.

Overall, XRF data maintain the basic structure of the pollution features previously identified (Hoffman, 2002; Kim et al., 2003). Lack of a distinctive sulfate source feature is due to sulfate emulating carbon, in that it is present in all features. With the improvement of pollution control strategies, sulfur combustion products may occur more equally among all combustion features and not in a single concentrated process, a feature that could also have affected the entire Kevex period (previous studies only modeled the first three years of Kevex data). The differences in JV and Kevex models are likely due the detector problems of the JV instrument and assumptions made about MDLs values used in assigning uncertainty.

2.3.2 Validation of the Source Features with bootstrap and t tests

Bootstrap $PM_{2.5}$ mass uncertainties and the corresponding mean contributions, given in **Table 2**, highlight instrumental and temporal effects. The first values in **Table 2** are the means, in terms of $\mu g/m^3$ and per cent, and the parenthetic values are the limits of variation, for the 5% and 95% confidence intervals. **Table 3** shows the same results for samples containing the highest twenty per cent masses, samples of most concern for regulation. At first glance, it may look as though the more problematic Jordan Valley data dominates Kevex at assigning contribution to $PM_{2.5}$ by feature; however, recall that JV data have uniform uncertainty, limiting the variation that PMF would use in modeling. **Table 4** shows the Pearson correlations for source features and species, values which

clarify the relationship among the $PM_{2.5}$ constituents. Understanding the nature of the source feature is a combination of the contribution to mass by each feature as well as understanding how the species interact.

The ranking of the features and the mean mass contributions may be more useful to see how the groups compare to one another. Airborne soil seems to have the overall smallest uncertainty. Soil is determined by several, large trace element quantities, so its structure is not difficult for PMF to capture, as is seen by the significant positive correlations to several crustal elements. For Kevex, metal processing and Cl-rich incineration also have relatively small uncertainties, due to their unique presence in their respective features. (Note also that these features have high correlations with single specie contributors.) JV shows similar attributes for metal, but not for Cl. Even with all the corrections made, JV had trouble accurately characterizing Cl; detector loss indicated a considerable shift in raw output data, which has affected the PMF profile. Uncertainties were amplified by several factors in an attempt to force a separation between Cl and the exhaust markers, but no appreciable effect was observed. Another possible influence would be that the Jordan Valley did not have a calibration standard available for Cl. While an XRF measurement may have seemed reasonable in the raw concentration output, relative to its fellow trace elements, Cl values may be inaccurate. In spite of the noted problems, the model attempts to separate Cl (with some exhaust markers) into its own source feature, a validation of the Kevex data and overall source feature. Vegetation and nitrate have larger uncertainties because they are subject to seasonal and short event emission increases and decreases. Unlike the other combustion sources, vehicle exhaust should have smaller uncertainties because of its consistent emission; however, a temporal

shift in pollutant composition, notably Pb, has changed the source profile over time, affecting the $PM_{2.5}$ estimate. The $PM_{2.5}$ bootstraps suggest that seasonal and shorter term source features can be more unpredictable in their PM contributions, while source features with strong markers are easier to characterize effectively.

Mass contributions as reported by Kim et. al and Schreuder et. al generally agree with the results found in this study. The Schreuder and Kim studies, along with these results indicate similar rankings of mass contributions for vehicle exhaust, biomass burning, metal processing, and Cl-rich environments. The biomass burning and vehicle exhaust were among the highest contributors for $PM_{2.5}$ mass, while metal and Cl-rich environments yield the lowest mass (the Schreuder work did not conclude any significant Cl). These results agree strongly with Kim et. al regarding airborne soil, while Schreuder et. al contend a higher mass donation. The latter work also indicates a fairly high standard deviation; different criteria for eliminating abnormally high PM samples could be affecting the PMF models in this case. For nitrate, this work concludes a strong nitrate mass input, whereas the prior publications indicate more moderate masses. This discrepancy could be due to interference of Cl-rich mass into the nitrate source, in the later, Jordan Valley data.

The t test results were useful in highlighting both temporal fluctuations and instrumental variations in source feature behavior. **Table 4** shows which group tests yielded proven and failed hypotheses for agreement between group averages. As previously mentioned, the JV detector began to degrade at the end of the study, slowly shifting elemental measurement. The t test hypotheses failed in several comparisons of the second JV dataset to other groups. The first and second Kevex groups did not

correspond as well to other groups as did the third Kevex group. Reviewing the bootstrap intervals (**Table 2**), a much larger variation was attributed to biomass burning (the largest variation within the whole study period), suggesting a lot of mass variation was captured by agricultural activities. Isolating so much of the variation By feature, Cl-rich, metal processing, and nitrate had the highest incidence of passed t test hypotheses. These source features are marked by exclusive species, which are easier for the model to distinguish. Biomass burning and vehicle exhaust source features share several common species; airborne share similarities with crustal elements, and like biomass burning, can vary by season.

2.4 **Conclusions**

Eastern Washington is susceptible to particulate matter pollution violations, and therefore correctly modeling the source features is essential in remedying the problem. PMF models using a suite of XRF, ionic, and carbon species have found six definite source features of pollution: motor vehicle combustion; chlorine-emitting combustion, from the municipal incinerator; biomass burning, from agriculture; airborne soil, from agriculture, traffic, and climate; metal processing, due to a downwind metal smelter; and nitrate, from secondary processes associated with other source features. The Kevex and Jordan Valley data treat the features somewhat differently as separate datasets; however, the same markers can be found in these results and agree with earlier work. Kevex produces slightly more defined profiles, due to the longer time period and reliable instrument. In spite of the instrumental failures, the Jordan Valley model shows that the present and previous Kevex analyses produce a correct, realistic profile of Spokane

particulate pollution. The discontinuity suggests that more work needs to be done on resolving the instrumental differences in species studied, either revisiting the correction factors or using companion trace elements analyses as profile standards for JV. Future work will include running PMF with carbon, ion, XRF data with more trace element instrumental analyses data, as they are completed. All of these projects will be performed with the intent of resolving the most accurate, representative picture possible of Eastern Washington particulate matter air pollution.

Acknowledgements

This study was funded in part by the Integrative Graduate Education and Research Training (IGERT) grant from the National Science Foundation to Washington State University under grant DGE-0072817, the Washington State Department of Ecology, Mickey Leland Nation Urban Air Toxics Research Center, University of Washington/EPA Northwest Research Center for Particulate Matter and Health Center (R827355) and University of Rochester/EPA Particulate Matter and Health Center (R827354). This research has not been subjected to EPA's required peer and policy review and does not necessarily reflect the views of EPA.

References

Chow J., Watson J., Fujita E., Lu Z., Lawson D., Ashbaugh L., 1994. Temporal and spatial variations of $PM_{2.5}$ and PM_{10} aerosol in the Southern California Air Quality Study. Atmospheric Environment 28, 2061-2080.

Claiborn, C.S., Finn, D., Larson, T.V., Koenig, J., 2000. Windblown dust contributes to high PM2.5 concentrations. Journal of Air and Waste Management Association 55, 1440-1445.

Dockery, D.W., Pope, C.A., Xu, X.P., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air-pollution and mortality in 6 United States cities. New England Journal of Medicine 329, 1753-1759.

Gauderman, W.J., McConnell, R., Gilliland, F., London, S., Thomas, D., Avol, E., Vora, H., Berhane, K., Rappaport, E.B., Lurmann, F., Margolis, H.G., Peters, J., 2000. Association between air pollution and lung function growth in southern California children. American Journal of Respiratory and Critical Care Medicine 162, 1383–1390.

Gray H.A., Cass G.R., 1986. Characteristics of atmospheric organic and elemental carbon particle concentrations in Los Angeles. Environmental Science and Technology 20, 580-588.

Haller, L., Claiborn, C., Koenig, J., Larson, T., Norris, G., and Edgar, R., 1999, Airborne particulate matter size distributions in an arid urban area. Journal of the Air & Waste Management Association 49, 161-168.

Henry, R.C., 1987. Current factor analysis models are ill-posed. Atmospheric Environment 21, 1815-1820.

Henry, R.C., 2002. Multivariate receptor models – current practices and future trends. Chemometrics and Intelligent Laboratory Systems 60, 43-48.

Hien, P.D., Bae, V.T., Thinh, N.T.H, 2004. PMF receptor modeling of fine and coarse PM10 in air masses governing monsoon conditions in Hanoi, northern Vietnam. Atmospheric Environment 38, 189-201.

Hoffman, M., Finn, D., Claiborn, C.S., 2001. Elemental analysis and receptor modeling of airborne particulate matter collected in Spokane. Annual conference of the Pacific Northwest International Section of the Air and Waste Management Association, Big Sky, MT.

Hoffman, M. (2002). Elemental analysis and receptor modeling of airborne particulate matter in Spokane, Washington, M.S. thesis, Washington State University, Pullman, WA.

Hopke, P.K., 1991. Receptor Modeling for Air Quality Management. Elsevier, Amsterdam, The Netherlands.

Huang, S., Rahn, K.A., Arimoto, R., 1999. Testing and optimization two factor-analysis techniques on aerosol at Narragansett, Rhode Island. Atmospheric Environment 33, 2169– 2185.

Kantamaneni, R., Adams, G., Bamesberger, L., Allwine, E., Westberg, H., Lamb, B., Claiborn, C.S., 1996. The measurement of roadway PM10 emission rates using atmospheric tracer ratio techniques. Atmospheric Environment 30 (24), 4209–4223.

Kellogg, R. 1994, Standard Operating Procedure for the Analysis of Aerosols on Filters with the Source Apportionment Research Branch Kevex Energy Dispersive X-ray Fluorescence Spectrometer, 39 p. ManTech Environmental Technology, Inc., Research Triangle Park, NC.

Kim, E., Larson, T.V., Hopke, P. K., Slaughter, C., Sheppard, E., Claiborn, C., 2003. Source identification of PM2.5 in an arid northwest U.S. city by positive matrix factorization. Atmospheric Research 66, 291-305.

Lee, E., Chan, C.K., Paatero, P., 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmospheric Environment 33, 3201-3212.

Liu, W., Hopke, P.K., Han, Y-j, Yi, S-M, Holsen, T.M., Cybart, S., Kozlowski, K., Milligan, M. Application of receptor modeling to atmospheric constituents at

Potsdam and Stockton, NY. Atmospheric Environment 37, 4997-5007.
Norris, G., 1998. Air pollution and the exacerbation of asthma in an arid, western, US city, PhD dissertation, University of Washington, Seattle, WA.
Norris, G., T. Larson, J. Koenig, C. Claiborn, L. Sheppard, D. Finn, 2000. Asthma aggravation, combustion, and stagnant air. Thorax 55, 466-470.
Paatero, P., 2000. User's guide for positive matrix factorization programs PMF2 and PMF3, Part I: tutorial.
Paatero, P., Hopke, P.K., Song, X.H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. Chemometrics and Intelligent Laboratory Systems 60, 253-264.
Phalen, R.,F., 1998. Uncertainties relating to the health effects of particulate air pollution: The US EPA's particle standard. Toxicology Letters 96,97, 263-267.
Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: 1. Elemental composition and sources. Journal of Geophysical Research 103 (D15), 19045– 19057.
Pope III,C.A., 2000. Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk? Environmental Health Perspectives 108 (Suppl. 4), 713–723.
Ramadan, Z., Song, X.H., Hopke, P.K., 2000. Identification of sources of Phoenix aerosol by positive matrix factorization. Journal of the Air & Waste Management Association 50, 1308–1320.
Qin, Y., Oduyemi, K., Chan, L.Y., 2002. Comparative testing of PMF and CFA models. Chemometrics and Intelligent Laboratory Systems 61, 75-87.
Schwartz, J.G., G. Norris, T. Larsen, L. Sheppard, C. Claiborn, J. Koenig, 1999. Episodes of high coarse particle concentrations are not associated with increased mortality. Environmental Health Perspectives 107, 339-342.
Schreuder, A.B., Larson, T.V., Sheppard, L., Claiborn, C., 2006. Ambient woodsmoke and associated respiratory emergency department visits in Spokane, WA, Journal of Occupational and Environmental Health 12:2, 147-153.
Song, X-H, Polissar, A. V., Hopke, P.K., 2001. Sources of fine particle composition in the northeastern US. Atmospheric Environment 35, 5277-5286.
Villasenor, R., Claiborn, C., Lamb, B., O'Neill, S., 2001. Mesoscale modeling of wintertime particulate matter episodes in eastern Washington, USA. Atmospheric Environment 35, 6479-6491.
Willis, R.D., 2000. Workshop on UNMIX and PMF as applied to PM2.5. EPA 600-A-00-048.
Xie, Y.L., Hopke, P.K., Paatero, P., Barrie, L.A., Li, S.M., 1999. Identification of source nature and seasonal variations of Arctic aerosol by positive matrix factorization. Journal of Atmospheric Science 56, 249–260.

XRF	Minimum Detection Level		Uncertainty	
Element	Kevex	JV	Kevex	JV
Si	0.0084	.0275	0.0499	0.1126
Cl	0.0050	.0161	0.0045	0.0245
K	0.0066	.0025	0.0099	0.0299
Ca	0.0094	.0006	0.0079	0.0124
Mn	0.0008	.0013	0.0008	0.0020
Fe	0.0008	.0069	0.0166	0.0374
Cu	0.0008	.0001	0.0031	0.0021
Zn	0.0010	.0004	0.0018	0.0040
Br	0.0006	.0007	0.0006	0.0011
Pb	0.0016	.0019	0.0015	0.0023

Table 1. Instrumental limits of detection (MDL) for the Kevex and Jordan Valley XRF. Decrease in sensitivity in detection limits for JV data is accounted for with increased uncertainty values.

$PM_{2.5}$	Kevex Data			Jordan Valley Data	
Source	1/1/95-5/25/96	5/26/96-11/10/97	11/11/97-3/31/99	4/1/99-10/28/00	10/29/00-5/15/02
Airborne Soil	1.06 / 9.70 (0.89; 1.16)	0.53 / 5.0 (0.00; 0.665)	0.49 / 4.60 (0.07; 0.61)	1.71 / 19.9 (1.65; 1.86)	1.07 / 10.5 (1.01; 1.14)
Nitrate	2.96 / 27.0 (2.61; 3.46)	2.31 / 21.5 (2.09; 2.57)	1.81 / 17.0 (1.56; 2.07)	1.78 / 20.7 (1.63; 1.92)	1.66 / 16.4 (1.59; 1.74)
Cl-rich	0.905 / 8.20 (0.81; 1.04)	0.66 / 6.20 (0.60; 0.73)	0.12 / 1.00 (0.09; 0.12)	1.19 / 13.9 (1.13; 1.301)	4.70 / 46.1 (4.56; 4.81)
Metal processing	0.55 / 5.00 (0.48; 0.60)	0.95 / 8.90 (0.85; 1.01)	0.44 / 4.20 (0.40; 0.49)	0.50 / 5.82 (0.47; 0.60)	0.40 / 3.92 (0.34; 0.42)
Biomass burning	3.71 / 33.8 (3.33; 3.87)	2.66 / 24.9 (2.42; 2.88)	5.30 / 49.8 (5.10; 5.42)	1.04 / 12.1 (0.87; 1.06)	1.36 / 13.3 (1.26; 1.42)
Vehicle exhaust	1.81 / 16.4 (1.52; 2.00)	3.60 / 33.7 (3.21; 3.78)	2.48 / 23.3 (2.37; 2.60)	2.37 / 27.6 (2.09; 2.58)	0.98 / 9.60 (0.59; 1.06)
Mass Total	**11.0**	**10.7**	**10.6**	**8.59**	**10.2**

Table 2. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset. Leading numbers are mean $PM_{2.5}$ mass for each feature, in $\mu g/m^3$ as well as per cent, relative to overall source feature mass; parenthetic numbers are 5% and 95% bootstrap confidence intervals, respectively.

$PM_{2.5}$ Source	INAA-Kevex	INAA-JV
Airborne soil	1.54 / 7.00 (1.35; 1.66)	1.85 / 9.83 (1.72; 1.97)
Nitrate	3.08 / 14.0 (2.52; 3.36)	2.65 / 14.1 (2.01; 2.93)
Cl-rich	1.21 / 5.50 (1.05; 1.63)	2.56 / 13.6 (2.39; 2.73)
Metal processing	1.28 / 5.80 (1.10; 1.52)	0.38 / 2.00 (0.32; 0.44)
Biomass burning	8.95 / 40.6 (8.22; 9.71)	8.03 / 42.7 (7.87; 8.28)
Vehicle exhaust	5.99 / 27.2 (5.20; 7.78)	3.35 / 17.8 (3.08, 3.88)
Mass Total	**22.05**	**18.82**

Table 3. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset, for the upper fifth mass fraction. Leading numbers are given as mean $PM_{2.5}$ mass / per cent of total source mass for each feature, in $\mu g/m^3$; parenthetic numbers are 5% and 95% bootstrap confidence intervals, respectively.

JV	Kevex																		
	Bio. Burn	AB Soil	Metal	Nitrate	Veh. Exh.	Cl-rich	TC	NO3	SO4	NH4	Si	Cl	K	Ca	Mn	Fe	Cu	Zn	Br	Pb
Bio. Burn					0.34		0.63	0.33		0.36			0.72					0.37	0.49	0.31
AB Soil											0.99		0.76	0.96	0.92	0.98				
Metal																	1.00			
Nitrate					0.34								0.30		0.36	0.31		0.38	0.59	0.58
Veh. Exh.							0.47						0.42					0.98	0.31	0.47
Cl-rich				0.33								1.00								
TC				0.35		0.87		0.36					0.44					0.47	0.36	
NO3				0.92		0.43	0.44			0.61										
SO4				0.42				0.31		0.30										
NH4				0.81				0.66	0.41									0.30		
Si		0.97											0.78	0.94	0.91	0.98				
Cl						0.33	0.33													
K	0.30	0.75				0.42	0.40				0.71			0.75	0.76	0.80		0.45	0.41	0.37
Ca		0.96									0.89		0.69		0.87	0.93				
Mn		0.88									0.87		0.69	0.82		0.93				
Fe		0.97									0.96		0.76	0.92	0.88					
Cu			1.00																	
Zn						0.83	0.56	0.31					0.35						0.37	0.47
Br													0.39							
Pb					0.97	0.30	0.32													

Table 4. Pearson correlations for Kevex and Jordan Valley data, between sources and species. The left half corresponds to JV data, and the right half represents Kevex data.

$PM_{2.5}$ Source	K1-K2	K1-K3	K1-JV1	K1-JV2	K2-K3	K2-JV1	K2-JV2	K3-JV1	K3-JV2	JV1-JV2
Airborne soil				**X**	**X**					
Nitrate					**X**			**X**	**X**	**X**
Cl-rich			**X**		**X**	**X**		**X**		
Metal processing		**X**	**X**		**X**			**X**	**X**	
Biomass burning	**X**	**X**			**X**					
Vehicle exhaust			**X**		**X**			**X**		

Table 5. Independent t-tests for determining whether PMF derived contributions are representative of one another. An "X" indicates that the t score verified the hypothesis that one group is similar enough to another to representative of one another. The null hypothesis could be due to natural variability in the source features (such as strong source emission events during one time period), or instrumental differences, such as step changes between Kevex and Jordan Valley, or detector degradation in Jordan Valley.

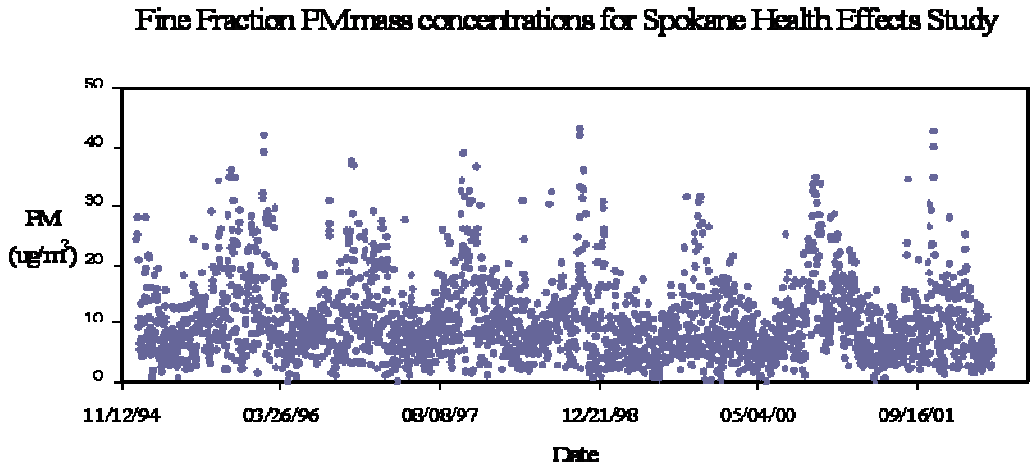


Figure 1. Fine fraction PM for Spokane, WA, from 1/1/1995 – 5/15/2002. Notice the regular peaks in $PM_{2.5}$ corresponding to summer time dust events and winter time combustion.

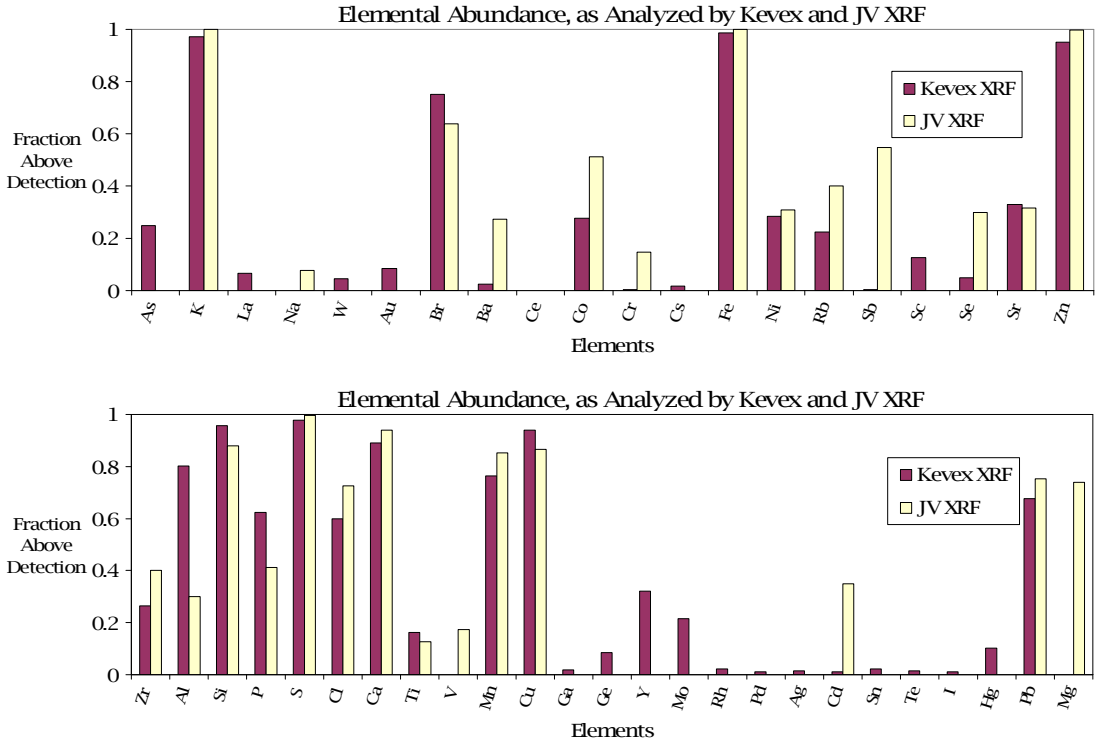


Figure 2. Abundance frequency for Kevex and Jordan Valley instrument analyses on Spokane Health Effects data. This abundance helped determine which species were fit for PMF modeling.

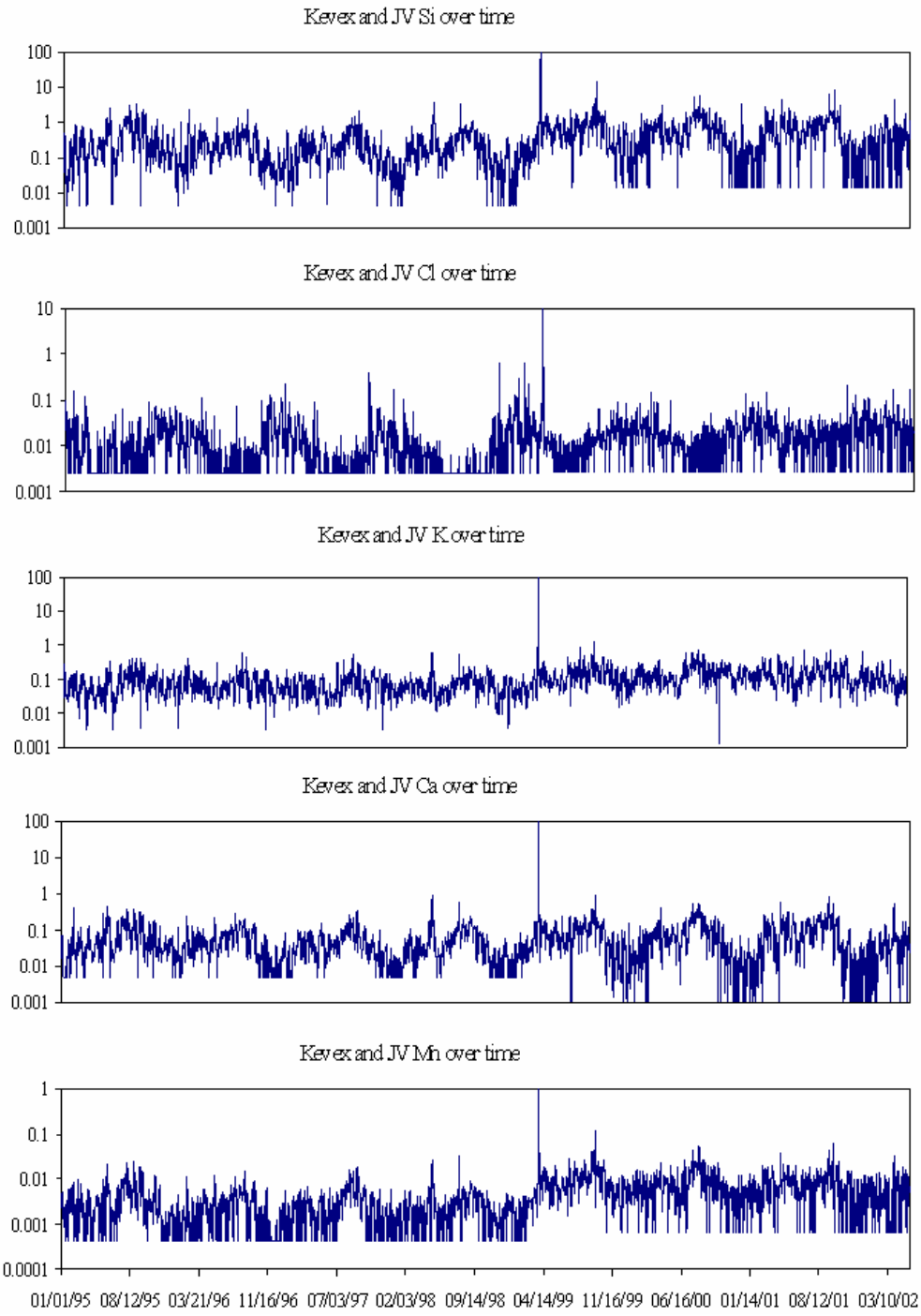


Figure 3a. Time series plots for XRF elements analyzed by PMF. The maxima lines in the middle of the graphs mark where Kevex and Jordan Valley meet. Notice the changes in data vary by species, some nearly seamless and others are more noticeable.

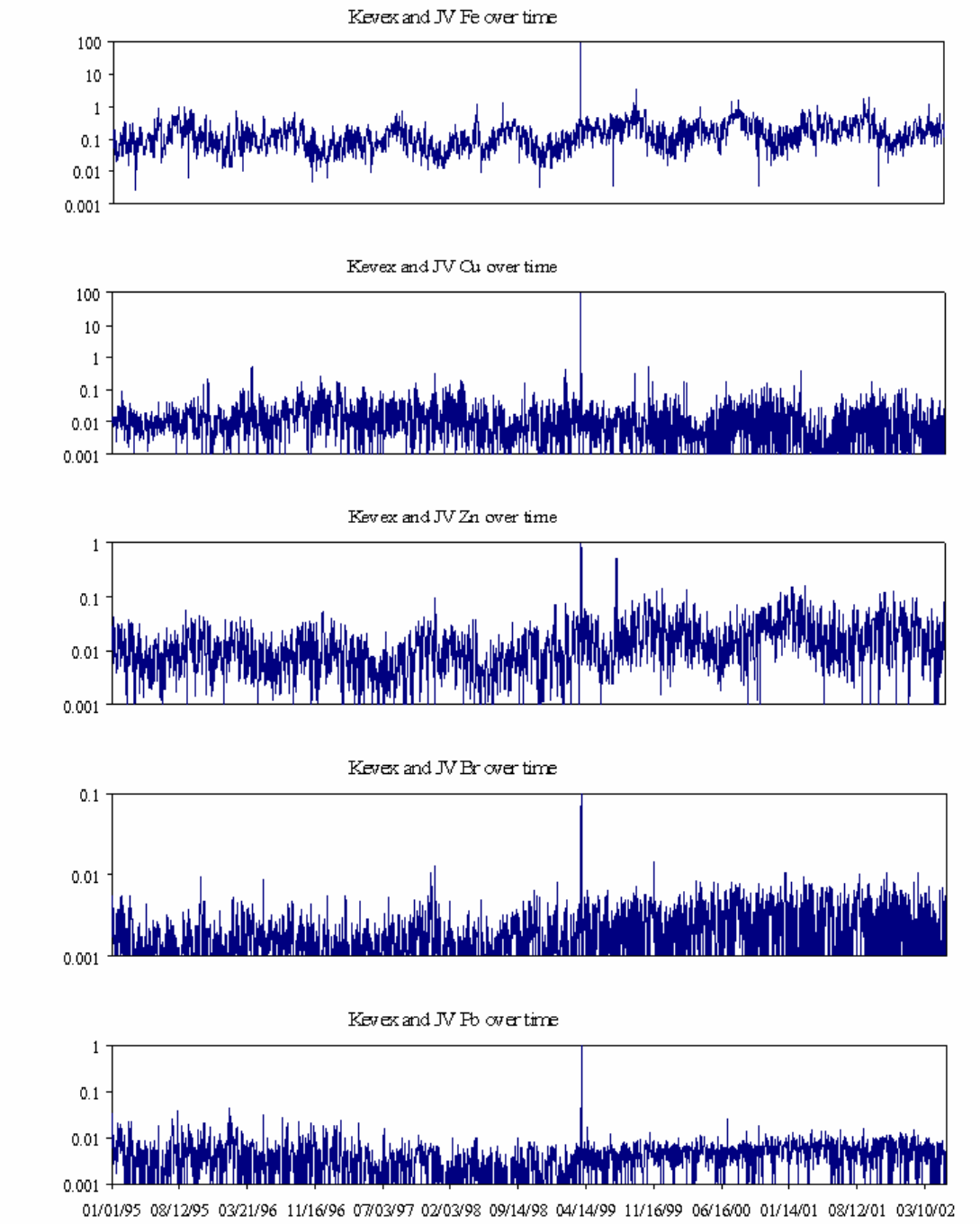


Figure 3b. Time series plots for XRF elements analyzed by PMF. The maxima lines in the middle of the graphs mark where Kevex and Jordan Valley meet. Notice the changes in data vary by species, some nearly seamless and others are more noticeable.

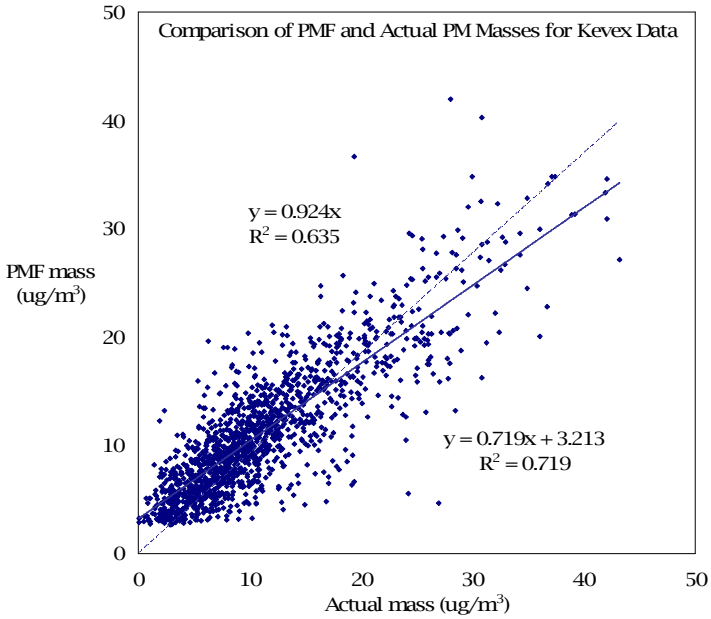


Figure 4. PMF-generated PM mass for Kevex compared to the actual mass measured by VAPS and TEOM instruments. The degree of correspondence is good, especially for lower concentrations.

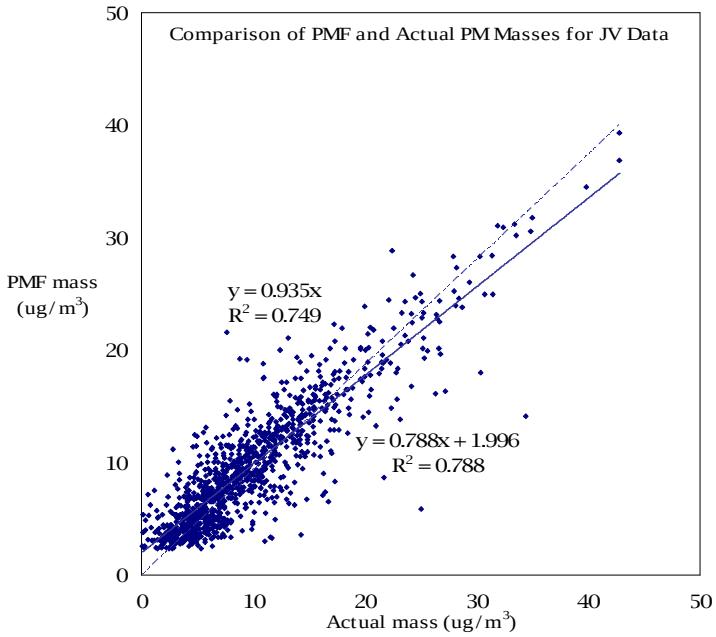


Figure 5. PMF-generated PM mass for Jordan Valley compared to the actual mass measured by VAPS and TEOM instruments. The degree of correspondence is good, especially for lower concentrations.

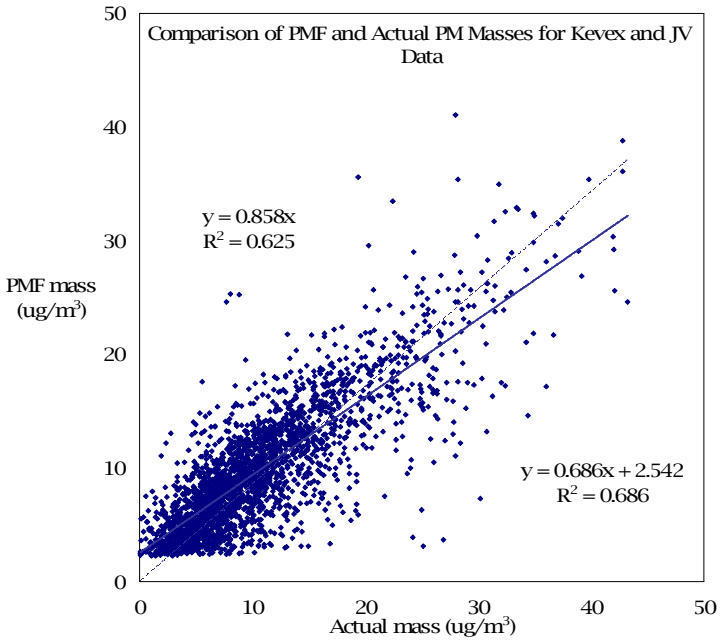


Figure 6. A comparison of PMF-generated PM mass for the combined Kevex and Jordan Valley dataset, to actual mass measured by VAPS and TEOM instruments. Correspondence is fairly good, especially for lower concentrations. Data fan out more noticeably here, indicating some disagreement between Kevex and Jordan Valley.

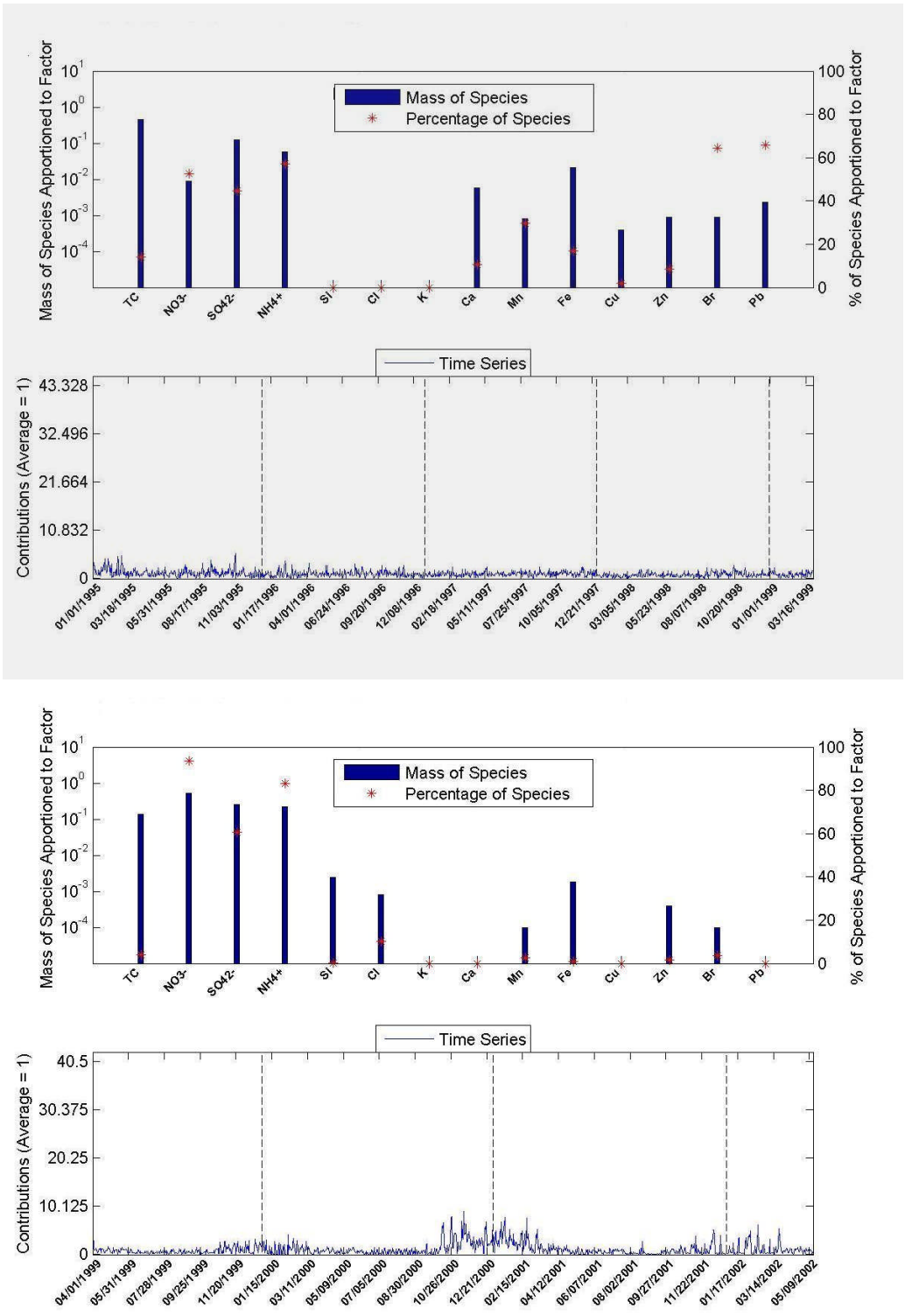


Figure 7. Nitrate for Kevex (top) and JV (bottom) models.

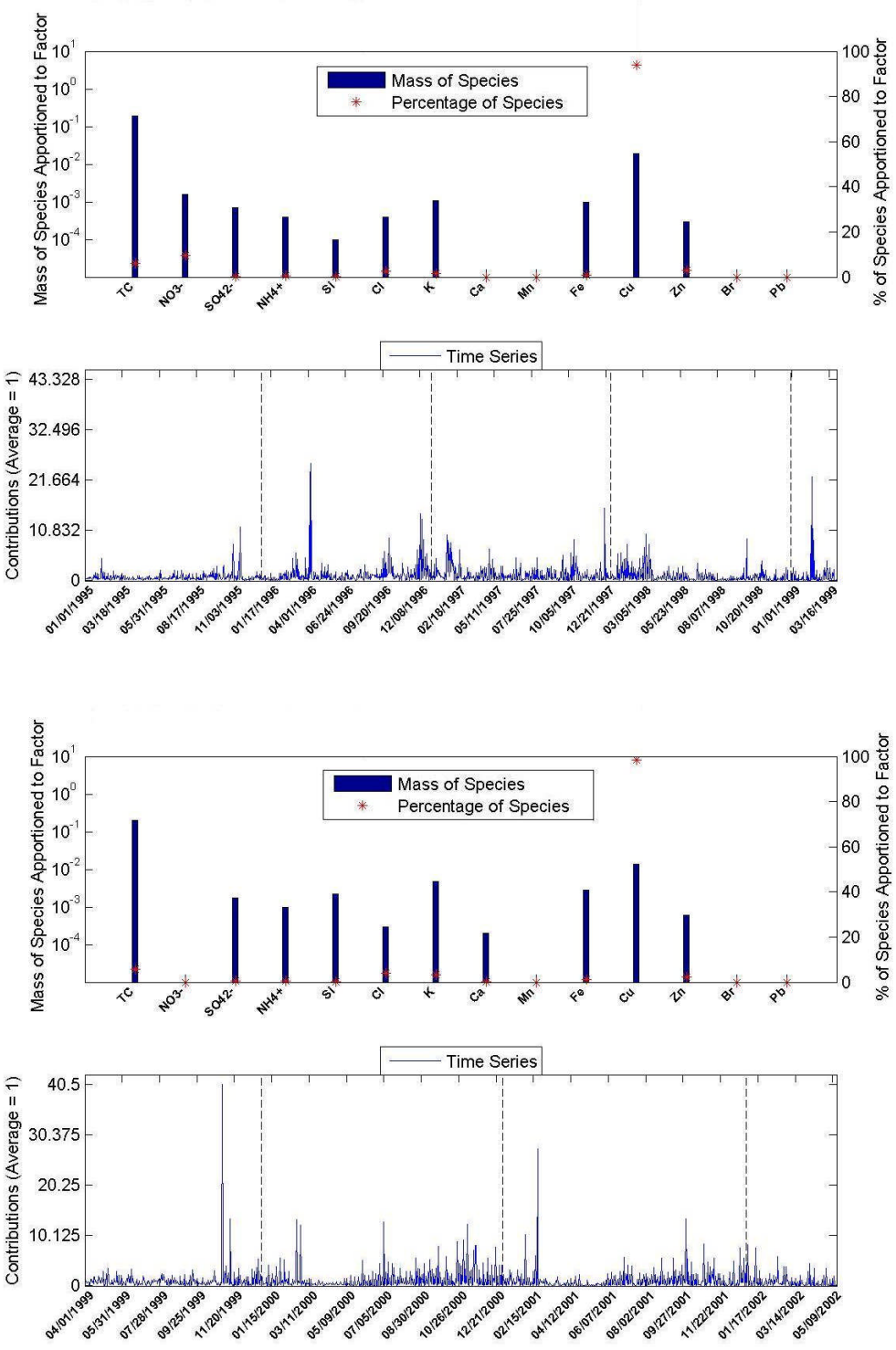


Figure 8. Metal processing for Kevex (top) and JV (bottom) models.

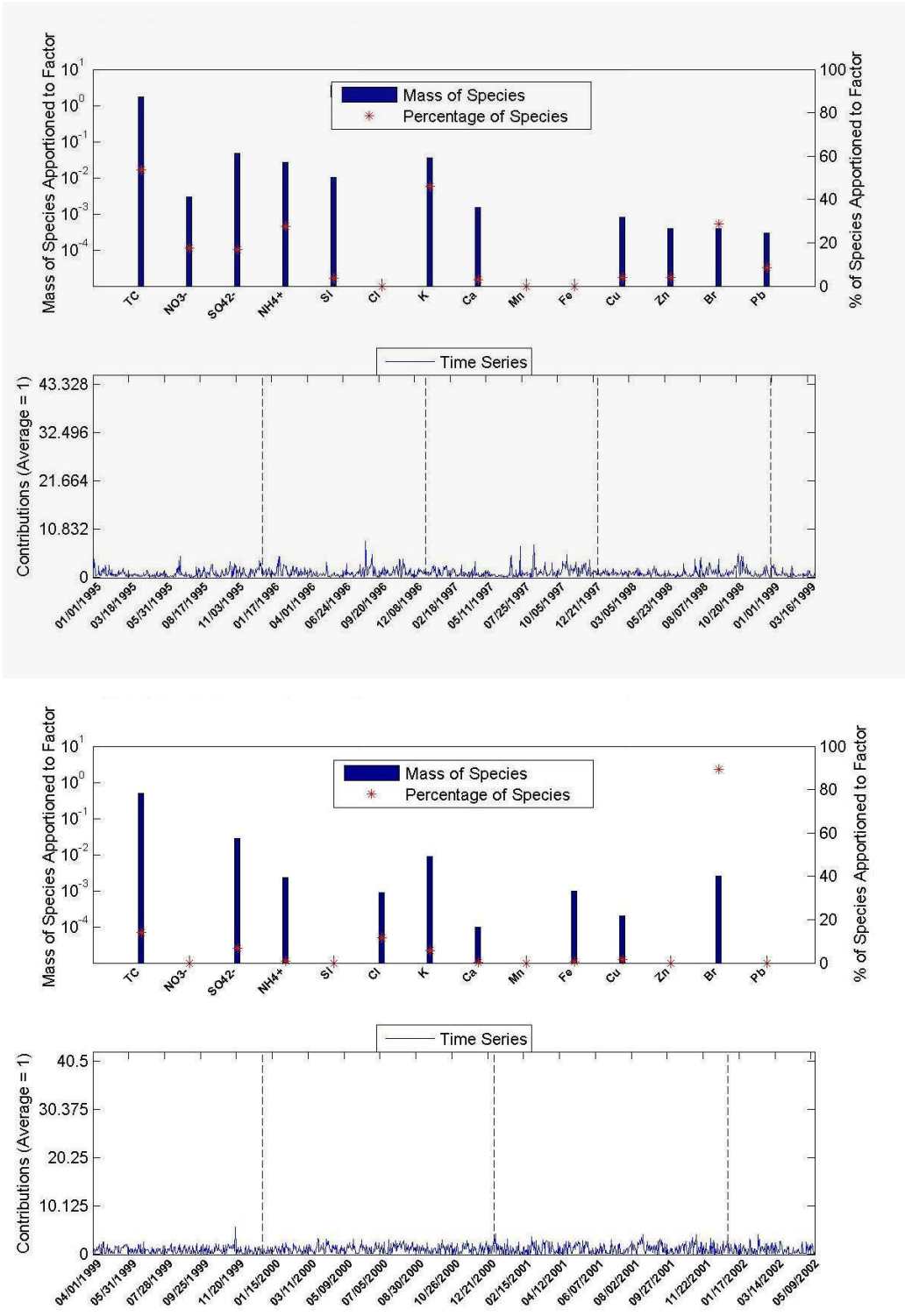


Figure 9. Biomass burning for Kevex (top) and JV (bottom) models.

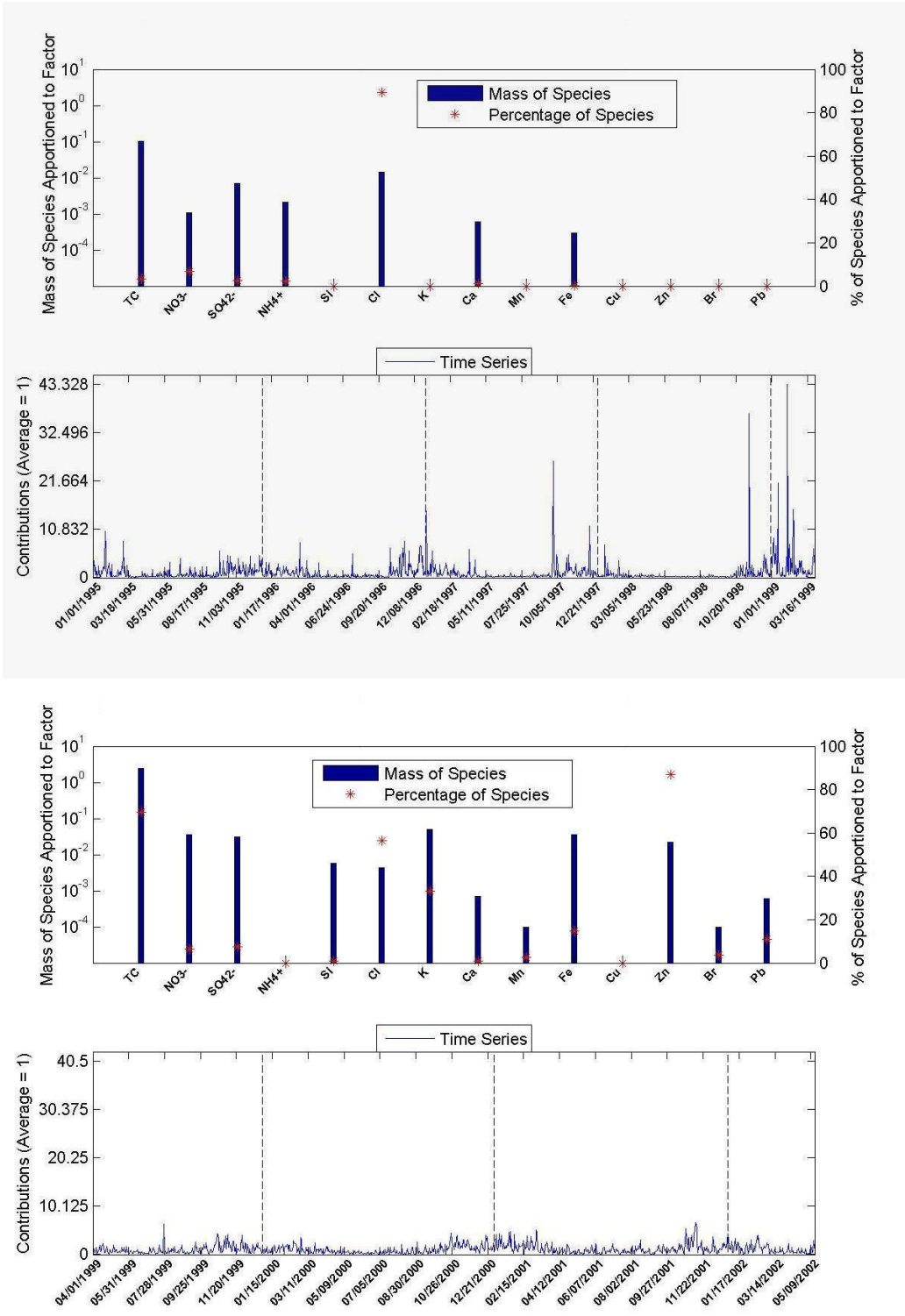


Figure 10. Cl-rich source for Kevex (top) and JV (bottom) models.

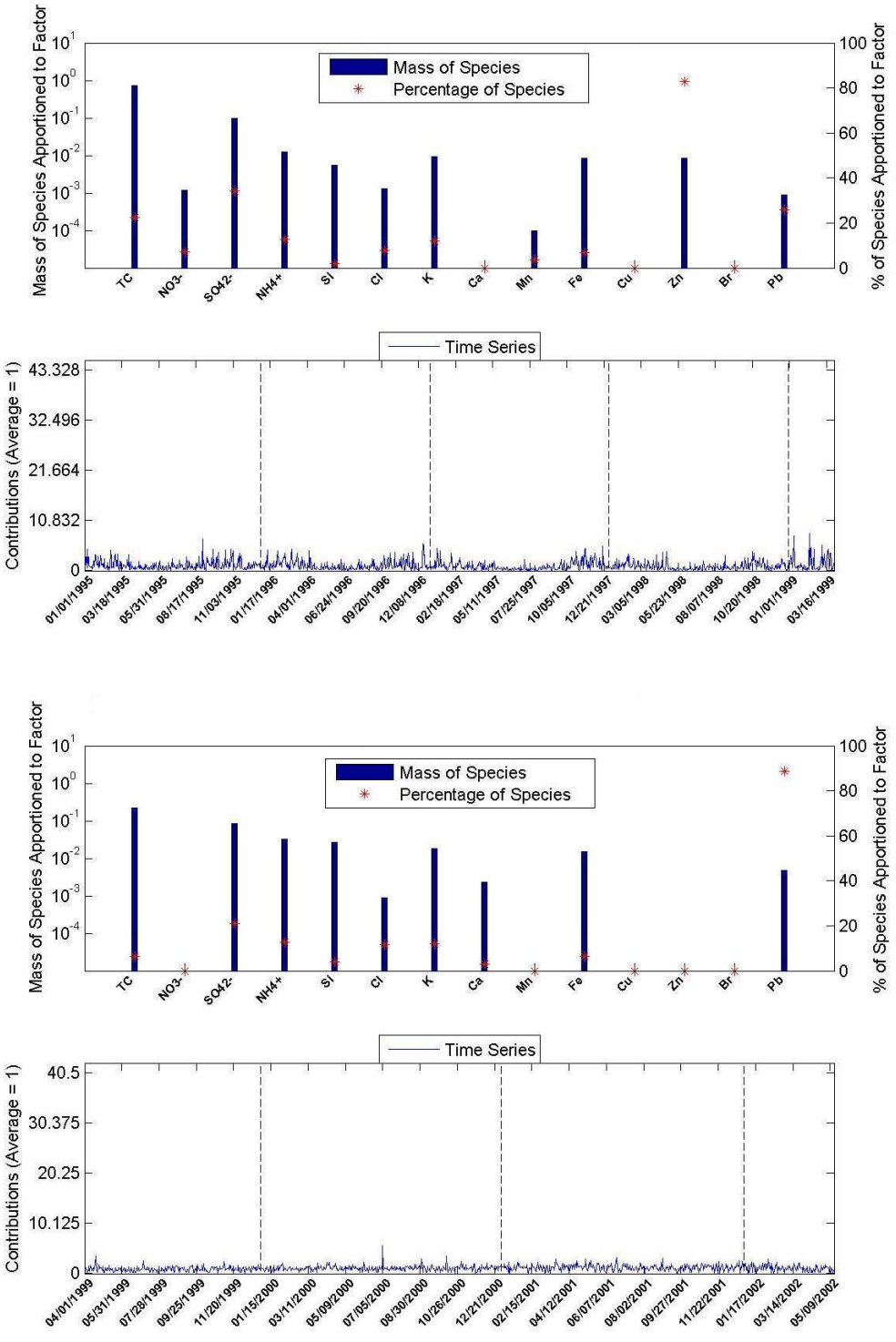


Figure 11. Vehicle exhaust for Kevex (top) and JV (bottom) models.

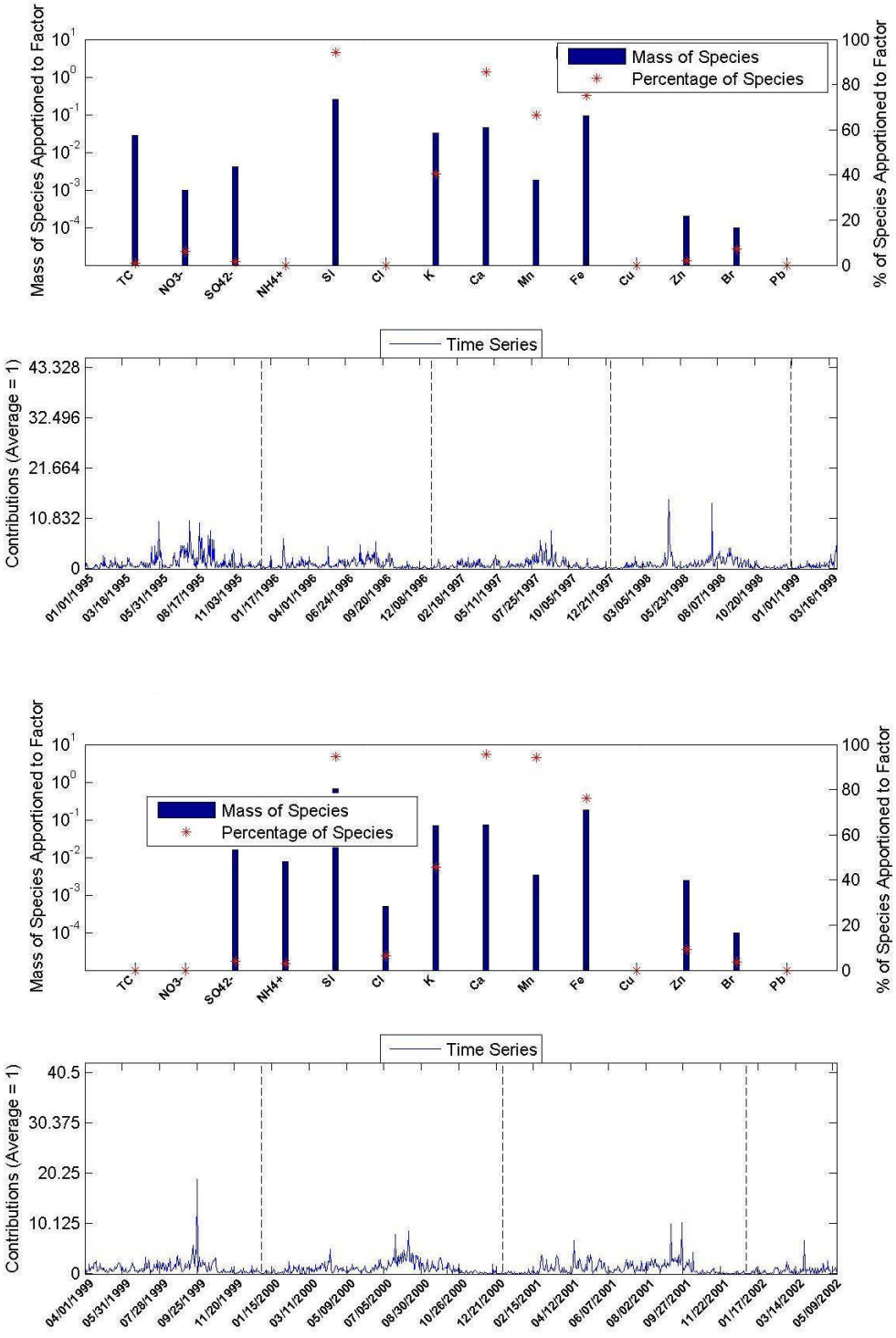


Figure 12. Airborne soil for Kevex (top) and JV (bottom) models.

CHAPTER 3

PMF-derived $PM_{2.5}$ Sources in Spokane, Washington, with emphasis on Trace Metals Analysis

This chapter is presented as a paper that has been prepared for publication in a peer-reviewed journal.

PMF-derived $PM_{2.5}$ Sources in Spokane, Washington, with emphasis on Trace Metals Analysis

Jennifer Shaltanis, Candis Claiborn
Laboratory for Atmospheric Research, Washington State University, Pullman, WA

Timothy Larson
Department of Civil Engineering, University of Washington, Seattle, WA

Correspondence:
Jennifer Shaltanis
Laboratory for Atmospheric Research
WSU/NSF IGERT Center for Multiphase Environmental Research
Department of Civil and Environmental Engineering
Washington State University
PO Box
Pullman WA 99164-2714
Phone: (509) 335-6248
Fax: (509) 335-
Email: jshaltanis@wsu.edu

Abstract

Data from the Spokane Health Effects Study were analyzed for several trace elements, using the Kevex and Jordan Valley EDX-771 energy dispersive X-ray fluorescence spectrometers (XRF) and Washington State University's Instrumental Neutron Activation Analysis (INAA). Pollution samples from 3/1/1996-9/10/2001 were amassed by Spokane Health Effects Study, currently largest of its kind, and analyzed by the trace element methods. Data were modeled to identify particulate sources with Positive Matrix Factorization (PMF), Environmental Protection Agency version 1.1. 20 species chosen for source apportionment analysis were: total carbon, NO_3^-, SO_4^{2-}, NH_4^+, As, Br, Ca, Cl, Co, Cr, Cu, Fe, K, Mn, Pb, Sb, Sc, Si, and Zn.

PMF analysis models were performed on INAA alone as well as INAA-XRF composite datasets. An initial INAA model with carbon and ionic species was not sufficient to resolve most components of the profile. INAA data were grouped with Kevex and Jordan Valley data, by date, and evaluated with PMF using five to nine sources. The analyses revealed seven identifiable sources of air pollution: airborne soil, metal processing, biomass burning, vehicle exhaust, nitrate, and Cl-rich, and Cr-rich sources. These sources agree with previous, smaller studies using Spokane Health Effects data, with exception of the Cr-rich and Pb-rich sources, features uniquely characterized by these composite groups. Bootstrap analyses clarify the strengths of these sources on the overall fine fraction load. INAA and XRF data together have provided an enhanced temporal and instrumental analysis of the Spokane, WA air pollution profile.

Keywords: PMF, source apportionment, XRF, INAA, trace elements, $PM_{2.5}$

3.1 Introduction

Fine fraction particulate matter, $PM_{2.5}$, stems from several sources of pollution, with potential to cause a variety of potential health problems (Dockery et al., 1993; 1994). Chemical composition can vary and is unique for each locale. Species can be associated with particle size and can have various influences on health. The current National Ambient Air Quality Standard (NAAQS) for fine fraction PM has 24 hour and annual limits, which are 35 and 15 $\mu g/m^3$, respectively, and are based purely on total concentration, not composition specific masses.

Health studies indicate that basing standards on mass concentration alone is insufficient to improve the atmosphere for general health (Rizzio et al., 1999; Green et al., 2002), and that health effects could be more accurately determined by PM composition. Epidemiology studies have associated increased respiratory distress with more prominent pollution particles, such as carbon and metals (Ghio et al., 1992, 1996; Norris et al., 2000; Aarnio et. al, 2005). Ecological studies have shown that particles of certain types can have varying degrees of damage to vegetation and indirect effects on climate; ecological distress alone is a serious environmental problem, but it can, in turn, have detrimental effects on human health (Dailey et al., 1997; Grantz et al., 2003). Research has turned from studying overall PM mass to the composition of PM, in an attempt to more accurately characterize what level of what type of PM is safe for human health and welfare.

Spokane, WA, is a medium sized city, nestled in the Eastern Washington hills surrounded by diverse agricultural lands with a moderate climate, and a history of

moderate air pollution. The blend of urban, industrial and farming environments in the region produce a variety of air pollution issues. Historically, Spokane has experienced episodes of $PM_{2.5}$ exceeding the EPA standards (Villasenor et al., 2001). Windblown dust during the dry summer and fall and stagnation during fall and winter, as well as continual occurrence of vehicle combustion and industrial processes, can be major contributors to elevated $PM_{2.5}$ concentrations. The Spokane Health Effects Study compiled a massive pollution and health database for the 1995-2002 period, and was intended to provide insight to the sources of air pollution and associated health effects (Kantamaneni et al., 1996; Norris, 1998; Claiborn et al., 1998; Haller et al., 1999; Norris et al., 2000; Villasenor et al., 2001; Hoffman, 2002).

Instrumental Neutron Activation Analysis, or INAA, effectively appraises particulate concentrations because of its extreme sensitivity, proven reliability, and range of elemental coverage. The low detection levels and coverage enable researchers to test for several trace elements not available with other methods. Of the various instrumental methods used in the Spokane Health Effects Database, INAA can identify 28 elements in this database. Unlike other studies performed on the Spokane dataset, INAA adds valuable details about unique species, which may increase knowledge about particulate sources. Its many benefits to particulate identification are well known; INAA has been used in numerous trace element analyses worldwide (Landsberger and Wu, 1995; Rizzio et al., 1999; Farinha et al., 2001; Suarez et al., 2002; Cao et al., 2002; Almeida et al., 2003; Bem et al., 2003).

The Spokane Health Effects study was a collaborative effort among the Mickey Leland Nation Urban Air Toxics Research Center, Washington State University,

University of Washington, Spokane County Air Pollution Control Authority (SCAPCA), the American Lung Association of Washington, and the Harvard School of Public Health. Among its extensive objectives to assess air pollution and its effects on human health, one major goal was to amass PM data, for a long period of time using several measurement techniques, in order to create an exhaustive, long-term PM dataset. The purpose of this research was to create a more complete view of the particulate pollution sources in Eastern Washington.

Previous research on the Spokane Health Effects data has involved only subsets of the study period or instrumental analyses. Building on prior work using trace analysis by X-Ray Fluorescence (XRF) and Positive Matrix Factorization (PMF), this research was intended to enhance knowledge of trace elements and particulate sources. Previous studies have successfully used INAA and XRF analyses in concert to resolve pollution data (Bradley et al., 1995; Farinha et al., 2001; Almeida et al., 2003; Graney et al, 2004). This research has used INAA, XRF, carbon, and ion analyses as part of a larger study to resolve particulate sources using a more complete piece of the vast PM dataset.

3.2 Methods

Filters were collected from January 1st, 1995- May 15th, 2002, using a Versatile Air Pollutant Sampler, or VAPS (Stevens et al., 1993; Sommerville et al., 1994). Quartz, polytetrafluoroethylene (Teflon), and Nucleopore filters were exposed for a 24 hour period, from midnight to midnight. Filters were kept at standard conditions before and after exposure in desiccators before and after sampling, and quartz filters were purified of

any carbon contamination before exposure by baking them for several hours above 800ºF. Pre- and post-exposure weights were used to calculate VAPS masses. VAPS were regularly checked for leaks, consistent flow rates, and other hardware problems; in addition, VAPS masses were compared to data collected by a collocated Tapered Element Oscillating Microbalance (TEOM) instrument (Rupprecht and Patashnick, TEOM Series 1400a) for any suspicious samples. In previous research using only XRF trace element, the entire data period was used, with replacements accounting for approximately 2.5% of samples; replacements here would be less than 2.5%.

Samples included in this study were analyzed for carbon, SO_4^{2-}, NO_3^-, NH_4^+, and trace elements. Carbon was measured with Thermal Manganese Oxidation (TMO) and Thermal Optical Transmittance (TOT), for total carbon and carbon fractions. Sulfate and nitrate were determined via ion chromatography of extracted quartz filter particles, and ammonium ions were extracted from quartz filters with citric acid and measured with colorimetry. Trace elements were analyzed by three methods, the EPA Kevex Energy Dispersive X-Ray Fluorescence, or ED-XRF, instrument in Research Triangle Park, NC, the Jordan Valley ED-XRF instrument operated at WSU Laboratory for Atmospheric Research (LAR), and Instrumental Neutron Activation Analysis through the WSU Nuclear Radiation Center (NRC).

Trace element data were taken from the measurement output and formatted for PMF analysis. Trace element data were reported in terms of mass per square centimeter or per filter and converted to mass concentrations, using the VAPS flow volume. The Kevex instrument has clearly defined minimum detection limits (MDLs), as reported by Kellogg (1994), as well as uncertainties. Uncertainties for INAA were published with

concentration, but MDLs were not clearly defined. MDLs are often based on instrumental method and the general ability of the instrument to discern concentration peaks from the background noise. JV data were not reported with MDLs or uncertainties; therefore, the average blank concentrations, per specie, were taken as the species MDLs, and uncertainties were based on these MDLs and the specie standard deviations of the samples, according to the following:

$$E_x = \sigma_x + \frac{1}{3} MDL_x \quad (1)$$

Blanks are subjected to the same conditioning and are handled with the same level of care as exposed filters. Salma and Zemplén-Papp (1999) used more detailed criteria for determining MDLs; however, because the elements chosen for this study have strong abundance frequencies coupled with reasonable uncertainties to be used in PMF, a simpler, blanket rule is used in this study.

Positive Matrix Factorization (PMF) was used to model particulate sources of air pollution using the Spokane Health Effects Database. The version 1.1 released by United States Environmental Protection Agency is a user-friendly form, based on the original script version accepted in the research community in the last decade. Its algorithm is based on the multi-linear engine version of PMF, originally released in the 1990s (Paatero, 1997; Huang et al., 1999; Lee et al., 1999; Xie et al., 1999; Ramadan et al., 2000; Willis, 2000; Polissar et al., 1998; Song et al., 2001; Henry, 2002; Paatero et al., 2002; Qin et al., 2002; Kim et al., 2003). The premise behind PMF is that factors derived by the model can effectively explain the variability of data in terms of contaminant

trends; essentially the factors are specific, identifiable sources of pollution. EPA's version supplements the proven method by adding on external controls to speed up the iterative model process: by accepting several formats of data, providing graphical-user interface selection of inputs, and adding post-analysis graphical and uncertainty bootstraps. Its default settings are designed to treat the dynamic nature of environmental data. For more information, the product manual by Eberly (2005) discusses the model algorithms and advanced settings.

3.3 Results and Discussion

3.3.1 INAA Results

Preliminary PMF models were run on a dataset including total carbon, SO_4^{2-}, NH_4^+, NO_3^-, and INAA data, including trace concentrations, in order to determine whether INAA data alone can support a conclusive PMF model or if XRF data are also necessary. The dataset consisted of particles collected from 3/1/1996-9/10/2001, for a total of 1914 samples, for 21 species, or approximately 95% of available dates. One of many EPA PMF features is the ability to alter uncertainty between initial data construction and modeling. Classifications of data as "weak" will downweight the data by a factor of 3. This choice was selected for many elements which had poor signal to noise ratios. Because many of these trace element exist in very small quantities, quantifying the potential for error is especially important (Rizzio et al., 2000; Farinha et al., 2001).

Selection of elements subject to apportionment modeling was based on abundance frequency, which is defined as the recurrence of samples above instrumental detection,

per specie. Carbon, SO_4^{2-}, NO_3^-, and NH_4^+ are generally much higher in concentration than trace element and have consistent abundance throughout the database. For trace element, inclusion was based on **Figure 1**. This plot shows the fraction of samples above detection level, per specie. An abundance criterion of 50% was set as the lower threshold for inclusion. While some of these less frequent elements could have significant impacts, the level of uncertainty imposed by instrumental detection limits and ability to separate specie peaks from background noise was considered too influential on the final profile.

The sources found by INAA alone shows some agreement to previous studies on this dataset (Hoffman 2002, Kim et al., 2003), but are incomplete due to absence of key XRF species not available through INAA. The profile also shows unique features not seen before with this data. For sake of article space, these results are presented only in textual summary, not graphically. The INAA modeled data resolved airborne soil, vehicle exhaust, nitrate, and possibly biomass burning. Unique factors included a somewhat ambiguous factor, with chromium as a cornerstone and other species showing minimal specie contribution. Nearly all species were present in this source, most notably sodium; this proposed factor appears as a Na-rich source. In other studies, Na has been tied to a sea salt source, not reasonable for a region 250 miles from the coast (Xie et. al, 1999). Instead, this may actually be systematic contamination from skin contact on the filter.

3.3.2 INAA-XRF Results

With preliminary INAA models and previous XRF studies as guides, PMF models were run on a composite data set containing INAA and XRF data, and SO_4^{2-}, NO_3^-, NH_4^+, and TC, in attempt to identify the INAA and XRF species important to the

Spokane particulate profile. The addition of XRF data helped to clarify whether species were legitimate source components or erroneous and negligible. Initial model runs for both Kevex and JV data were used to determine individual significance of certain elements. The questionable Na from the INAA model showed up ubiquitously in the XRF-INAA profiles, which was sufficient to declare Na erroneous. No other species were deemed as such, so the overall species coverage for the INAA-XRF models are: TC, NO_3^-, SO_4^{2-}, NH_4^+, As, Br, Ca, Cl, Co, Cr, Cu, Fe, K, Mn, Pb, Sb, Sc, Si, and Zn.

INAA and XRF instruments overlap on key species – notably Fe, Br, Zn, K, and As. A view of INAA and XRF agreement is seen in **Figure 2**. These plots represent corresponding samples of Br, K, Fe, and Zn. For Fe and Zn, the instrumentation concurs on relative terms. Br and K have more scatter; this disparity could be due to abundance levels. Recall from **Figure 1** that Zn and Fe are very abundant for all three instruments. For K, INAA is below 50% in abundance frequency while the XRF instruments show good detection of these elements; the opposite is true for Br. JV has known problems with certain elements. A companion study carried out alongside this research investigated XRF-produced sources, with consideration for instrumental step changes midway through the 1995-2002 period, and detector degradation in the JV instrument affecting the latter portion of the dataset. Those issues will not be addressed here; rather JV and Kevex data will be treated separately, assuming JV corrections are sufficient for PMF modeling. Of these common elements, the instrument with the greatest abundance in each was the selected data for PMF modeling.

Current and previous XRF PMF models have found 6-7 sources, and the inclusion of INAA data have supplemented these factors by supplying additional elemental details,

as well as introducing an additional source. Previous XRF studies have determined the following particulate profile: airborne soil, biomass burning, vehicle exhaust, a Cl rich source, some form of nitrate and/or sulfate, and metal processing (Kim et al., 2003). These INAA-XRF sources are given in that order in **Figures 3**-**8**. With inclusion of INAA data, the composite INAA-Kevex model derived 7 sources, adding a unique Cr-rich source. The Cr has small overall contribution, is episodic, and peaks in 1996, 1998, and 2000, seen in **Figure 9**. INAA elements, Co and Sc, had abundance frequencies of 87% and 93% but did not have any influence on the source profile. INAA data added valuable details, providing further explanation of previously determined sources, highlighting new features, and dismissing other elements as insignificant.

Arsenic and Sb are common in the burning of fossil fuels and are often linked to colder seasons (Beceiro-González et. al, 1997; Tsai et. al, 2003). The INAA-Kevex shows a nice spread of As and Sb over the vegetative, vehicle sources, and nitrate, all of which are due to combustion (seen in **Figures 4**, **5**, and **7**, respectively). For the INAA-JV model, most of these species are tied into the biomass burning source, with a trace amount of Sb found in airborne soil. A large amount of the vegetative and nitrate contributions occur during fall and winter, corresponding to the association with cool weather.

Bromine is often associated with burning of fuels and wear on the vehicle parts (Shendell and Naeher, 2002; Pekney et. al, *In Press*). For the INAA-Kevex model, Br is tied equally to the three fuel combustion sources — vegetation, vehicles, and nitrate (**Figures 4**, **5**, and **7**), while the INAA-JV source associates the Br to airborne soil,

biomass burning, nitrate, and vehicles (**Figures 3**,**4**,**5**, and **7**). Unlike As and Sb, Br is not known to have a seasonal appearance.

The addition of Cr signifies a potentially new industrial process for Spokane which adds to the $PM_{2.5}$ load. Cr is associated with metal smelting, either through the use of high temperature cutting or molding of metallic alloys. The Cr source is seen in **Figure 9**. Included in this source are the time series contributions for each dataset. One can see the distinctive maxima over time, indicating episodes of elevated Cr. The irregularity supports the general association between Cr and specialty metal processing. Even though Ni was not included in this overall study, a small PMF model was run again, for dates during which Ni was above detection, in order to determine any association between Ni and this Cr-source. Like Cr, Ni is a metal which can be tied to specialty metal processing. No appreciable relationship was observed between the two elements or Ni with the source feature.

Of the two major forms, Cr(III) and Cr(VI), hexavalent chromium can be deadly to humans. While trace element analysis does not delineate between the two forms, overall Cr content may be indicative of the dangerous form. Work by Talebi (2003) suggests Cr(VI) can account for as much as 25% of overall Cr in industrial settings. Considering the episodic nature of this Cr-rich source, the level of Cr(VI) may also be sporadically high as well.

Because regulatory agencies are especially concerned with $PM_{2.5}$ when it approaches or breaches the primary standard, a small PMF model was run for the top fifth (by mass) samples, to determine if the samples most concerning to control agencies reflect what is given by all data. For the INAA-Kevex data, the approximately two

hundred samples very closely resembled the source features given in this source profile. For the INAA-JV data, most of the sources were identifiable, with soil being represented by two distinctive features, while confusing the signal between Cl-rich and combustion. In spite of this limitation, these results indicate the high level days can represent the overall $PM_{2.5}$ impacts and all source features would be of concern to regulatory agencies.

3.3.3 Tying in Sources to $PM_{2.5}$ Mass

Confidence intervals were established by using $PM_{2.5}$ mass in a second run of the optimum PMF model and reviewing the bootstrapped uncertainties for this model. $PM_{2.5}$ masses were downweighted by increasing their uncertainties by a factor of at least 30 (essentially, these points would have no effect on the source profile); the bootstrap program in EPA's version 1.1 determines uncertainty in the model result by looking at variances in matching random starting point runs to the original model. Ideally the greater the variation explained, with minimal spread in the random runs, the more "accurate" the model. The 95^{th} percentile was taken as the performance measure for this uncertainty assessment; data are found in **Table 1**. Samples corresponding to the top fifth highest masses are given in **Table 2**. These data would represent the days of most concern to regulatory agencies. The first numbers listed are the bootstrapped variations, in $\mu g/m^3$ for each source, to show how accurately the sources represent $PM_{2.5}$. The proceeding numbers in parentheses represent the order of least to greatest variation by source, to show how the sources fine fraction estimates compare to one another. Data are roughly divided into four 500 point groups, two each of Kevex and JV data, by date, in order to show any subtle time or instrument effects not found in the larger time period models.

Cr- and Cl-rich sources are marked by episodes of increased $PM_{2.5}$ due largely to single species. For chromium and chlorine, the $PM_{2.5}$ mass varies over time and instrument, in response to single, intense peaks in the source profiles, individual events that greatly affect the general PM representation. Lower rankings among the four data groups correspond to periods in the data where these sources are not especially active and their PM contributions do not vary greatly; likewise, higher variability is in response to periods where background levels are intermingled with days of increased chromium and chlorine. Additionally, Cl provided by the Jordan Valley method is problematic. The detector suffered degradation, which affected its ability to accurately measure Cl; along with the poor detector, the JV instrument does not have a calibration for Cl, which inhibits it ability to effectively measure it in the presence of other trace elements. Based on temporal trends, a lot of the mass attributed to the Cl-rich source in the JV bootstrap data (**Table 1**) may be due to vehicle exhaust and biomass burning, given the correlation between the Cl-rich source and Zn (**Table 3**).

Soil, nitrate, and metal processing are fairly consistent among instruments and time, and represent medium variations in $PM_{2.5}$. Metal processing has a slightly lower average variation because its peak events are dependent on only one major specie that does not have a large range of concentrations. Geological elements in soil and the ions and carbon in nitrate can vary more significantly in concentration, affecting the $PM_{2.5}$ mass.

Vegetation and vehicle exhaust are the most variable due to the nature of their formation and transport. Each of these combustion sources can comprise a variety of sites, which introduces a chance of contamination by different species of the combustion

material (wood, grass, or other fuel sources), or gradual changes in fuel composition (such as decrease of Pb in gasoline). Additionally, because these sources can be transported from many locations, entrainment of other species can slightly alter the profiles. This issue may be more prominent here than in the airborne soil profile, because in the airborne soil, the geological elements exist in higher concentrations and are less effected by subtle invasions.

3.4 Conclusions

Trace element analyses are essential to understanding particulate pollution. More concentrated particulate matter, such as carbon and ions, only tell part of the story when trying to determine what causes the $PM_{2.5}$ to occur. The crustal elements, as a group, help determine airborne soil, overall a small contribution to $PM_{2.5}$, but with strong influence during the summer. In spite of some instrumental limitations, Pb, Cl, and Cu differentiate carbon rich sources into very different and significant anthropogenic pollution sources, vehicles, municipal incineration, and metal processing. Similarly, Cr has proven to be a singly significant identifier of a new source, most likely a specialty metal processing plant, and its irregular emissions. Sb and As help discern mostly wintertime combustion, nitrate, and biomass burning. While many of these elements are found only in trace amounts, they can be diverse in origin and physiologically potent. The use of coupled trace element analyses in source apportionment is a valuable tool in capturing the full fine particulate landscape.

Future work with this data will include breaking down the carbon into organic and elemental form, for the time period using the TOT instrument. The added carbon

information will help in clarifying and separating agricultural, vehicle, and industrial combustion into more specific sources. The addition of detailed carbon information will be a final step in understanding the fine fraction particulate pollution in Spokane, WA.

Acknowledgements

This study was funded in part by the Integrative Graduate Education and Research Training (IGERT) grant from the National Science Foundation to Washington State University under grant DGE-0072817, the Washington State Department of Ecology, Mickey Leland Nation Urban Air Toxics Research Center, University of Washington/EPA Northwest Research Center for Particulate Matter and Health Center (R827355) and University of Rochester/EPA Particulate Matter and Health Center (R827354). This research has not been subjected to EPA's required peer and policy review and does not necessarily reflect the views of EPA.

References

Aarnio, P., Yli-Tuomi, T., Kousa, A., Mäkelä, T., Hirsikko, A., Hämeri, K., Räisänen, M.,Hillamo, R., Koskentalo, T., Jantunen, M., 2005. The concentrations and composition of and exposure to fine particles ($PM_{2.5}$) in the Helsinki subway system. Atmospheric Environment, 5059-5066.

Almeida, S.M., Reis, M.A., Freitas, M.C., Pio, C.A., 2003. Quality assurance in elemental analysis of airborne particles. Nuclear Instruments & Methods in Physics Research B: Beam Interactions with Materials & Atoms 207, 434-446.

Beceiro-González, E., González-Soto, E., López-Mahía, Prada-Rodríguez, D., 1997. Total arsenic and selenium levels in atmospheric particulate matter of La Coruña (Spain). The Science of the Total Environment 208, 207, 211.

Bem, H., Gallorini, M., Rizzio, E., Krzeminska, M. 2003. Comparative studies on the concentrations of some elements in the urban air particulate matter in Lodz City of Poland and in Milan, Italy. Environment International 29, 423-428.

Bradley, D.A., NG, K.H., Green, S., Mountford, P.J., Shukri, A., Evans, J. Applications

for XRF, NAA, and low-kV radiographic techniques in the study of body composition and diseased tissue. Radiation Physical Chemistry 47:5, 745-749.
Cao, L., Tian, W., Ni, B., Zhang, Y., Wang, P., 2002. Preliminary study of airborne particulate matter in a Beijing sampling station by instrumental neutron activation analysis. Atmospheric Environment 36, 1951-1956.
Claiborn, C.S., Lamb, B.K., Miller, A., Beseda, J., Clode, B., Vaughan, J., Kang, L., Newvine, C., 1998. Regional measurements and modeling of windblown agricultural dust: the Columbia Plateau PM10 program. Journal of Geophysical Research 103, 19753–19767.
Dailey, G.C., Introduction: what are ecosystem services? In: Dailey, G.C., editor. Nature's Services: societal dependence on natural ecosystems. Washington (DC): Island Press; 1997, 1-10.
Dockery, D.W., Pope, C.A., III, Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air pollution and mortality in six U.S. cities. New England Journal of Medicine 329, 1753-1759.
Dockery, D.W., Pope, C.A., III, 1994. Acute respiratory effects of particulate air pollution. Annual Review of Public Health 15, 107-132.
Eberly, S. 2005. EPA PMF 1.1 User's Guide. US Environmental Protection Agency.
Farinha, M.M., Freitas, M.C., Almeida, S.M., Reis, M.A., 2001. Some improvements in air particulate matter analysis by INAA. Radiation Physics and Chemistry 61, 659-661.
Ghio, A.J., Kennedy, T.P., Whorton, R., Crumbliss, A.L., Hatch, G.E., Hoidal, J.R., 1992. Role of surface complexed iron in oxidant generation and lung inflammation induced by silicates. American Journal of Physiology 263, L511–L518.
Ghio, A.J., Stonehuerner, J., Pritchard, R.J., Piantadosi, C.A., Quigley, D.A., Dreher, K.L., Costa, D.L., 1996. Humic-like substances in air pollution particulates correlate with concentrations of transition metals and oxidant generation. Inhalation Toxicology 8, 479–494.
Gold, D.R., Damokosh, A.I., Pope. C.A., III, Dockery, D.W., McDonnell, W.F., Serrano, P., Retama, A., Castillejos, M., 1999. Particulate and ozone pollutant effects of the respiratory function of children in southwest Mexico City. Epidemiology 10, 8-16.
Graney, J.R., Landis, M.S., Norris, G.A., 2004. Concentrations and solubility of metals from indoor and personal exposure $PM_{2.5}$ samples. Atmospheric Environment 38, 237-247.
Grantz, D.A., Garner, J.H.B, Johnson, D.W., 2003. Ecological effects of particulate matter. Environment International 29, 213-239.
Green, L.C., Crouch, E.A., Ames, M.R., Lash, T.L., 2002. What's Wrong with the National Ambient Air Quality Standard (INAAQS) for Fine Particulate Matter ($PM_{2.5}$)? Regulatory Toxicology and Pharmacology 35, 327-337.
Haller, L., Claiborn, C., Koenig, J., Larson, T., Norris, G., and Edgar, R., 1999, Airborne particulate matter size distributions in an arid urban area. Journal of the Air & Waste Management Association 49, 161-168.
Henry, R.C., 2002. Multivariate receptor models – current practice and future trends. Chemometrics and intelligent laboratory systems 60, 43-48.

Hoffman, M. (2002). Elemental analysis and receptor modeling of airborne particulate matter in Spokane, Washington, M.S. thesis, Washington State University, Pullman, WA.

Huang, S., Rahn, K.A., Arimoto, R., 1999. Testing and optimization two factor-analysis techniques on aerosol at Narragansett, Rhode Island. Atmospheric Environment 33, 2169– 2185.

Kantamaneni, R., Adams, G., Bamesberger, L., Allwine, E., Westberg, H., Lamb, B., Claiborn, C.S., 1996. The measurement of roadway PM10 emission rates using atmospheric tracer ratio techniques. Atmospheric Environment 30 (24), 4209–4223.

Kellogg, R.: 1994, Standard Operating Procedure for the Analysis of Aerosols on Filters with the Source Apportionment Research Branch Kevex Energy Dispersive X-ray Fluorescence Spectrometer, 39 p. ManTech Environmental Technology, Inc., Research Triangle Park, NC.

Kim, E., Larson, T.V., Hopke, P. K., Slaughter, C., Sheppard, E., Claiborn, C., 2003. Source identification of PM2.5 in an arid northwest U.S. city by positive matrix factorization. Atmospheric Research 66, 291-305.

Landsberger, S., Wu, D., 1995. The impacts of heavy metals from environmental tobacco smoke on indoor air quality as determined by Compton suppression neutron activation analysis. The Science of the Total Environment 173/174, 323-337.

Lee, E., Chan, C.K., Paatero, P., 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmospheric Environment 33, 3201-3212.

Norris, G., Larson, T., Koenig, J., Claiborn, C., Sheppard, L., 2000. Asthma aggravation, combustion, and stagnant air. Thorax 55, 466-470.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and intelligent laboratory systems 37, 23-35.

Paatero, P. Hopke, P.K., Song, X.-H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. Chemometrics and intelligent laboratory systems 60, 253-264.

Pekney, N.J., Davidson, C.I., Bein, K.J., Wexler, A.S., Johnston, M.V. Identification of sources of atmospheric PM at the Pittsburgh supersite, Part I: Single particle analysis and filter-based positive matrix factorization. Atmospheric Environment *In Press.*

Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: 1. Elemental composition and sources. Journal of Geophysical Research 103 (D15), 19045– 19057.

Qin, Y., Oduyemi, K., Chan, L.Y., 2002. Comparative testing of PMF and CFA models. Chemometrics and Intelligent Laboratory Systems 61, 75-87.

Ramadan, Z., Song, X.H., Hopke, P.K., 2000. Identification of sources of Phoenix aerosol by positive matrix factorization. Journal of the Air & Waste Management Association 50, 1308–1320.

Rizzio, E., Giaveri, G., Arginelli, D., Gini, L., Profumo, A., Gallorini, M., 1999. Trace elements total content and particle sizes distribution in the air particulate matter of a rural-residential area in north Italy investigated by instrumental neutron activation analysis. The Science of the Total Environment 226, 47-56.

Rizzio, E., Giaveri, G., Gallorini, M., 2000. Some analytical problems encountered for trace elements determination in the airborne particulate matter of urban and rural areas. The Science of the Total Environment 256, 11-22.

Salma, I., Zemplén-Papp, É., 1999. Instrumental neutron activation analysis for studying size-fractionated aerosols. Nuclear Instruments & Methods in Physics Research 435, 462-474.

Shendell, D. G., Naeher, L.P., 2002. A pilot study to assess ground-level ambient air concentrations of fine particles and carbon monoxide in urban Guatemala. Environment International 2002, 375-382.

Somerville, M., Mukerjee, S., Fox, D., and Stevens, R.K. 1994. Statistical approaches in wind sector analyses for assessing local source impacts. Atmospheric Environment 28, 3483-3493.

Stevens, R.K., Pinto, J., Mamane, Y., Ondov, J., Abdulraheem, M., Al-Majed, N., Sadek, M., Cofer, W., Elenson, W., and Kellogg, R., 1993. Chemical and physical properties of emissions from Kuwaiti oil fires. Water Science &Technology, 27 7-8, 223-233.

Song, X-H, Polissar, A. V., Hopke, P.K., 2001. Sources of fine particle composition in the northeastern US. Atmospheric Environment 35, 5277-5286.

Suarez, A.E., Ondov, J.M., 2002. Ambient Aerosol Concentrations of Elements Resolved by Size and by Source: Contributions of Some Cytokine-Active Metals from Coal- and Oil-Fired Power Plants. Energy & Fuels 16, 562-568.

Talebi, S.M., 2003. Determination of total and hexavalent chromium concentrations in the atmosphere of the city of Isfaham. Environmental Research 92, 54-56.

Tsai, Y.I,, Kuo, S-C., Lin, Y-H., 2003. Temporal characteristics of inhalable mercury and arsenic aerosols in the urban atmosphere in southern Taiwan. Atmospheric Environment 37, 3401-3411.

Villasenor, R., Claiborn, C., Lamb, B., O'Neill, S., 2001. Mesoscale modeling of wintertime particulate matter episodes in eastern Washington, USA. Atmospheric Environment 35, 6479-6491.

Willis, R.D., 2000. Workshop on UNMIX and PMF as applied to PM2.5. EPA 600-A-00-048.

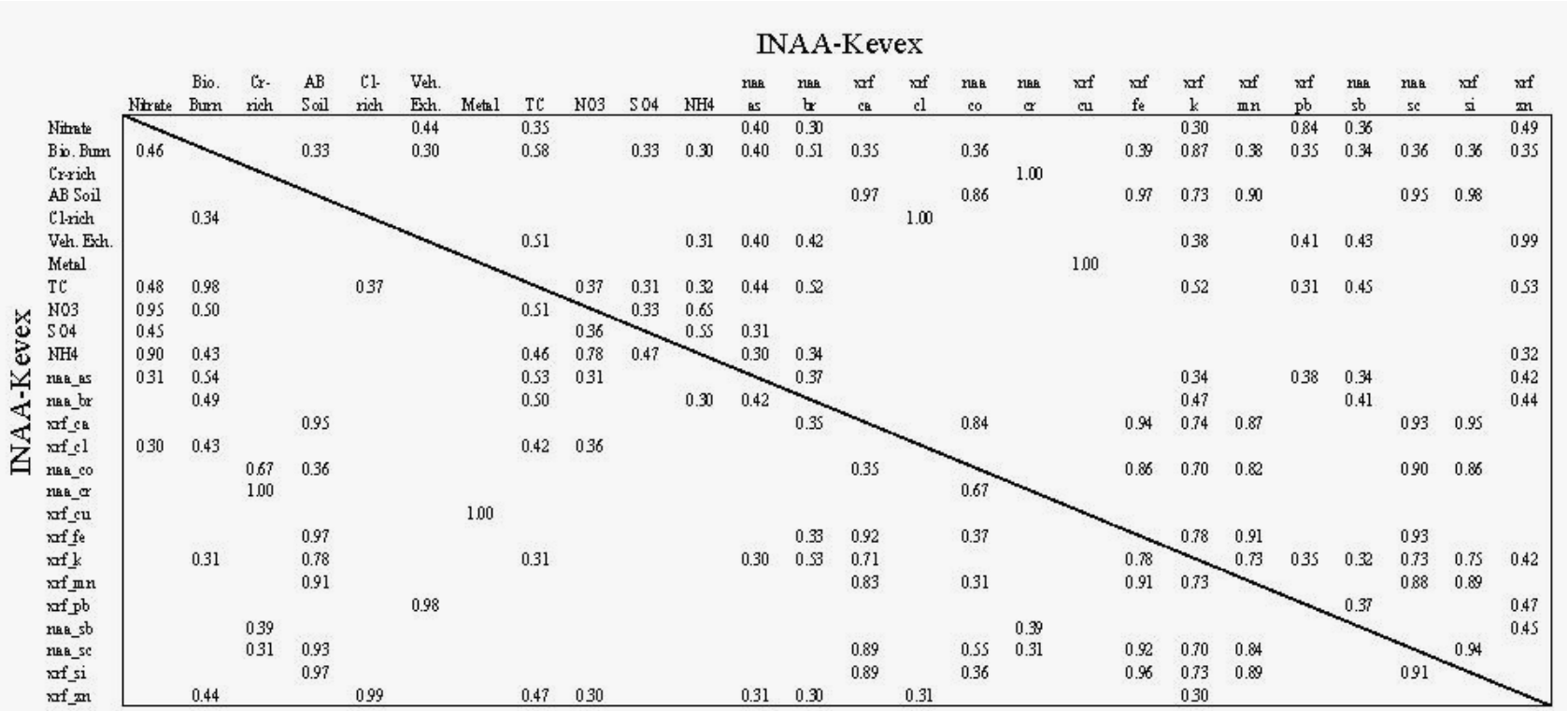
Xie, Y.-L., Hopke, P.K., Paatero, P., Barrie, L.A., Li, S.-M. 1999. Identification of source nature and seasonal variations of Arctic aerosol by the multi-linear engine. Atmospheric Environment 33, 2549-2562.

$PM_{2.5}$ Source	INAA-Kevex Data		INAA-Jordan Valley Data	
	3/1/96-9/29/97	9/30/97-3/31/99	4/1/99-6/23/00	6/24/00-9/10/01
Airborne Soil	0.74 / 7.80 (0.31; 0.85)	0.16 / 1.50 (0.00; 0.35)	1.63 / 20.0 (1.41; 1.89)	1.63 / 15.4 (1.41; 1.75)
Nitrate	0.83 / 8.70 (0.66; 0.98)	2.80 / 25.1 (2.39; 2.96)	0.81 / 9.90 (0.53; 0.91)	2.55 / 24.3 (2.02; 3.00)
Cl-rich	0.75 / 7.80 (0.64; 0.84)	0.16 / 1.40 (0.13; 0.17)	3.35 / 41.1 (2.83; 3.78)	0.55 / 5.30 (0.47; 0.92)
Metal processing	0.77 / 8.10 (0.68; 0.86)	0.37 / 3.40 (0.29; 0.42)	0.25 / 3.10 (0.17; 0.44)	0.25 / 2.40 (0.16; 0.60)
Biomass burning	2.29 / 24.0 (2.14; 2.54)	5.18 / 46.6 (5.04; 5.32)	0.34 / 4.20 (0.24; 0.44)	4.67 / 44.3 (3.71; 5.15)
Vehicle exhaust	4.02 / 42.1 (3.66; 4.36)	2.38 / 21.4 (2.21; 2.50)	1.75 / 21.5 (1.53; 2.51)	0.00 / 0.00 (0.00; 0.08)
Cr-rich	0.16 / 1.60 (0.10; 0.20)	0.08 / 0 .70 (0.00; 0.12)	0.01 / 0.20 (0.00; 0.04)	0.87 / 8.30 (0.64; 1.44)
Mass Total	**9.56**	**11.13**	**8.14**	**10.52**

Table 1. Mean mass contributions and 5 and 95 % bootstrap confidence intervals for each source, for four time subsets of XRF data. Leading numbers are mean contributions, for $PM_{2.5}$ for each source, in $\mu g/m^3$, as well as per cent of total source mass; parenthetic numbers are 5 and 95% confidence intervals, also in $\mu g/m^3$.

$PM_{2.5}$ Source	INAA-Kevex	INAA-JV
Airborne soil	1.00 / 4.82 (0.50; 0.1.15)	1.54 / 8.933 1.33; 7.71 (0.91; 2.20) (0.68; 1.55)
Nitrate	5.90 / 28.4 (5.41; 6.38)	2.09 / 12.1 (1.71; 4.07)
Cl-rich	1.56 / 7.51 (1.05; 1.69)	/ /
Metal processing	0.54 / 2.60 (0.38; 0.60)	0.71 / 4.12 (0.53; 1.39)
Biomass burning	4.84 / 23.3 (4.59; 5.21)	3.92 / 22.7 (2.36; 5.30)
Vehicle exhaust	5.88 / 28.3 (5.29; 6.32)	7.77 / 45.0 (6.41, 8.42)
Cr-rich	1.04 / 5.01 (0.93; 1.20)	1.22 / 7.07 (0.60; 1.48)
Mass Total	**20.76**	**17.25**

Table 2. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset, for the upper fifth mass fraction. Leading numbers are given as mean $PM_{2.5}$ mass / per cent of total source mass for each feature, in $\mu g/m^3$; parenthetic numbers are 5% and 95% bootstrap confidence intervals, respectively.



	Nitrate	Bio. Burn	Cr-rich	AB Soil	Cl-rich	Veh. Exh.	Metal	TC	NO3	SO4	NH4	naa as	naa br	xrf ca	xrf cl	naa co	naa cr	xrf cu	xrf fe	xrf k	xrf mn	xrf pb	naa sb	naa sc	xrf si	xrf zn
Nitrate						0.44		0.35				0.40	0.30							0.30		0.84	0.36			0.49
Bio. Burn	0.46			0.33		0.30		0.58		0.33	0.30	0.40	0.51	0.35		0.36			0.39	0.87	0.38	0.35	0.34	0.36	0.36	0.35
Cr-rich																	1.00									
AB Soil														0.97		0.86			0.97	0.73	0.90			0.95	0.98	
Cl-rich		0.34													1.00											
Veh. Exh.								0.51			0.31	0.40	0.42							0.38		0.41	0.43			0.99
Metal																		1.00								
TC	0.48	0.98			0.37				0.37	0.31	0.32	0.44	0.52							0.52		0.31	0.45			0.53
NO3	0.95	0.50						0.51		0.33	0.65															
SO4	0.45								0.36		0.55	0.31														
NH4	0.90	0.43						0.46	0.78	0.47		0.30	0.34													0.32
naa_as	0.31	0.54						0.53	0.31				0.37							0.34		0.38	0.34			0.42
naa_br		0.49						0.50			0.30	0.42								0.47			0.41			0.44
xrf_ca				0.95									0.35			0.84			0.94	0.74	0.87			0.93	0.95	
xrf_cl	0.30	0.43						0.42	0.36																	
naa_co			0.67	0.36										0.35					0.86	0.70	0.82			0.90	0.86	
naa_cr			1.00													0.67										
xrf_cu							1.00																			
xrf_fe				0.97									0.33	0.92		0.37				0.78	0.91			0.93		
xrf_k		0.31		0.78				0.31				0.30	0.53	0.71					0.78		0.73	0.35	0.32	0.73	0.75	0.42
xrf_mn				0.91										0.83		0.31			0.91	0.73				0.88	0.89	
xrf_pb						0.98																	0.37			0.47
naa_sb			0.39														0.39									0.45
naa_sc			0.31	0.93										0.89		0.55	0.31		0.92	0.70	0.84				0.94	
xrf_si				0.97										0.89		0.36			0.96	0.73	0.89			0.91		
xrf_zn		0.44			0.99			0.47	0.30			0.31	0.30		0.31					0.30						

Table 3. Pearson correlations for INAA-Kevex and INAA-JV data, for sources and species. The left half represents values for INAA-JV, while the right half represents INAA-Kevex.

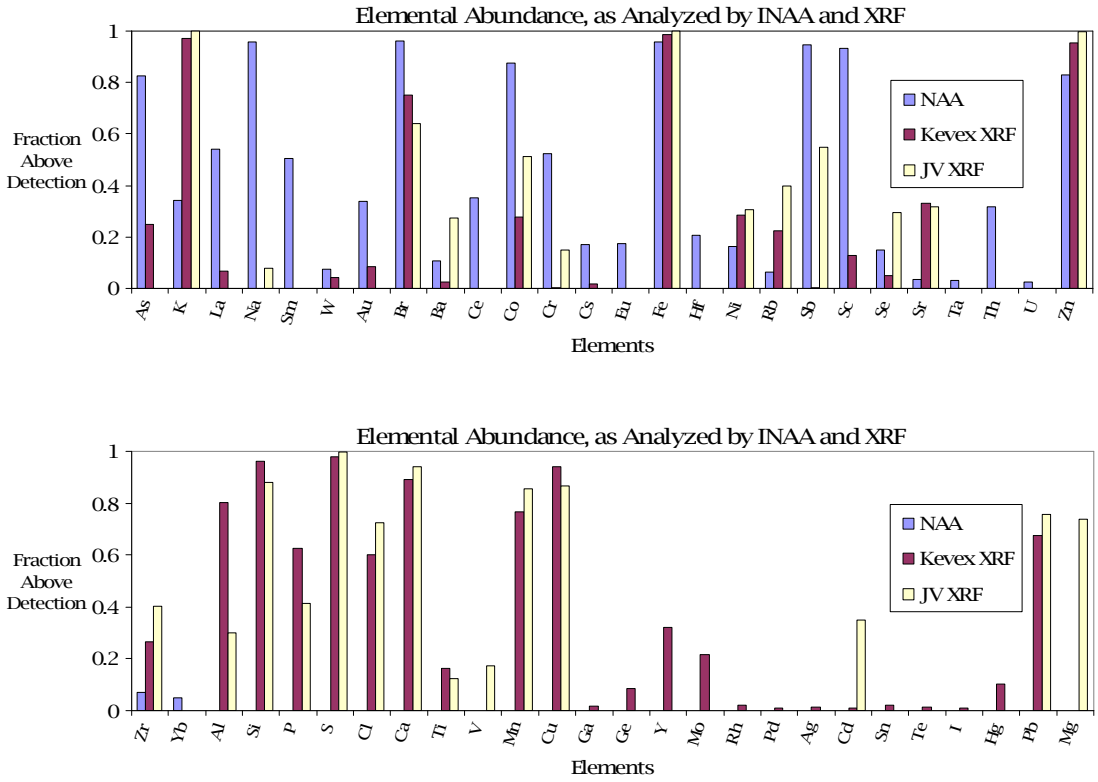


Figure 1. Elemental abundance frequency for INAA, Kevex XRF, and JV XRF. Plots represent what fraction of data, per specie, are above detection for the respective instrumental method.

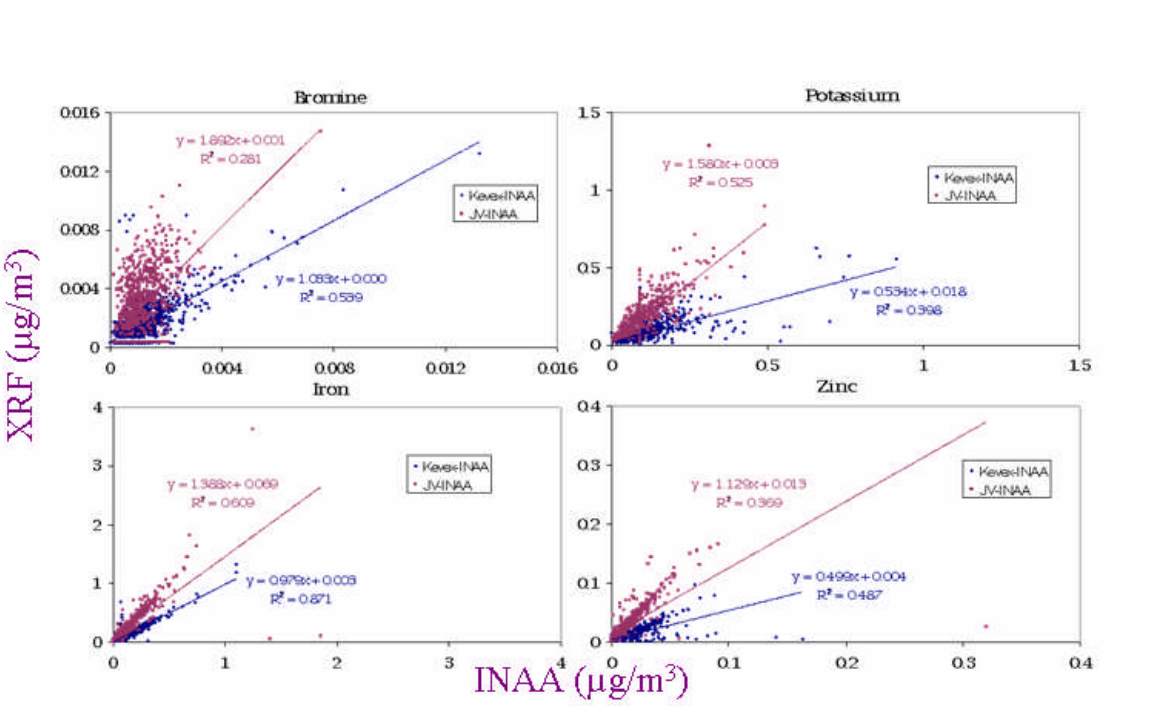


Figure 2. Comparison between INAA and XRF elements, showing how well they correspond. These plots were used as indicators whether common elements could be used interchangeably in the PMF or if models were best suited with specific instrument inputs.

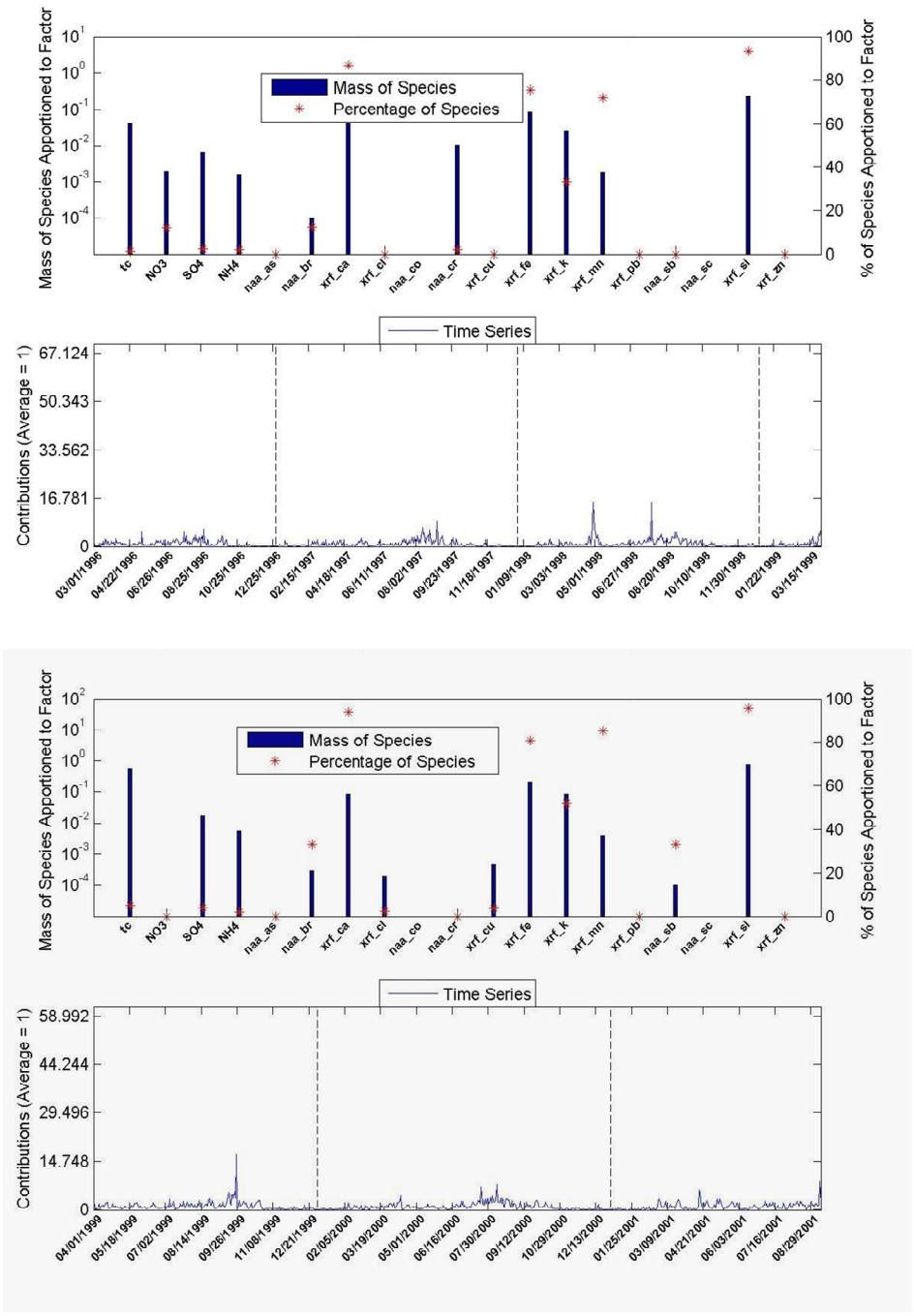


Figure 3. Airborne soil for INAA-Kevex (top) and INAA-JV data (bottom).

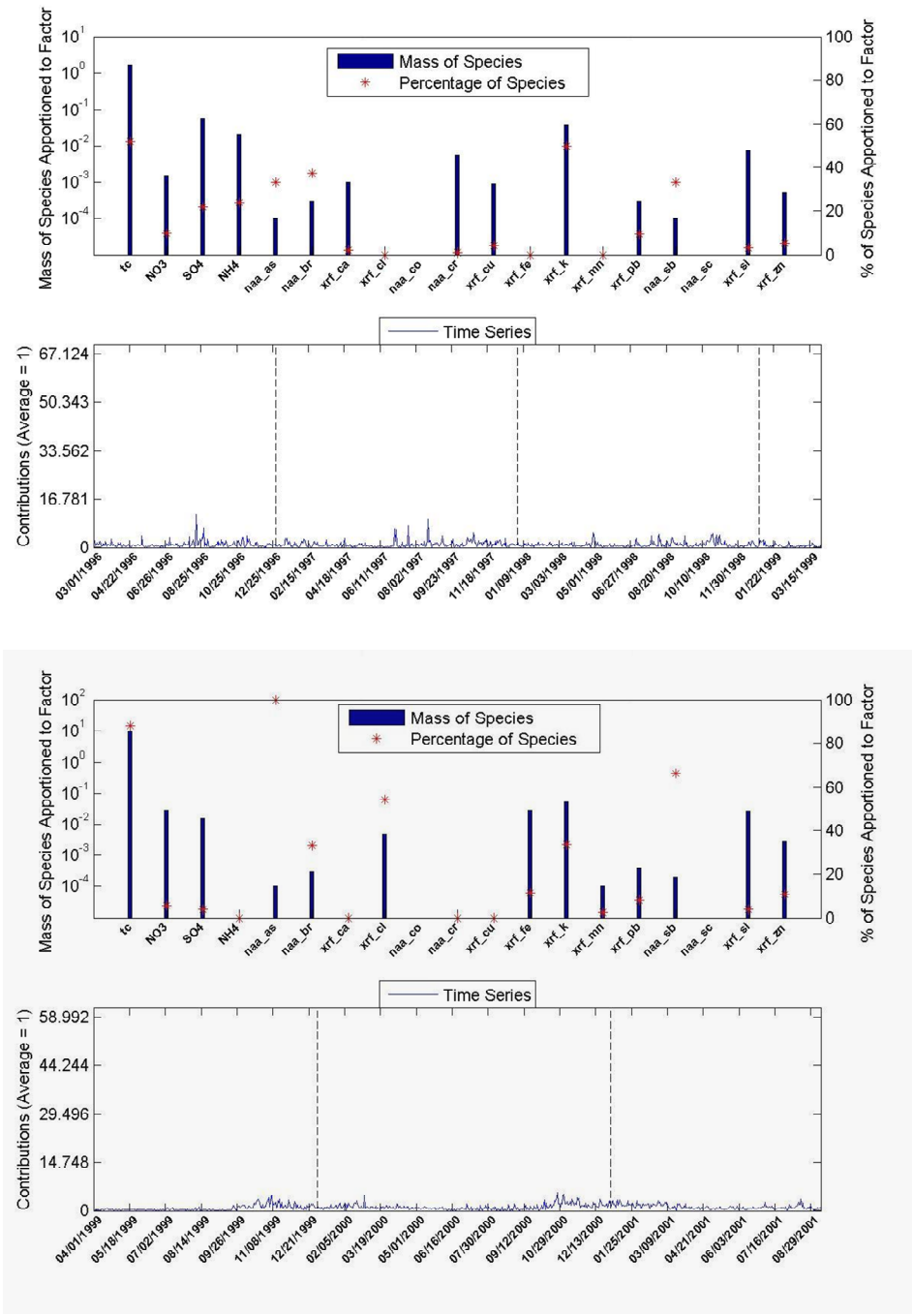


Figure 4. Biomass burning for INAA-Kevex (top) and INAA-JV data (bottom).

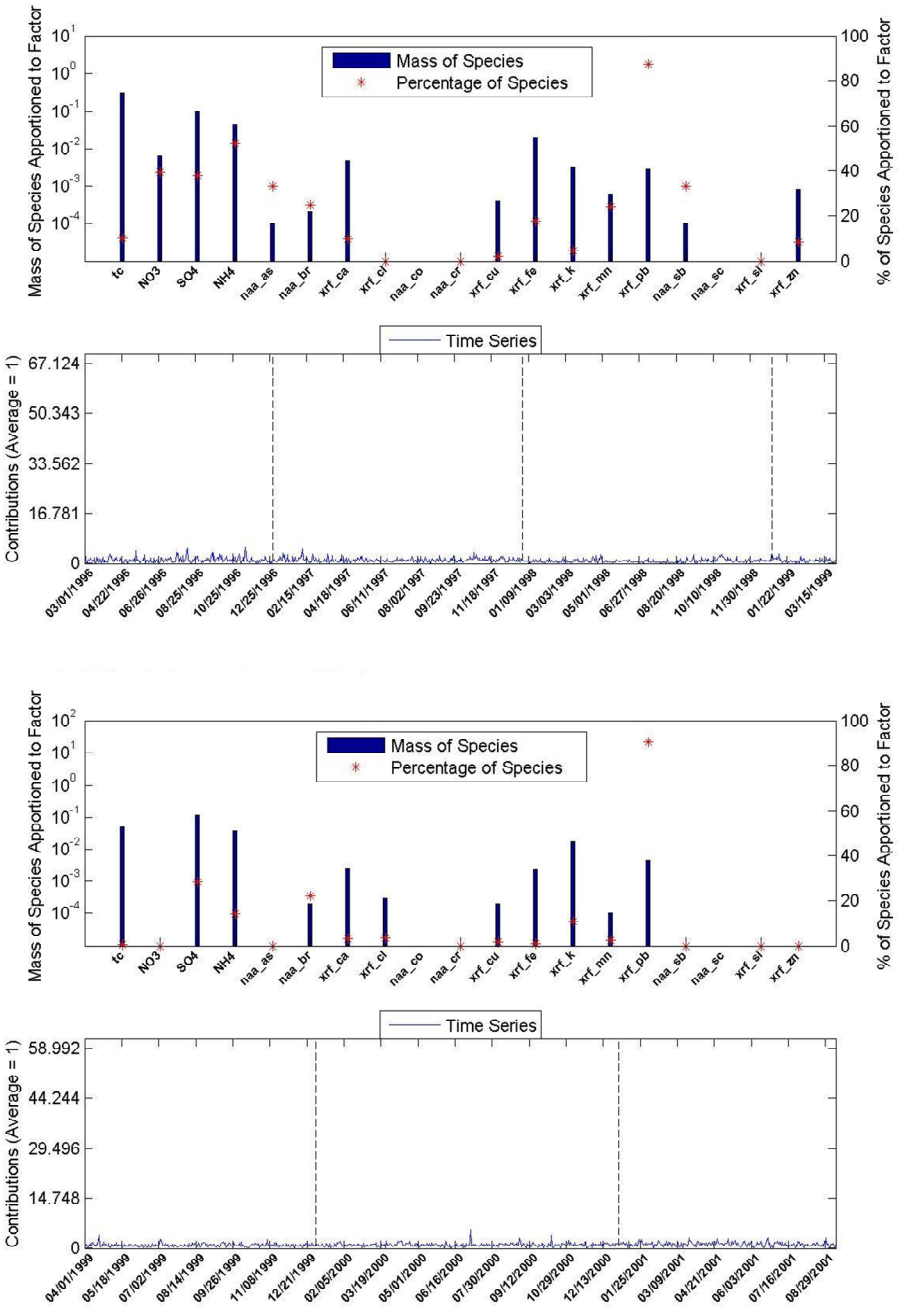


Figure 5. Vehicle exhaust for INAA-Kevex (top) and INAA-JV data (bottom).

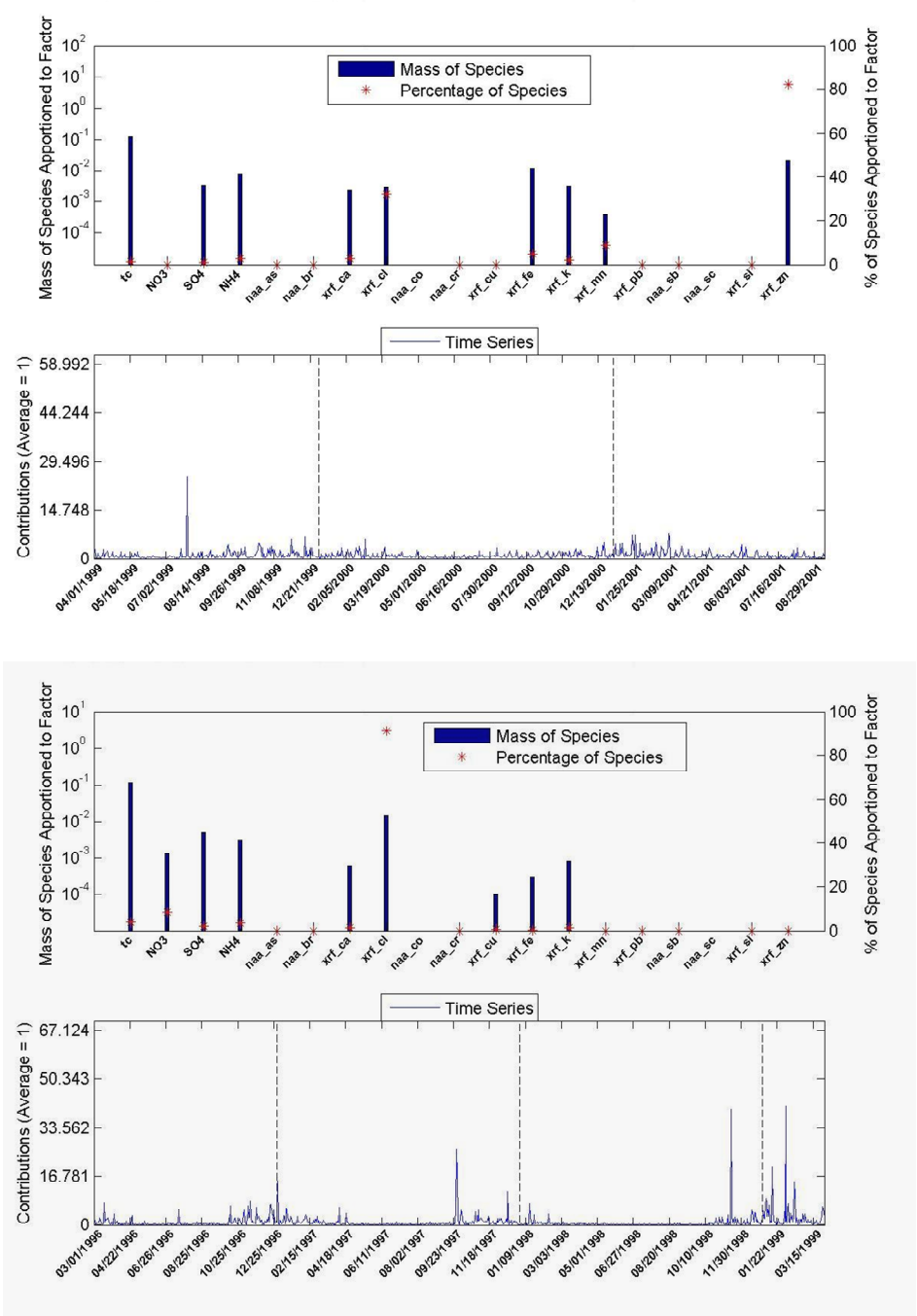


Figure 6. Cl-rich source for INAA-Kevex (top) and INAA-JV data (bottom).

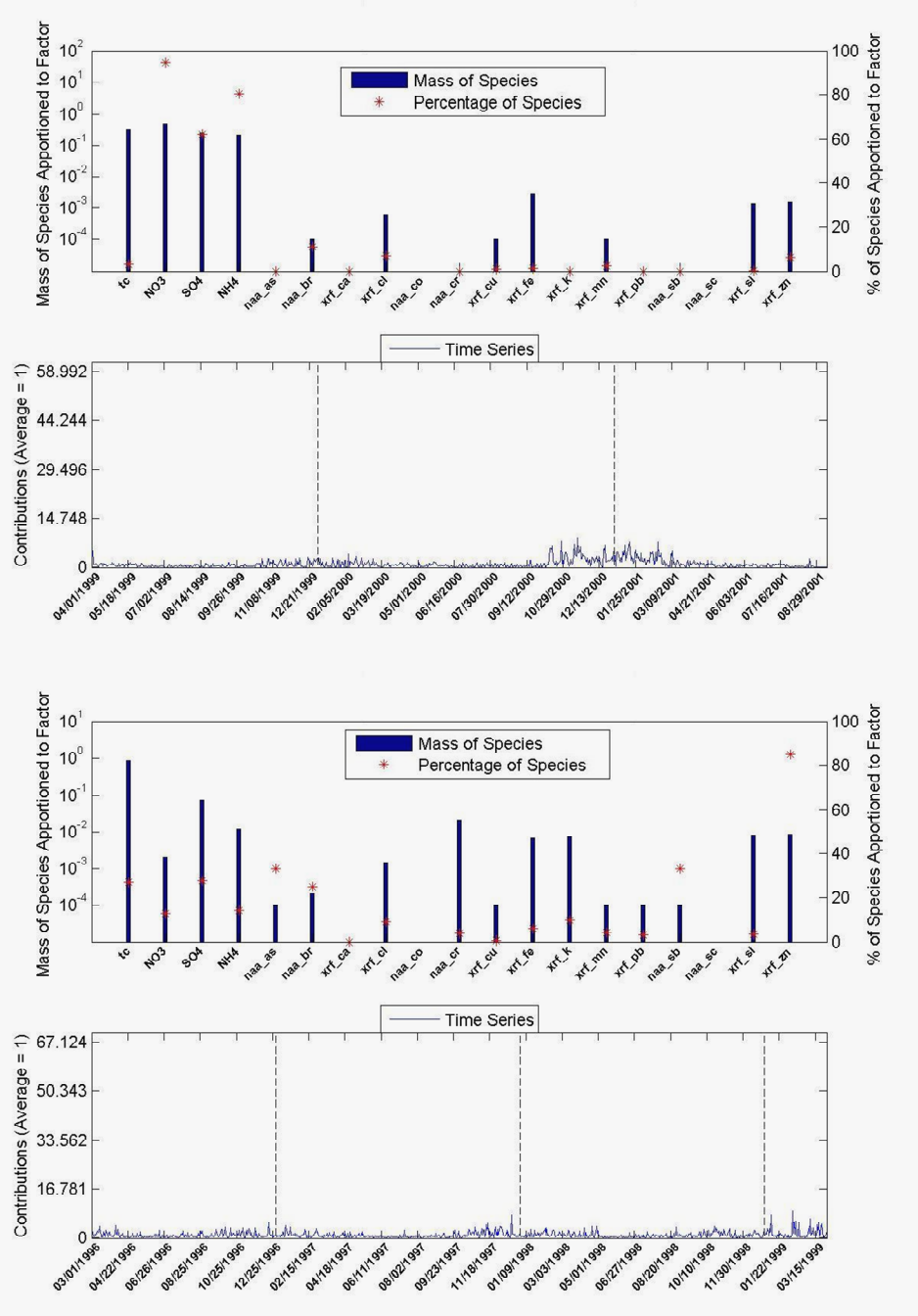


Figure 7. Nitrate for INAA-Kevex (top) and INAA-JV data (bottom).

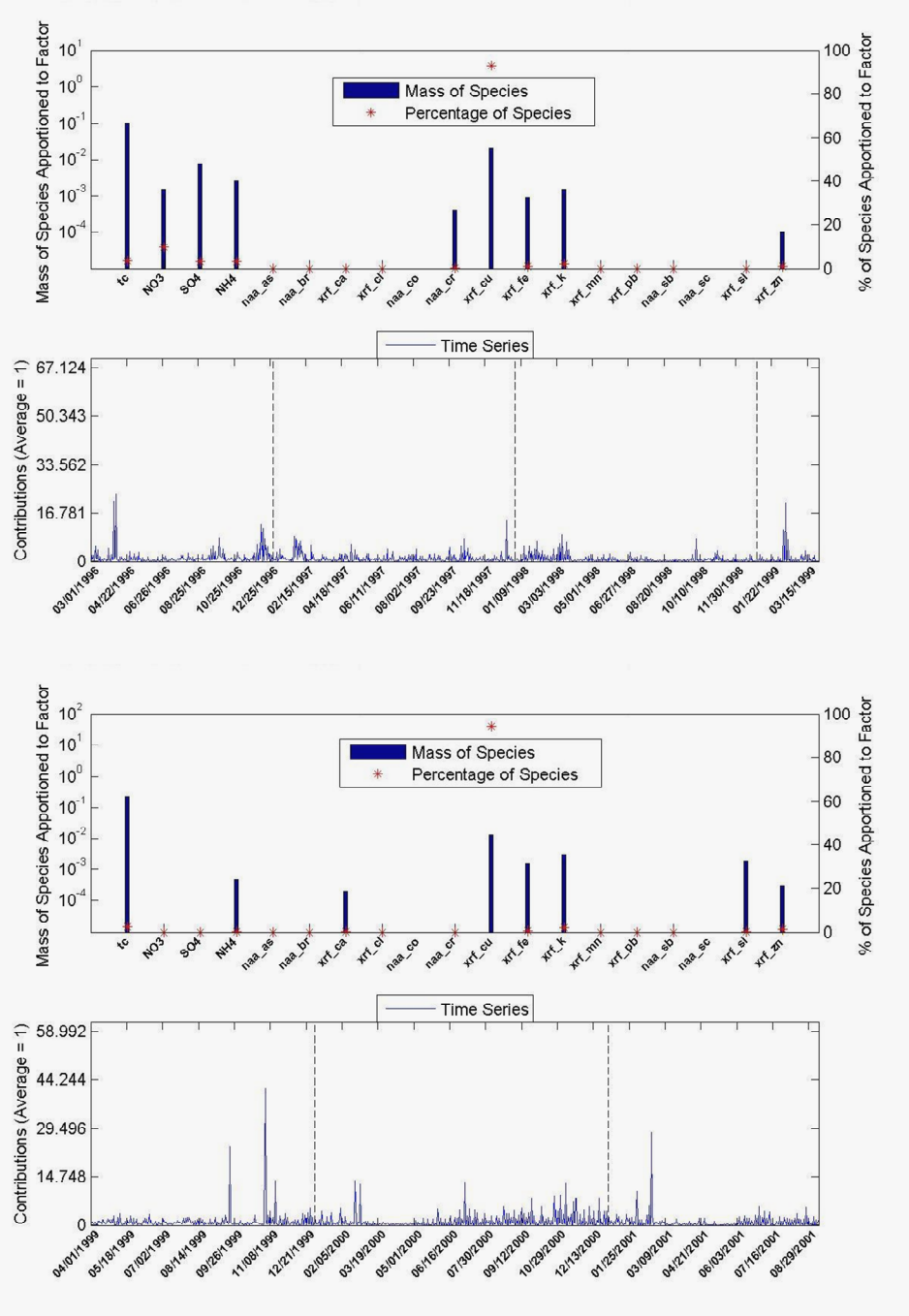


Figure 8. Metal processing for INAA-Kevex (top) and INAA-JV data (bottom).

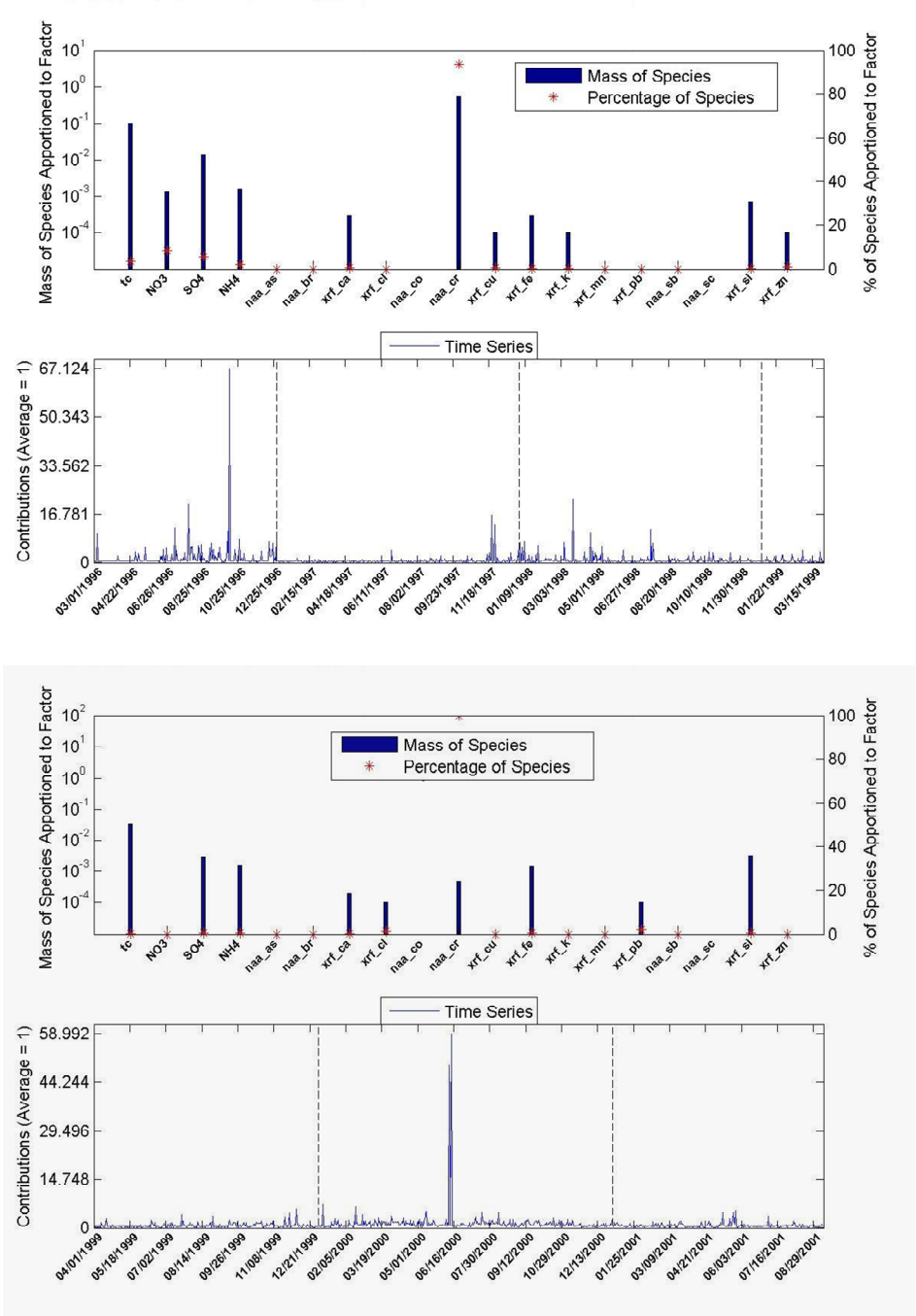


Figure 9. Cr-rich source for INAA-Kevex (top) and INAA-JV data (bottom); attached to each source profile are the time series for the respective datasets. Note specific episodes throughout the study period (1996, 1998, 2000).

CHAPTER 4

Sources of $PM_{2.5}$ modeled by EPA PMF version 1.1 using Trace Elements and Temperature-resolved Carbon

This chapter is presented as a paper that has been prepared for publication in a peer-reviewed journal.

Sources of $PM_{2.5}$ modeled by EPA PMF version 1.1 using Trace Elements and Temperature-resolved Carbon

Jennifer Shaltanis, Candis Claiborn
Laboratory for Atmospheric Research, Washington State University, Pullman, WA

Timothy Larson
Department of Civil Engineering, University of Washington, Seattle, WA

Correspondence:
Jennifer Shaltanis
Laboratory for Atmospheric Research
WSU/NSF IGERT Center for Multiphase Environmental Research
Department of Civil and Environmental Engineering
Washington State University
PO Box 642714
Pullman WA 99164-2714
Phone: (509) 335-7205
Fax: (509) 335-7632
Email: jshaltanis@wsu.edu

Abstract

Detailed carbon PM data from the Spokane Health Effects Study were modeled with EPA PMF version 1.1, along with, ionic data, and extensive trace element data provided by Instrumental Neutron Activation Analysis (INAA) and X-ray fluorescence (XRF), to create an exhaustive and long-term particulate profile for Eastern Washington. Pollution samples from March 1st, 1996 – September 10th, 2001 were analyzed for carbon, ionic, and trace element concentrations. Species included in this six year study were: elemental carbon (EC), organic carbon (OC) — including five subgroups by mass, NO_3^-, SO_4^{2-}, NH_4^+, As, Br, Ca, Cl, Co, Cr, Cu, Fe, K, Mn, Pb, Sb, Sc, Si, and Zn, for a total of 25 species.

Models were performed in a series of two studies, one using only total elemental and organic concentrations, and one smaller study, focusing on elemental and five organic carbon groups. PMF models evaluated data using a range of 5-9 sources, the number of sources hypothesized to be 6-7, based on previous PMF models using this dataset. EC-OC models revealed 7 identifiable sources of air pollution: airborne soil, metal processing with OC; biomass burning with both EC and OC; vehicle exhaust with both; nitrate with OC; and Cl-rich and Cr-rich sources with EC and OC, respectively. EC-OC fraction group models identified 6 plausible sources, eliminating metal processing as a source, but otherwise agreeing with the EC-OC and previous trace element sources. Clear OC markers include: OC_2 and OC_3 for biomass burning; OC_3 and OC_4 for vehicles; OC_5 for nitrate; OC_1 for the Cl-rich source. These carbon details were a final chapter in the fine fraction PM PMF analyses for the Spokane Health Effects Study, providing a strong understanding of the Spokane, WA air pollution profile.

Keywords: PMF, source apportionment, carbon fractions, elemental carbon, organic carbon, temperature-resolved carbon, XRF, trace elements, $PM_{2.5}$

4.1 Introduction

Carbon is a prominent feature of nearly all particulate pollution. Because of its diverse and abundant existence, it can be found in many forms, forms which can help identify specific pollution sources in concentration data. Specifically, carbon in PM can exist in the form of elemental carbon (EC), or as organic carbon (OC). Both are important because many forms pose potential threats to health and welfare of living organisms and the environment.

Carbon encompasses myriad forms and can be used to pinpoint several origins of particulate matter, specifically combustion such as domestic and agricultural burning and fossil fuel combustion (Schauer et al., 1996; Schauer and Cass, 2000; Park et al., *In Press*). Elemental carbon can be due to combustion of diesel fuels and coal and can decrease visibility because of its ability to absorb incident light. It can also indirectly weaken respiratory health in humans; diesel particles can synergistically interact with bioaerosols, resulting in elevated allergen sensitivity (Parnia et al., 2002; Adhikari et al., 2006). The complex and multiple natures of organic carbon can complicate the understanding of particulate behavior and cloud assessment of the local air quality. Organic carbon can stem from primary or secondary sources. The occurrence of OC may influence the overall ability of an aerosol to form into a cloud droplet, directly affecting aerosol concentration; scattering potential can also be affected, affecting visibility (Park et al., *In Press*).

Studies using only total carbon or even incorporating the OC/EC split may not be adequate to distinguish between certain combustion sources, especially gasoline and diesel exhaust, two common and large particulate sources (Polissar et al., 2001; Song et

al., 2001; Zhao and Hopke, 2006). Recent work has found associations between certain OC and EC subgroups and gasoline and diesel sources, notably work by Kim and Hopke (2004; 2004; 2005) and Kim et al. (2004), which suggest EC, OC_3 and OC_4, are tied to gasoline, and EC, OC_1 and OC_2 to diesel. Common patterns were found in the carbon profiles at differing sites, supporting the conclusions that diesel and gasoline can be uniquely traced by these resolved fractions.

Spokane is a medium size city in the eastern hills of Washington, with a distinctive environment, influenced by agricultural and natural settings, urban and industrial inlets. Hotter months bring forth dry, dusty days, while cool transition months have short periods of rain and temperatures that change quickly. Winter weather can last for several months, with short days and winter precipitation. Diversity in climate and land use can contribute to a variety of air pollution sources, historically creating a PM level exceeding the Environmental Protection Agency (EPA) National Ambient Air Quality Standards (NAAQS) limit (Villasenor et al., 2001).

Spokane's variety of particulate air pollution has been shown to cause distress in human health. Epidemiological studies have linked increased hospital visits for respiratory distress with increased PM (Dockery et al., 1993; Norris et al., 2000). Although these studies have not demonstrated ties between particular species and health problems, others conclude higher toxicity from certain chemicals found in pollution, suggesting some sources of particulate pollution are more dangerous than others (Green et al., 2002; Adhikari et al., 2006).

The concern over air pollution in Eastern Washington and its possible negative ramifications led to a massive campaign to collect weather, air pollution, and health data

collection from 1995-2002, known as the Spokane Health Effects study. This study was collaboration among the Mickey Leland Nation Urban Air Toxics Research Center, Washington State University, University of Washington, Spokane County Air Pollution Control Authority (SCAPCA), the American Lung Association of Washington, and the Harvard School of Public Health. One tenant of this long-term research included creating an exhaustive PM dataset, enabling researchers to comprehensively study fine fraction PM and determine a very accurate and complete particulate profile for the region.

Previous particulate models using the Spokane Health Effects Study have emphasized trace elements, SO_4^{2-}, NO_3^-, and NH_4^+, and total carbon (Kim et al., 2003), which is a somewhat incomplete approach. Similar to the trace element analyses in the Spokane dataset, which used more than one method, the Spokane Health Effects Study used carbon information measured by both the Thermal Manganese Oxidation (TMO) and the Thermal Optical Transmittance (TOT) instruments. The TMO method can estimate elemental and organic carbon, while the TOT instrument is capable of measuring the EC-OC split, further dividing the organic carbon into five mass grouped organic fractions.

The purpose of this research was to determine the fine fraction particulate source profile for Spokane, WA, using trace element, ionic, and carbon species, utilizing the detailed carbon information provided by TMO and TOT, by correcting for the limitations in instrumentation. It was hypothesized that the detailed carbon would not augment the profile with new sources; rather the details would further clarify and define the trace element and ion-derived sources in relation to OC and EC. This research would be a final

step in a comprehensive study of the fine fraction Spokane Health Effects particulate data.

4.2 Methods

From 1995-2002, ambient air and meteorological data were collected at a receptor north of Spokane's city center and downwind from a major road, to capture the essence of particulate pollution in Eastern Washington. The approach for this research first considered a full scope of the trace element and EC-OC study period, focusing on EC and OC, followed by a narrower review of EC and carbon fraction data available for the latter portion of the study.

Collection and most of the measurement methods were carried out by WSU researchers, according to EPA reference or equivalent methods. Versatile Air Pollutant Samplers (VAPS) collected 24 hour mass samples on quartz, polytetrafluoroethylene (Teflon), and nucelopore filters. Quartz filters were purified in an 800 ºF furnace for several hours, and stored in desiccators with the other filters prior to and after analysis to prevent contamination and humidity effects. Mass was measured with Cahn 32 and Cahn 33 microbalances. Carbon analyses were first performed on quartz filters with TMO, for samples from January 1, 1995 to September 20th, 1999, and were switched over to TOT starting on September 21st, 1999. Carrier gas flows were carefully monitored and sugar standards and lab blanks were included daily to minimize instrument error or drift. Trace elements were analyzed on Teflon and nucleopore filters, by the EPA Kevex and Jordan Valley Energy Dispersive X-Ray Fluorescence (ED-XRF) instruments and Instrumental Neutron Activation Analysis. Ionic extraction chromatography (SO_4^{2-} and NO_3^-) and

colorimetry (NH_4^+) were also performed on quartz filters. As with carbon analyses, all methods incorporated standards and blanks to minimize any instrument error or drift.

Carbon information were taken from raw oxidation analyses, in terms of μg/cm^2, and converted to mass concentration. For all carbon data, raw mass was multiplied by a factor of 11.34 cm^2, the area of exposure for the VAPS quartz filter, and normalized by quartz flow volume (m^3). Uncertainty for OC-EC split data was calculated in the same manner. For the OC fractions, uncertainty was not individually reported (only total OC and EC were given), so the ratio of each fraction to total OC was used to factor out uncertainty for each. That is, for each sample,

$$\sigma_i = \frac{OC_i}{OC} \times \sigma \qquad (1)$$

where σ_i represents uncertainty for OC_1-OC_5 and σ is the total OC uncertainty.

Corrections had to be considered when using the carbon analysis methods because of known instrument limitations, the greatest of these being the step change in instrumentation and error in analyzing the EC-OC split. Instrument change can be addressed fairly easily when one of the instruments used is assumed accurate and/or a reference method. Filter artifacts can be a more difficult problem to resolve. Tests conducted by Jordan et. al. (2006) suggested that the presence of EC in samples could be misrepresented, depending on the method and constraints surrounding that method. For OC, Park et al. (*In Press*) and Chow et al. (2006), give thorough summaries of several artifact studies (Eatough et al., 1990, 1993; McDow and Huntzicker, 1990; Tang et al., 1994; Chow et al., 1994, 1998; Malm et al., 1994; Turpin et al., 1994, 2000; Gundel et al., 1995; Andrews et al., 2000; Ding et al,., 2002; Mader et al., 2003; and Subramanian et al. 2004), in which they conclude that artifacts can contribute an error of ~ 3-10 % in

OC mass concentration. Several of the studies derived different schemes for correcting biases. Raw data time series plots indicated that the problems with EC and OC would pose a far smaller error than the instrument shift, and for this research, only the instrument shift is addressed.

An instrument correction was imposed on TMO data, for the EC-OC study, using TOT as the baseline. Time series data showed a step change within the TMO data, roughly mid-February, 1998 (the shift was most likely due to detector wear) and one starting September 21st, 1999, reflecting the TMO TOT transition. Total OC measured by the TMO instrument is assumed correct in spite of its difficulty correctly assessing the EC-OC split, which is validated by its seamless transition into TOT data. In addition, roughly 100 TMO period samples were re-analyzed with TOT, which showed general TC agreement to the TMO measurements. Several corrections were attempted to reconcile TMO to TOT. The most realistic transition was made with the following equations:

$$\Delta EC = average\left(\frac{EC}{TC}\right)_{TOT} - average\left(\frac{EC}{TC}\right)_{TMO} \quad (2a)$$

$$\left(EC_{TMO}\right)_{corr} = EC_{TMO} + TC_{TMO} * \Delta EC \quad (2b)$$

where ΔEC represents the decrease in EC/TC proportion from TMO to TOT needed to create a seamless EC concentration for the data period, $(EC_{TMO})_{corr}$ represents the new EC concentrations. These values were calculated for the two groups of TMO data to create a smooth transition among all TMO and TOT data. The corrected TMO EC data were calculated with sample specific EC and TC values (not group averages) to maximize variability and individual influence on PMF analysis.

Carbon information was analyzed by PMF, first for emphasis on elemental and organic carbon with supplemental information from ions and trace elements, followed by

separation of organic carbon into mass grouped fractions. For the EC-OC analysis, data covered twenty species: EC, OC, SO_4^{2-}, NO_3^-, NH_4^+; XRF elements Ca, Cl, Cu, Fe, K, Mn, Pb, Si, Zn; and INAA elements As, Br, Co, Cr, Sb, and Sc, for a total of 1684 samples, from the period of March 1st, 1996-September 10th, 2001. Data are separated into consecutive data groups, one by INAA-Kevex and one with INAA-JV. Separation of data by the two XRF instruments is related to an issue with agreement between the two XRF instruments (for discussion, please refer to the previous section, chapter 3). OC fractions studies only covered the TOT portion of analysis, or 586 samples, from September 21st, 1999-September 10th, 2001, during which time only the Jordan Valley and WSU INAA methods were used for trace elements.

The choice of apportionment model was the EPA Positive Matrix Factorization model, version 1.1. This model and its predecessors have been successfully used on the Spokane Health dataset in prior work (Kim et al, 2003). Its algorithms are based loosely on factor analysis principles found in classical statistics, with special constraints and bootstrap features added to treat the dynamic and unique nature of environmental data. The results found are reconstructed sources of the pollution, which are easily identified as a specific pollution-emitting activity or group of activities. Detailed discussion of the principles involved are found in: Paatero (1997); Huang et al. (1999); Lee et al. (1999); Xie et al. (1999); Ramadan et al. (2000); Willis, (2000); Polissar et al. (1998); Song et al. (2001); Henry (2002); Paatero et al. (2002); Qin et al. (2002); and Kim et al. (2003). Details regarding the EPA version's special features are available in the product manual by Eberly (2005). Special corrections, instrument adjustment, and data filtering were

done primarily before PMF analysis; PMF defaults were used with any exceptions noted in the results.

4.3 Results and Discussion

4.3.1 EC-OC Study

The final PMF model used all defaults, with exception of a weak classification for EC for both the INAA-Kevex and INAA-JV datasets. The weak classification in PMF v. 1.1 increases uncertainty by a factor of one-third and is used to treat species that may have be very sporadic or linger near detection levels. The extra uncertainty was included because the EC data are relatively small in concentration and have had a correction applied to part of them. Looking at the EC-OC split raw data, one can see organic carbon accounts for most of the total carbon sampled on the filters. In general, organic carbon is responsible for 60-90% of the total amount (uncorrected), with elemental carbon making up the remainder.

The PMF model derived 7 sources for the two datasets, in accordance with previous work done on this dataset. As hypothesized, no additional sources were distinctively found by the carbon split; rather they gave greater clarity to the known sources, which are vehicle exhaust, nitrate, biomass burning, airborne soil, Cl-rich and Cr-rich sources, and metal processing. **Tables 1** and **2** list the mass contributions for each source, for all data as well as samples in the highest twenty per cent mass group. **Table 3** lists the Pearson correlations for the sources and species (only values above 0.30 are given). In addition to the source profiles and mass contributions, these correlations help identify which species and sources are closely related.

Vehicle exhaust, nitrate, and biomass burning are results of different tyopes of carbon-based fuel combustion, and therefore contribute the largest amount of carbon particles. Vehicle exhaust, seen in **Figure 1**, shows little EC and some OC accompanying the historical markers of Pb, Zn, and Br. **Figure 2** shows nitrate, which is characterized by NO_3^-, SO_4^{2-}, NH_4^+, geological material, and OC. The INAA-JV data also show presence of EC, which the INAA-Kevex placed in soil. This shift in source profiles is most likely due to XRF sensitivity; this type of discord was not easily solved in previous work and is why the datasets remain separate. The more abundant, geological species are associated with EC but the XRF instruments measure these dust fragments differently which is what is motivating the PMF change. Likewise, biomass burning, seen in **Figure 3**, is marked by K, Sb, As, geological particles, OC, and, for INAA-JV, EC. Again here, where there are dust fragments present in low amounts, their association with EC may interfere with the soil profile, as analyzed by INAA-Kevex and PMF. K, As, Sb, and OC confirm these sources as vegetative in origin. K and OC are connected to one another as indicators of organic material burns, and As and Sb are also indicative of such combustion, notably in the colder weather, as concluded by Beceiro-González et. al (1997) and Tsai et. al (2003), which is when field burns take place.

Mechanical and industrial sources contain less, but specific carbon associations. Airborne soil contains high quantities of crustal elements, small traces of ions, and some EC, as seen in **Figure 4**. This source peaks in late summer or early fall, during which time, bare soil have collected plant matter, road dust, and pollution deposits. The EC-OC model using Kevex and INAA data indicate on average no mass is attributed to the average $PM_{2.5}$ mass; this result is likely due to the lack of carbon in the source. The

source provides a moderate amount of mass to the profile, with noticeable peaks but shadows in comparison to what carbon contributes. The Cl rich source contains Cl, some crustal elements, and EC, and is seen in **Figure 5**. The increase of lesser species for the INAA-JV profile suggests either Cl is not well defined by JV or it is a diminishing source; the presence of EC and Cl together however, show Cl is its own distinctive source. **Figure 6** shows a Cr-rich source, characterized uniquely by Cr and virtually nothing else. OC is minimal in both cases, suggesting the OC is a product of a carrier fuel used in processing the Cr metal. EC also appears in the INAA-Kevex case; time series contributions indicate more activity during these first years of study, whereas the INAA-JC case shows only one distinct event (in this case, there may not be enough data to accurately depict the carbon details for the Cr source). Metal processing, seen in **Figure** 7, reflects a similar situation as the Cr-rich source. This source is marked primarily by one metal, with trace amounts of other species, including OC for INAA-JV and both OC and EC for INAA-Kevex. In this case however, this source seems to increase activity during the INAA-JV period, indicating the PMF model for INAA-Kevex could be incorrect in its association of Cu and EC.

4.3.2 EC-OC fractions study

The PMF model used in this case applied weak classifications to EC, OC_4 and OC_5. The weak EC factor is carried over from the large EC-OC study, and OC_4 and OC_5 are “weak” because they are the lightest of the fractions; while the OC_4/OC ratio is not one of the lowest, there are several points near 0. One limitation with PMF is that because it is looking for patterns, it is not strictly tied to concentration; however, when the data in

question contains values covering several orders of magnitude (i.e., very high highs and very low lows), that freedom can become a limiting factor. The additional uncertainty helps relax that constraint.

When looking at carbon fractions over time (displayed in **Figure 8**), one can see both general trends and unique features in the data. Their relative contributions to OC, in terms of average per cent, are: 26% for OC_1; 25% for OC_2; 17% for OC_3; 21% for OC_4; and 11% for OC_5. Standard deviations for OC_i/OC ratios are consistently about 8-9%. Looking at the fraction masses relative to overall organic mass, one can see that the contributions are generally consistent. OC_2 and OC_3 tend to follow the same high and low contribution pattern, with a small increase for mid- to late summer of 2000, and sporadically elevated levels for 2001. OC_1 and OC_4 both decrease in their relative contributions for summer of 2000, in contrast to OC_2 and OC_3. OC_4 shows more variability than OC_1-OC_3 in its contribution to overall OC, which may indicate a specific source of OC_4. OC_5 can be due to a unique source but it is clouded by volatility; the lighter masses associated with OC_5 are more likely to be lost during sampling or storage, affecting the true source profile.

The PMF model successfully identified 6 of the originally established sources, with significant detail for elemental and carbon fraction influences. **Table 4** lists the Pearson correlations for the sources and species (only values above 0.30 are given). Vehicle exhaust, nitrate, biomass burning, airborne soil, Cl-rich, and Cr-rich sources are evident with unique carbon (or lack thereof). The first three are more are labeled traditional combustion sources, where either a petroleum product or vegetative material is

used as the fuel source, and the final three are industrial processes or mechanically induced.

Combustion sources show distinct EC-OC fraction signatures, in which each source accounts for the majority of at least one carbon group. **Figure 9** shows vehicle exhaust, with classic markers of Pb, Zn, Br, trace geological materials, and EC and OC. This exhaust source specifically identifies EC and OC_3 and OC_4 as the primary carbon inputs, which agrees with prior work by Kim and Hopke (2004). Unlike its predecessor, this PMF model does not separate diesel and gasoline markers. This may be due to the site location, which is downwind from a major highway that sustains both POV and commercial traffic; perhaps if this road were either a residential road (primarily POV) or interstate (higher concentration of commercial trucks), data could be separated further. Like the work by Kim and Hopke, the primary markers are the carbon data; however, these models also use Pb and Zn as identifiers. Pb is always available for vehicle identification, especially in more recently collected data. Zn is found in trace amounts in vehicle profiles which correspond to the findings here. Nitrate, seen in **Figure 10**, is characterized by NO_3^-, NH_4^+, OC_5 and trace amounts of geological materials. This source is a wintertime feature, corresponding to domestic heating increases via wood burning. Biomass burning is identified in **Figure 11**, containing Fe, K, SO_4^{2-}, small amounts of EC, and OC_2 and OC_3. Time series plot shows this source is more influential in cooler months, corresponding to post-harvest burns.

Industrial and mechanical sources include airborne soil, Cl-rich, and Cr-rich sources, also containing specific carbon signatures, either fairly high or noticeably lacking. Airborne soil (seen in **Figure 12**) contains the majority of geological elements,

and very carbon, just a trace amount of OC_2. This carbon is the second heaviest group of organics, which could correspond to organic material caught up in small soil particles. Airborne soil is at its peak during late summer or early fall, when post-harvest fields are laden with bare soil and crop debris and exposed to the hot, dry weather. The Cl-rich and Cr-rich profiles are not as clear as previous analyses, due to the shorter period of study and their episodic behavior. The Cl-rich source, seen in **Figure 13**, does show co-existence of Cl with OC_1 and some EC and OC_2. JV data shows a considerable association of Cl to Zn, which is mitigated but not resolved by the carbon details. This correlation is assumed incorrect, based on the Kevex profile, which links Zn to vehicle exhaust and Pb, not Cl. The mass from the Cl-source using JV XRF data likely contains some mass actually contributed by vehicle exhaust. Unfortunately the carbon information is not significant enough to correct the instrumental inaccuracy. The Cr-rich source shown in **Figure 14**, like soil, shows very little of any carbon, just small amounts of OC_2, OC_3, and OC_4. The organic carbon could be due to the fuel used to heat and process the Cr, and are essentially secondary markers in this source; that is, their presence is incidental and not indicative of the true nature of the source.

Metal processing is absent from this PMF analyses, likely due to its low overall contribution to $PM_{2.5}$ mass and small association with carbon. In the previous EC-OC study, only a trace amount of organic carbon was present with the copper marker. During peak emissions, still only a small percentage of mass is attributed to metal processing. Carbon contributes more mass than any other particle type and it is most likely overshadowing this smaller feature.

4.4 Conclusions

The addition of EC and OC details confirm prior results using only total carbon, ions, and trace elements, by maintaining their original structures, while adding valuable EC-OC markers. In general, combustion sources gained mass and presence strength in the profile, while features not highly correlated to carbon decreased in mass and profile influence. As with many trace elements, particular EC and OC groups are surfacing in research as markers for specific sources of pollution. Carbon can be difficult to quantify and the concentration data are intrinsically tied to the method used for analysis. The volatile nature of several organic compounds can lead to loss of sample and therefore a skewed model, underestimating the true abundance of these species..

This research brings to the PMF work on the fine fraction Spokane data full circle. What was started years ago with only partial datasets and some of the of instrumentation is now joined by all the trace element analyses, ions, and carbon details collected in the historical Spokane Health Effects Study. Future work includes adding on weather and gaseous pollutant data, which would address several of the criteria pollutants defined by EPA. Future work also includes moving on to coarse fraction and comparison of those results to the fine fraction profile. Current and planned health studies will investigate and conclude any epidemiological influences by the species found to be significant source contributors in this and the trace element studies. Past, current, and future work are meant to further knowledge about Eastern Washington air pollution and its ramifications on human health and welfare.

Acknowledgements

This study was funded in part by the Integrative Graduate Education and Research Training (IGERT) grant from the National Science Foundation to Washington State University under grant DGE-0072817, the Washington State Department of Ecology, Mickey Leland Nation Urban Air Toxics Research Center, University of Washington/EPA Northwest Research Center for Particulate Matter and Health Center (R827355) and University of Rochester/EPA Particulate Matter and Health Center (R827354). This research has not been subjected to EPA's required peer and policy review and does not necessarily reflect the views of EPA.

References

Adhikari, A., Reponen, T., Grinshpun, S.A., Martuzevicius, D., LeMasters, G., 2006. Correlation of ambient inhalable bioaerosols with particulate matter and ozone: A two-year study. Environmental Pollution 140, 16-28.

Andrews, E., Saxena, P., Musarra, S., Hildemann, L.M., Koutrakis, P., McMurry, P.H., Olmez, I.,White, W.H., 2000. Concentration and composition of atmospheric aerosols from the 1995 SEAVS experiment and a review of the closure between chemical and gravimetric measurements. Journal of Air and Waste Management Association 50, 648-664.

Beceiro-González, E., González-Soto, E., López-Mahía, Prada-Rodríguez, D., 1997. Total arsenic and selenium levels in atmospheric particulate matter of La Coruña (Spain). The Science of the Total Environment 208, 207, 211.Burtscher, H., 2005. Physical characterization of particulate emissions from diesel engines: a review. Aerosol Science, 896-932.

Chow, J.C., Watson, J.G., Fujita, E.M. Lu, Z., Lawson, D.R., Ashbaugh, L.L., 1994. Temporal and spatial variations of $PM_{2.5}$ and PM_{10} aerosol in the southern California air quality study. Atmospheric Environment 28, 2061-2080.

Chow, J.C., Watson, J.G., Lowenthal, D.H., Egami, R.T., Solomon, P.A., Thuillier, R.H., Magliano, K., 1998. Spatial and temporal variations of particulate precursor gases and photochemical reaction products during SJVAQS/AUSEPX ozone episodes. Atmospheric Environment 32, 2835-2844.

Chow, J.C., Watson, J.G., Lowenthal, D.H., Chen, L.-W.C., Magliano, K.L., 2006. Particulate carbon measurements in California's San Joaquin Valley. Chemosphere 62, 337-348.

Ding, Y. Pang, Y., Eatough, D.J., 2002. High-volume diffusion denuder sampler for the

routine monitoring of fine particulate matter: I. Design and optimization of the PC-BOSS. Aerosol Science and Technology 36, 369-382.

Dockery, D.W., Pope, C.A., Xu, X.P., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air-pollution and mortality in 6 United-States cities. New England Journal of Medicine 329, 1753-1759.

Eatough, D.J., Aghdaie, N., Cottam, M., Gammon, T., Hansen, L.D., Lewis, E.A., Farber, R.J., 1990. Loss of semi-volatile organic compounds from particles during sampling on filters. In: Mathai, C.V. (Ed.), Transactions, Visibility and Fine Particles. Air & Waste Management Association, Pittsburgh, PA, pp. 146–156.

Eatough, D.J., Wadsworth, A., Eatough, D.A., Crawford, J.W., Hansen, L.D., Lewis, E.A., 1993. A multiple-system, multi-channel defusion denuder sampler for the determination of fine-particulate organic material in the atmosphere. Atmospheric Environment 27, 1213–1219.

Eberly, S. 2005. EPA PMF 1.1 User's Guide. US Environmental Protection Agency.

Green, L.C., Crouch, E.A., Ames, M.R., Lash, T.L., 2002. What's Wrong with the National Ambient Air Quality Standard (INAAQS) for Fine Particulate Matter ($PM_{2.5}$)? Regulatory Toxicology and Pharmacology 35, 327-337.

Gundel, L.A., Stevens, R.K., Daisey, J.M., Lee, V.C., Mahanama, K.R.R., Cancel-Velez, H.G., 1995. Direct determination of the phase distributions of semi-volatile polycyclic aromatic hydrocarbons using annular denuders. Atmospheric Environment 29, 1719–1733.

Henry, R.C., 2002. Multivariate receptor models – current practice and future trends. Chemometrics and intelligent laboratory systems 60, 43-48.

Huang, S., Rahn, K.A., Arimoto, R., 1999. Testing and optimization two factor-analysis techniques on aerosol at Narragansett, Rhode Island. Atmospheric Environment 33, 2169– 2185.

Jordan, T.B., Seen, A.J., Jacobsen, G.E., Gras, J.L., 2006. Radiocarbon determination of woodsmoke contribution to air particulate matter in Launceston, Tasmania. Atmospheric Environment, 2575-2582.

Kim, E., Larson, T.V., Hopke, P. K., Slaughter, C., Sheppard, E., Claiborn, C., 2003. Source identification of PM2.5 in an arid northwest U.S. city by positive matrix factorization. Atmospheric Research 66, 291-305.

Kim, E., Hopke, P.K., Edgerton, E.S., 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. Atmospheric Environment 38, 3349-3362.

Kim, E., Hopke, P.K., 2004. Source apportionment of fine particles at Washington, DC utilizing temperature resolved carbon fractions in positive matrix factorization. Journal of the Air & Waste Management Association 54, 773-785.

Kim, E., Hopke, P.K., 2004. Improving source identification of fine particles in a rural northeastern U.S. area utilizing temperature resolved carbon fractions. Journal of Geophysical Research 109, D09204.

Kim, E., Hopke, P.K., 2005. Improving source apportionment of fine particles in the Eastern United States utilizing temperature-resolved carbon fractions. Journal of the Air & Waste Management Association 55,1456-1463.

Kim, E., Hopke, P.K., 2006. Characterization of fine particle sources in the Great Smoky Mountains area. Science of the Total Environment, *In Press*.

Lee, E., Chan, C.K., Paatero, P., 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmospheric Environment 33, 3201-3212.

Mader, B.T., Pankow, J.F., 2001. Gas/solid partitioning of semivolatile organic compounds (SOCs) to air filters. 3. An analysis of gas absorption artifacts in measurements of atmospheric SOCs and organic carbon (OC) when using Teflon membrane filters and quartz fiber filters. Environmental Science & Technology 35, 3422–3432.

Malm, W.C., Sisler, J.F., Huffman, D., Eldred, R.A., Cahill, T.A., 1994. Spatial and seasonal trends in particle concentrations and optical extinction in the United States. Journal of Geophysical Research 99, 1347–1370.

McDow, S.R., Huntzicker, J.J., 1990. Vapor adsorption artifact in the sampling of organic aerosol: face velocity effects: Face velocity effects. Atmospheric Environment 24A, 2563–2571.

Norris, G., Larson, T., Koenig, J., Claiborn, C., Sheppard, L., Finn, D., 2000. Asthma aggravation, combustion, and stagnant air. Thorax 55, 466-470.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and intelligent laboratory systems 37, 23-35.

Paatero, P. Hopke, P.K., Song, X.-H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. Chemometrics and intelligent laboratory systems 60, 253-264.

Park, S.S., Bae, M.-S., Schauer, J.J., Young, J.K., Cho, S.Y., Kim, S.J. In Press. Molecular composition of $PM_{2.5}$ organic aerosol measured at an urban site of Korea during the ACE-Asia campaign. Atmospheric Environment In Press.

Parnia, S., Brown, J.L., Frew, A.J., 2002. The role of pollutants in allergic sensitization and the development of asthma. Allergy 57, 1111-1117.

Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: 1. Elemental composition and sources. Journal of Geophysical Research 103 (D15), 19045– 19057.

Qin, Y., Oduyemi, K., Chan, L.Y., 2002. Comparative testing of PMF and CFA models. Chemometrics and Intelligent Laboratory Systems 61, 75-87.

Ramadan, Z., Song, X.H., Hopke, P.K., 2000. Identification of sources of Phoenix aerosol by positive matrix factorization. Journal of the Air & Waste Management Association 50, 1308–1320.

Schauer, J.J., Rogge, W.F., Hildermann, L.M., Mazurek, M.A., Cass, G.R., 1996. Source apportionment of airborne particulate matter using organic compounds as tracers. Atmospheric Environment 30, 3837-3855.

Schauer, J.J., Cass, G.R., 2000. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds. Environmental Science & Technology 34, 1821–1832.

Song, X.-H., Polissar, A.V., Hopke, P.K., 2001. Sources of fine particle composition in the northeastern US. Atmospheric Environment 35, 5277-5286.

Subramanian, R., Khlystov, A.Y., Cabada, J.C., Robinson, A.L., 2004. Positive and negative artifacts in particulate organic carbon measurements with denuded and undenuded sampler configurations. Aerosol Science & Technology 38, 27–48.

Tang, H., Lewis, E.A., Eatough, D.J., Burton, R.M., Farber, R.J., 1994. Determination of

the particle size distribution and chemical composition of semi-volatile organic compounds in atmospheric fine particles with a diffusion denuder system. Atmospheric Environment 28, 939–947.

Tsai, Y.I,, Kuo, S-C., Lin, Y-H., 2003. Temporal characteristics of inhalable mercury and arsenic aerosols in the urban atmosphere in southern Taiwan. Atmospheric Environment 37, 3401-3411.

Turpin, B.J., Huntzicker, J.J., Hering, S.V., 1994. Investigation of organic aerosol sampling artifacts in the Los Angeles Basin. Atmospheric Environment 28, 3061–3071.

Villasenor, R., Claiborn, C., Lamb, B., O'Neill, S., 2001. Mesoscale modeling of wintertime particulate matter episodes in eastern Washington, USA. Atmospheric Environment 35, 6479-6491.

Willis, R.D., 2000. Workshop on UNMIX and PMF as applied to PM2.5. EPA 600-A-00-048.

Xie, Y.-L., Hopke, P.K., Paatero, P., Barrie, L.A., Li, S.-M. 1999. Identification of source nature and seasonal variations of Arctic aerosol by the multi-linear engine. Atmospheric Environment 33, 2549-2562.

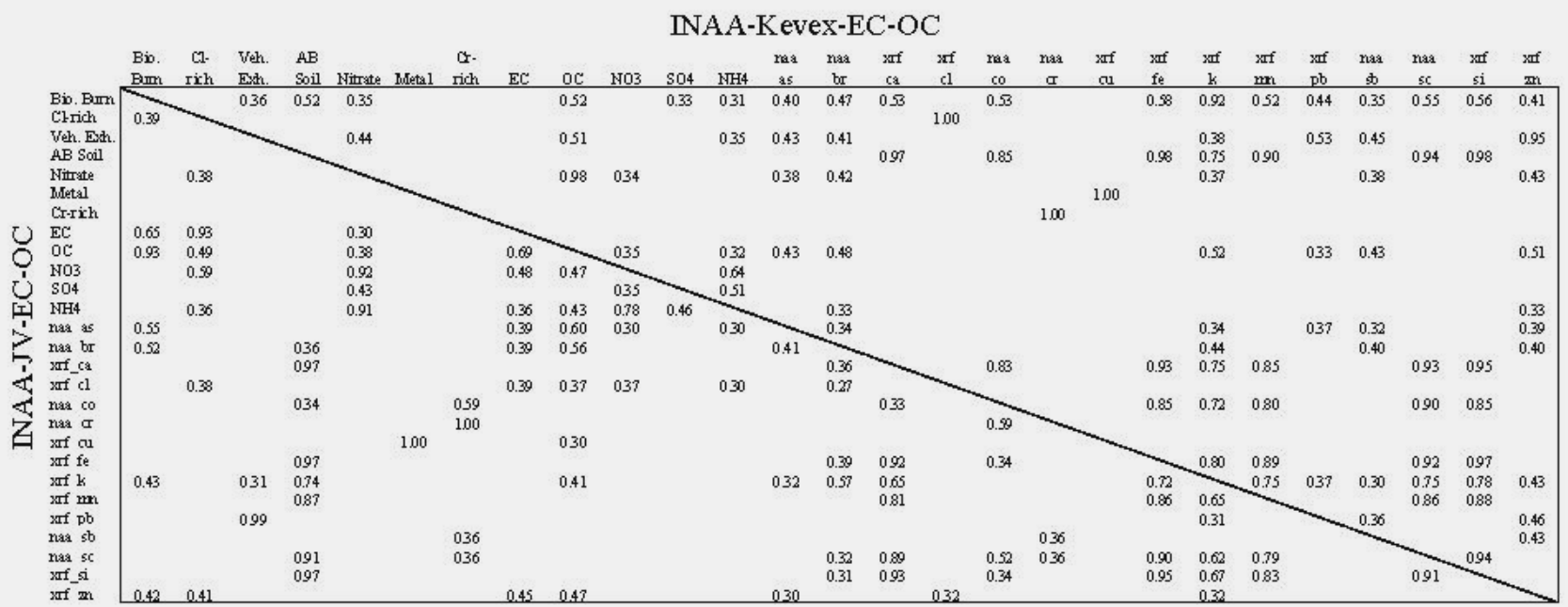
Zhao, W., Hopke, P.K. 2006. Source identification for fine aerosols in Mammoth Cave National Park. Atmospheric Research 80, 309-322.

$PM_{2.5}$ Source	INAA-XRF-EC-OC Data		INAA-XRF-EC-OC_{1-5} Data
	INAA-Kevex	**INAA-JV**	**INAA-JV**
Airborne soil	0.00 / 0.00 (0.00; 0.26)	1.51 / 15.5 (1.36; 1.60)	1.54 / 15.9 (1.26; 1.66)
Nitrate	4.71 / 45.9 (4.47; 4.97)	2.32 / 23.9 (2.23; 2.44)	2.02 / 20.7 (1.82; 2.13)
Cl-rich	0.19 / 1.90 (0.16; 0.20)	0.59 / 5.80 (0.50; 0.70)	1.29 / 13.3 (0.98; 1.71)
Metal processing	0.42 / 4.10 (0.40; 0.50)	0.31 / 3.20 (0.27; 0.40)	/ /
Biomass burning	3.36 / 32.7 (3.06; 3.42)	4.45 / 45.8 (4.28; 4.57)	2.26 / 23.3 (2.01; 2.38)
Vehicle exhaust	1.50 / 14.6 (1.23; 1.64)	0.46 / 4.70 (0.00, 0.55)	1.96 / 20.1 (1.46; 2.19)
Cr-rich	0.08 / 0.80 (0.04; 0.11)	0.07 / 0.70 (0.06; 0.09)	0.65 / 6.70 (0.57; 0.70)
Mass Total	**10.3**	**9.71**	**9.71**

Table 1. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset. Leading numbers are given as mean $PM_{2.5}$ mass / per cent of total source mass for each feature, in $\mu g/m^3$; parenthetic numbers are 5% and 95% bootstrap confidence intervals, respectively.

$PM_{2.5}$ Source	INAA-XRF-EC-OC Data		INAA-XRF-EC-OC_{1-5} Data
	INAA-Kevex	**INAA-JV**	**INAA-JV**
Airborne soil	1.00 / 4.82 (0.50; 0.1.15)	1.54; 1.33 / 8.93; 7.71 (1.36; 1.60)	1.54 / 15.9 (1.26; 1.66)
Nitrate	5.90 / 28.4 (5.41; 6.38)	2.09 / 12.1 (1.71; 4.07)	2.02 / 20.7 (1.82; 2.13)
Cl-rich	1.56 / 7.51 (1.05; 1.69)	/ /	1.29 / 13.3 (0.98; 1.71)
Metal processing	0.54 / 2.60 (0.38; 0.60)	0.71 / 4.12 (0.53; 1.39)	/ /
Biomass burning	4.84 / 23.3 (4.59; 5.21)	3.92 / 22.7 (2.36; 5.30)	2.26 / 23.3 (2.01; 2.38)
Vehicle exhaust	5.88 / 28.3 (5.29; 6.32)	7.77 / 45.0 (6.41, 8.42)	1.96 / 20.1 (1.46; 2.19)
Cr-rich	1.04 / 5.01 (0.93; 1.20)	1.22 / 7.07 (0.60; 1.48)	0.65 / 6.70 (0.57; 0.70)
Mass Total	**20.8**	**17.3**	**9.72**

Table 2. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset, for the upper fifth mass fraction. Leading numbers are given as mean $PM_{2.5}$ mass / per cent of total source mass for each feature, in $\mu g/m^3$; parenthetic numbers are 5% and 95% bootstrap confidence intervals, respectively.



	Bio. Burn	Cl-rich	Veh. Exh.	AB Soil	Nitrate	Metal	Cr-rich	EC	OC	NO3	SO4	NH4	naa as	naa br	xrf ca	xrf cl	naa co	naa cr	xrf cu	xrf fe	xrf k	xrf mn	xrf pb	naa sb	naa sc	xrf si	xrf zn
Bio. Burn			0.36	0.52	0.35				0.52		0.33	0.31	0.40	0.47	0.53		0.53			0.58	0.92	0.52	0.44	0.35	0.55	0.56	0.41
Cl-rich	0.39															1.00											
Veh. Exh.					0.44				0.51			0.35	0.43	0.41							0.38		0.53	0.45			0.95
AB Soil															0.97		0.85			0.98	0.75	0.90			0.94	0.98	
Nitrate		0.38							0.98	034			0.38	0.42							0.37			0.38			0.43
Metal																			1.00								
Cr-rich																		1.00									
EC	0.65	0.93			0.30																						
OC	0.93	0.49			0.38			0.69		035		0.32	0.43	0.48							0.52		033	0.43			0.51
NO3		0.59			0.92			0.48	0.47			0.64															
SO4					0.43					035		0.51															
NH4		0.36			0.91			0.36	0.43	0.78	0.46			0.33													0.33
naa as	0.55							0.39	0.60	030		0.30		0.34							0.34		037	0.32			0.39
naa br	0.52			0.36				0.39	0.56				0.41								0.44			0.40			0.40
xrf_ca				0.97										0.36			0.83			0.93	0.75	0.85			0.93	0.95	
xrf cl		0.38						0.39	0.37	037		0.30		0.27													
naa co				034			0.59								0.33					0.85	0.72	0.80			0.90	0.85	
naa cr							1.00										0.59										
xrf cu						1.00			0.30																		
xrf fe				0.97										0.39	0.92		0.34				0.80	0.89			0.92	0.97	
xrf k	0.43		0.31	0.74					0.41				0.32	0.57	0.65					0.72		0.75	037	0.30	0.75	0.78	0.43
xrf mn				0.87											0.81					0.86	0.65				0.86	0.88	
xrf pb			0.99																		0.31			0.36			0.46
naa sb							036											0.36									0.43
naa sc				0.91			036							0.32	0.89		0.52	0.36		0.90	0.62	0.79				0.94	
xrf_si				0.97										0.31	0.93		0.34			0.95	0.67	0.83			0.91		
xrf zn	0.42	0.41						0.45	0.47				0.30			0.32					0.32						

Table 3. Pearson correlations for INAA-Kevex-EC-OC and INAA-JV-EC-OC data, for sources and species. The left half corresponds to INAA-JV-EC-OC correlations, while the right half represents INAA-Kevex-EC-OC data.

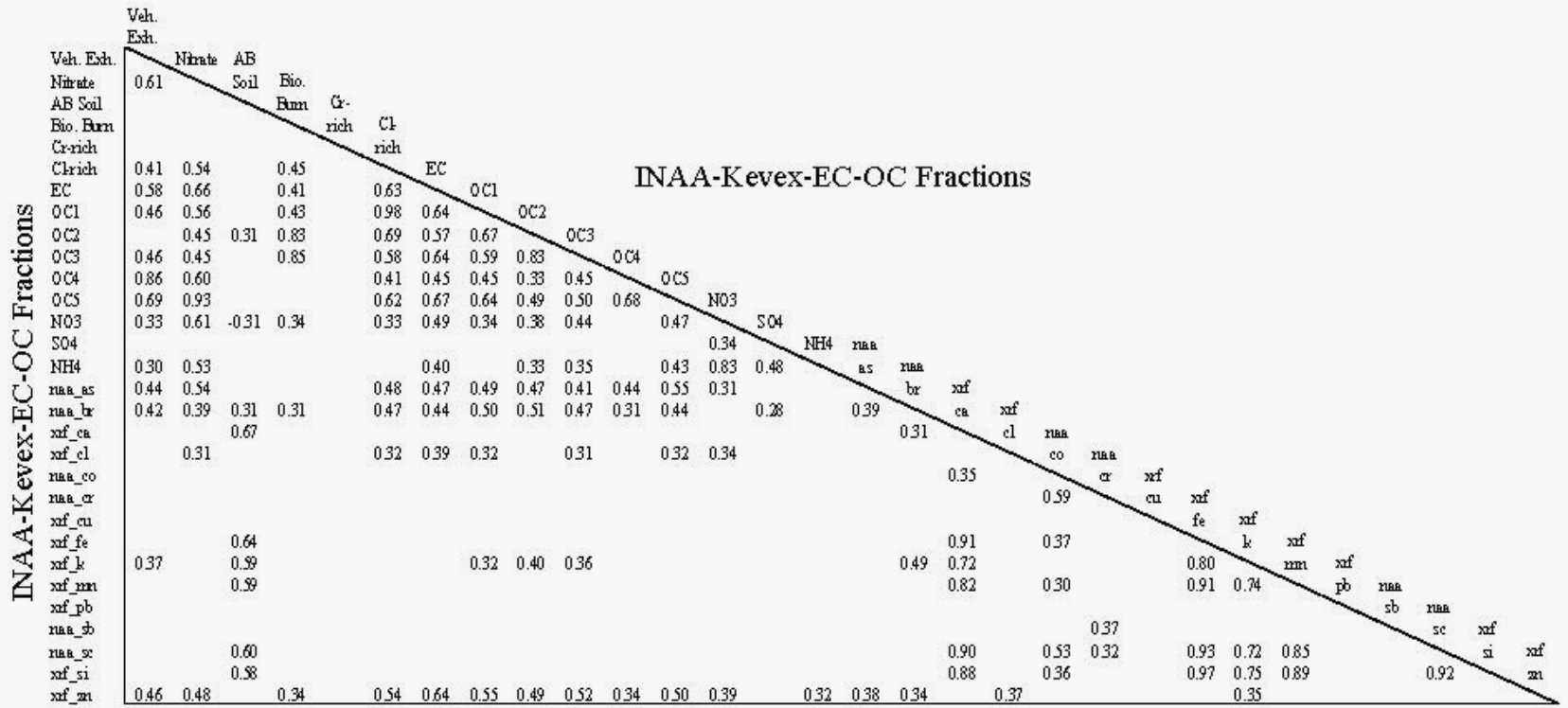


Table 4. Pearson correlations for INAA-JV-EC-OC fractions, for sources and species.

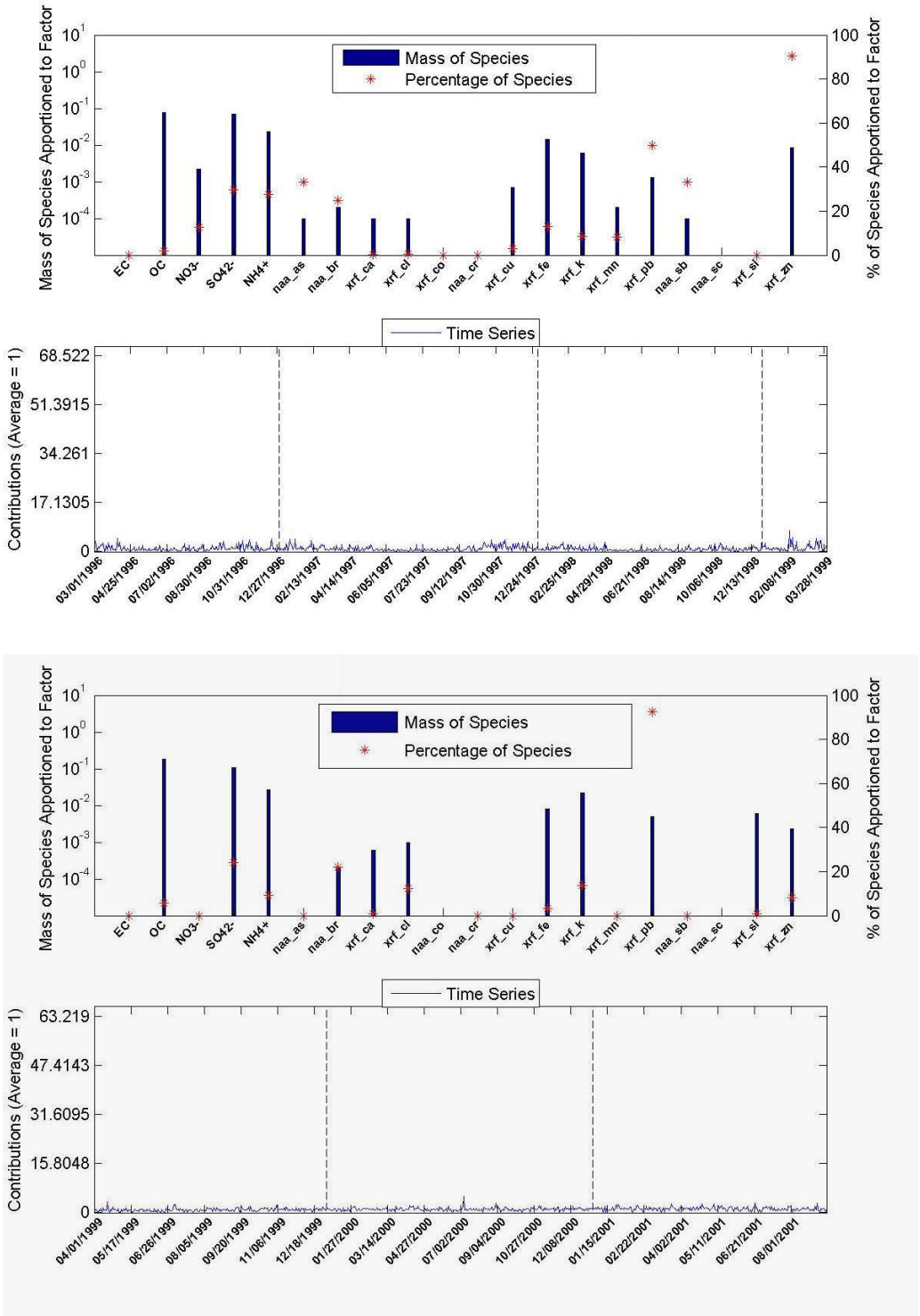


Figure 1. Vehicle exhaust source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by both EC and OC, as well as Pb, Zn, and Br.

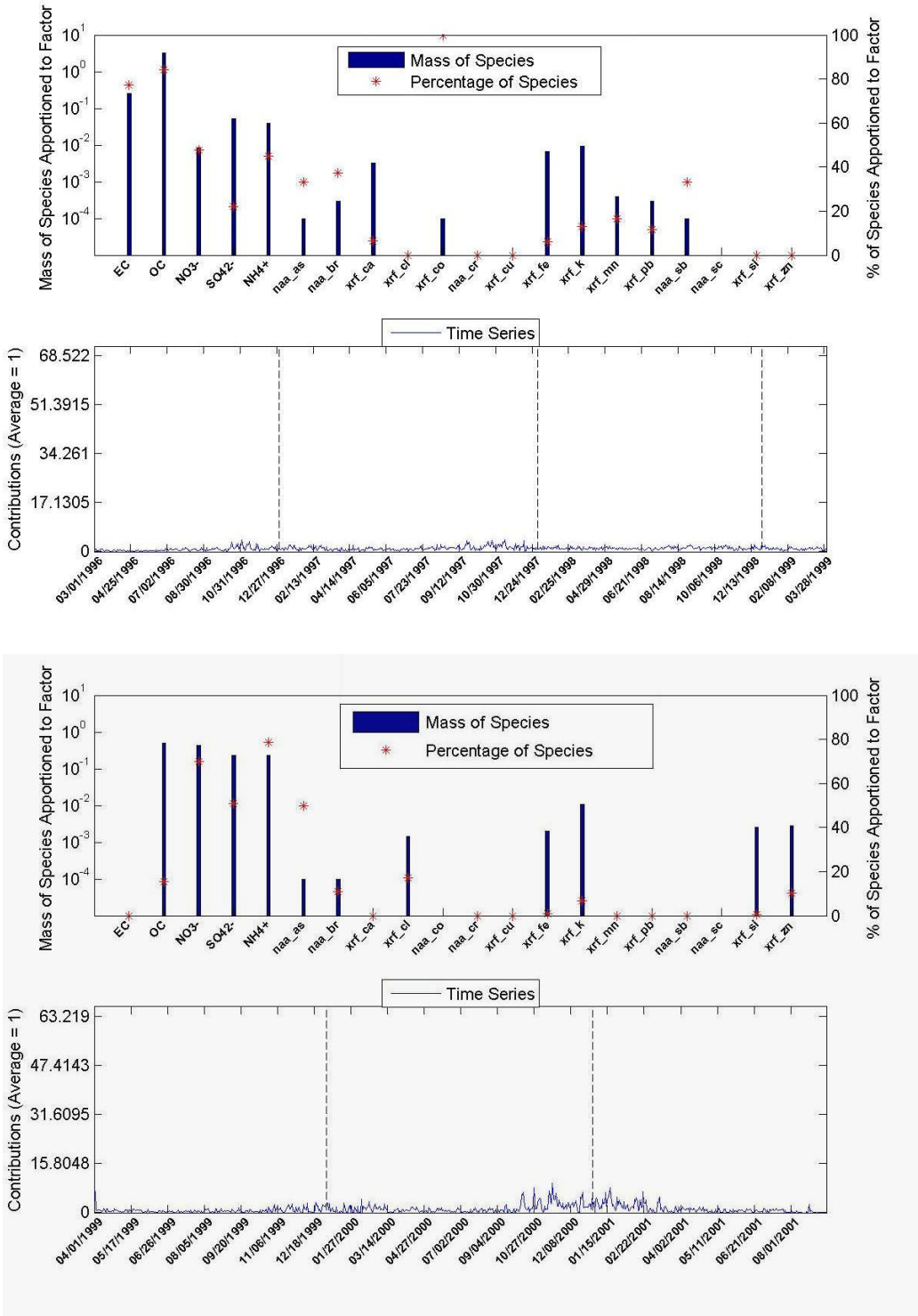


Figure 2. Nitrate source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by OC, ions, and geological material.

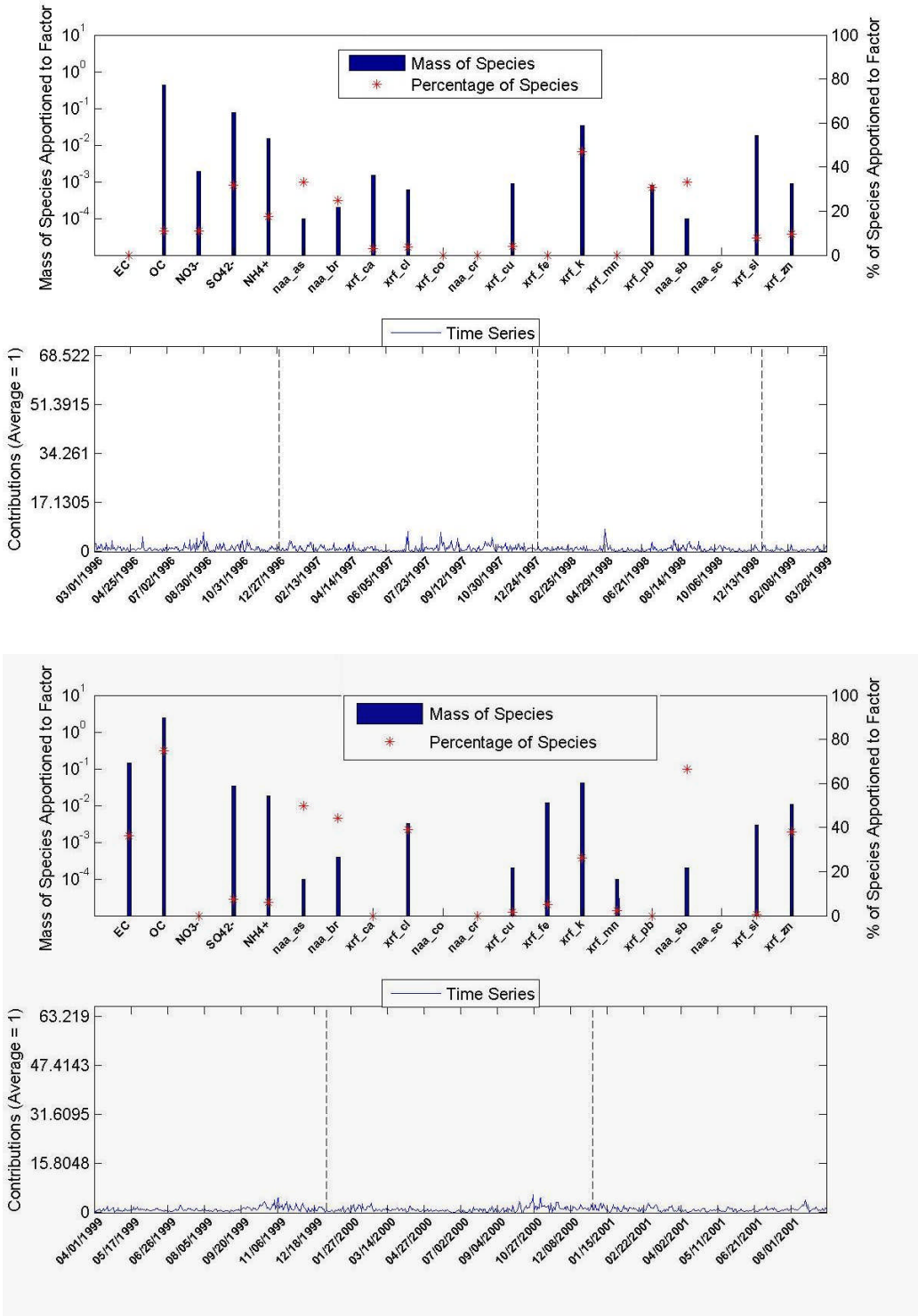


Figure 3. Biomass burning source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by OC, maybe EC, K, As, Sb, and geological material.

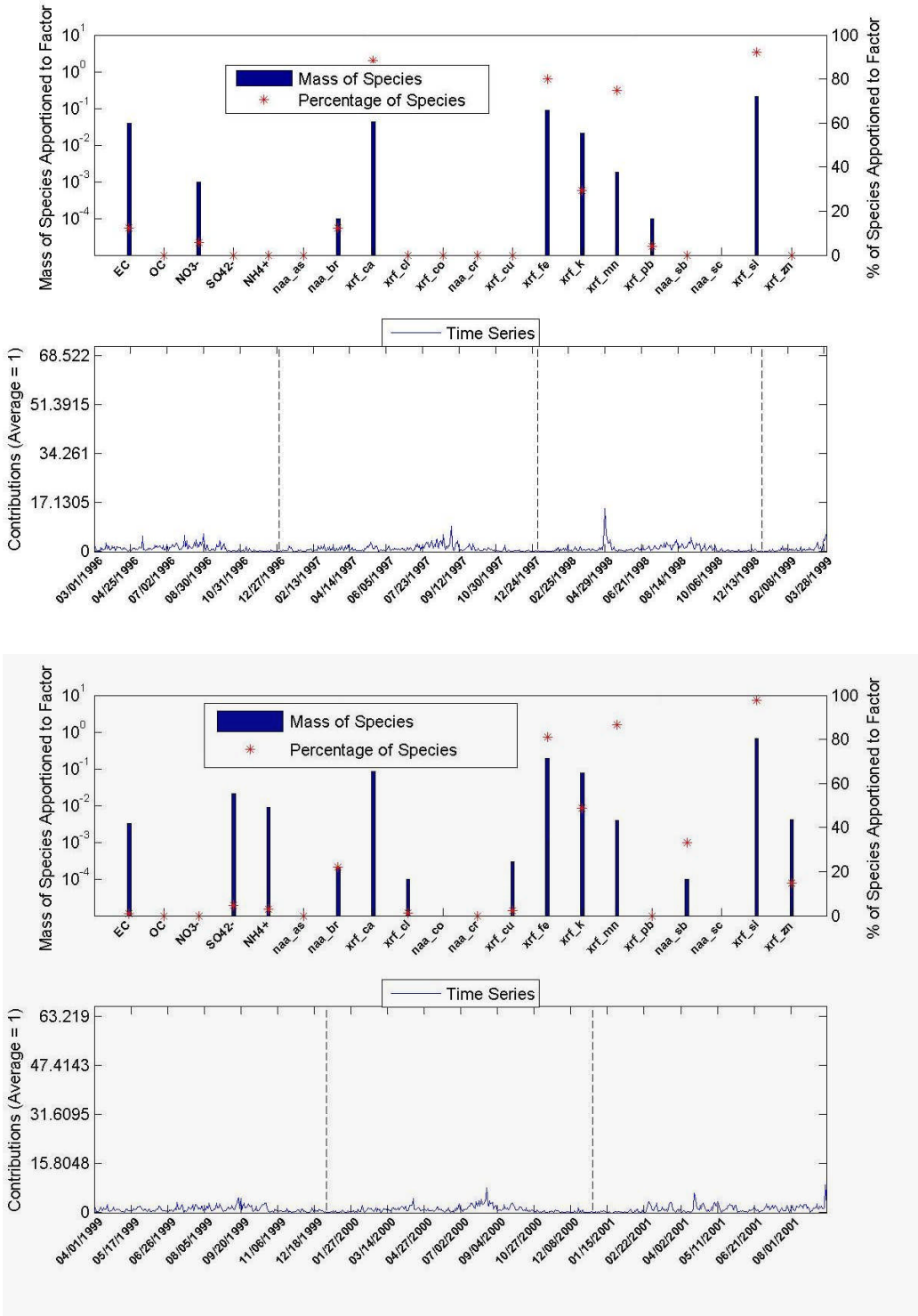


Figure 4. Airborne soil source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by possibly some EC, and geological materials.

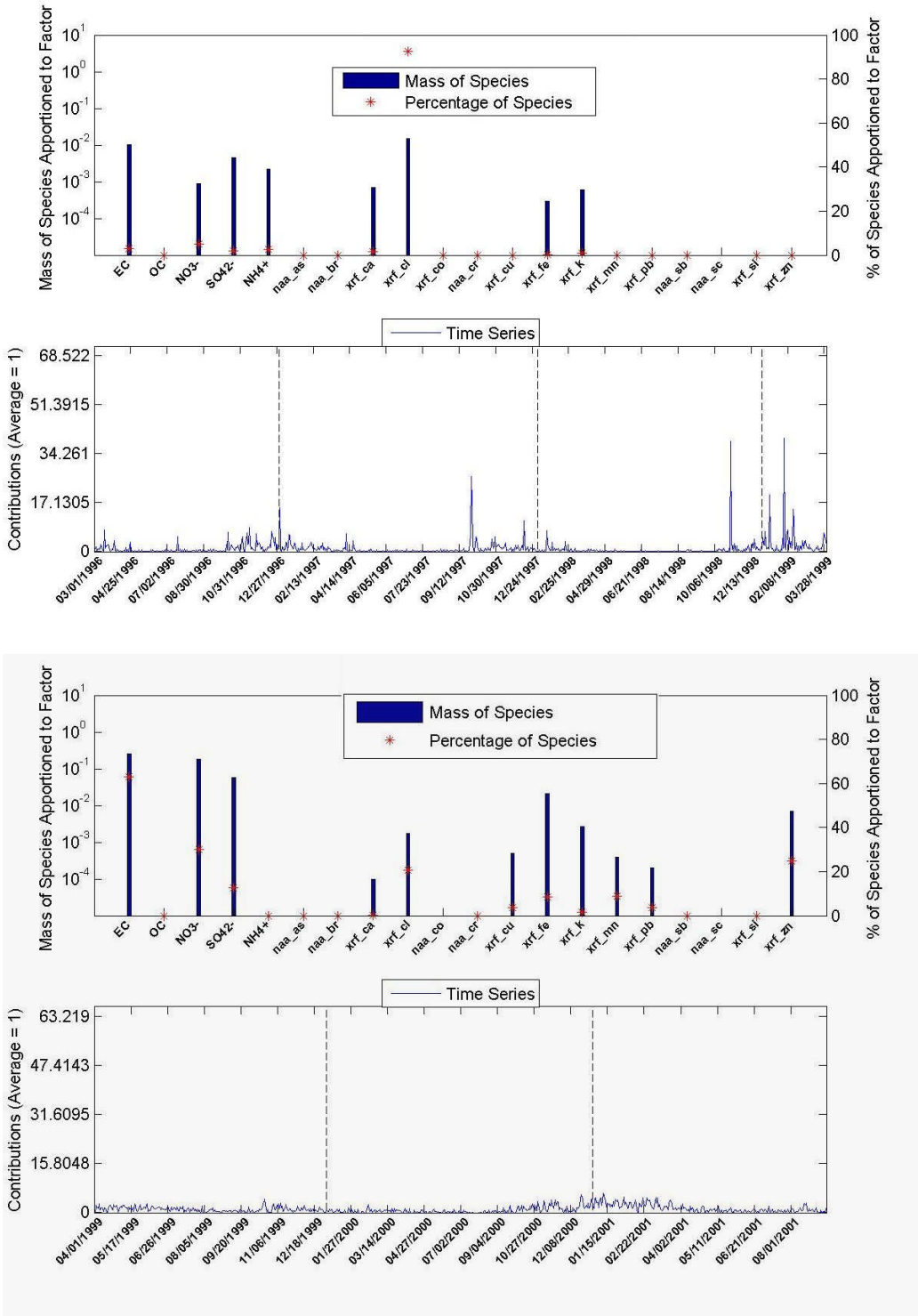


Figure 5. Cl-rich source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by EC, Cl, and geological materials.

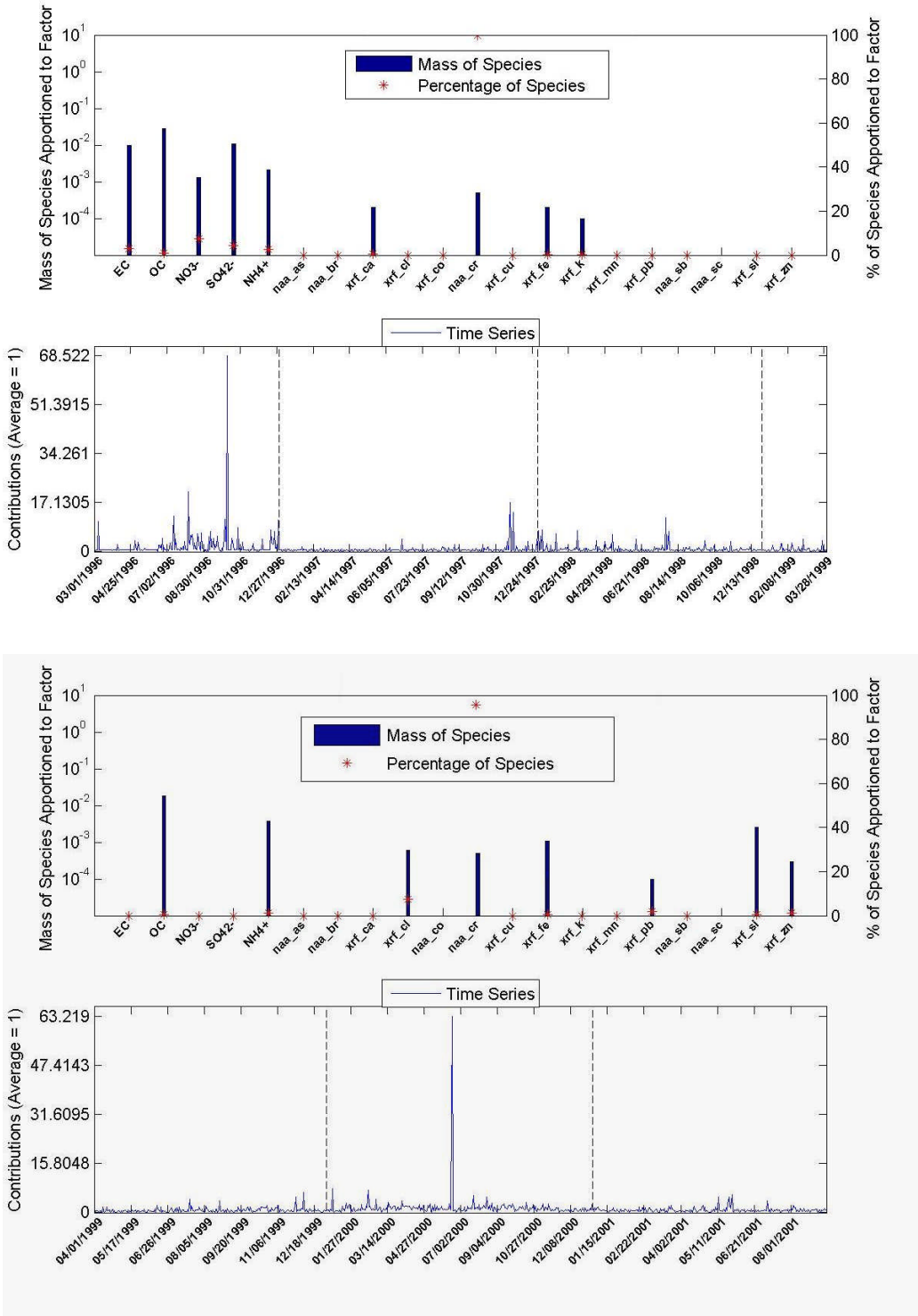


Figure 6. Cr-rich source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by possibly some OC and Cr.

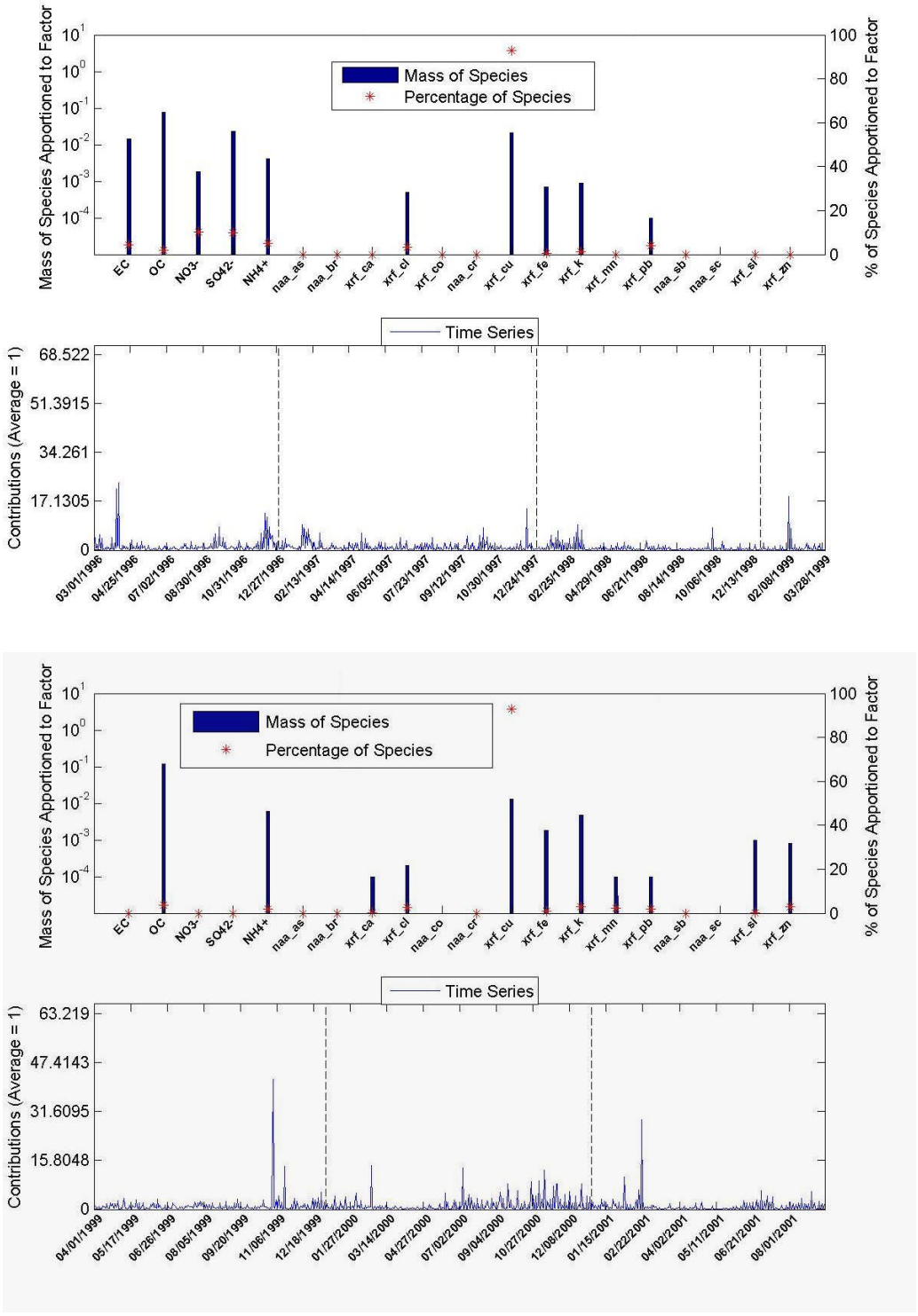


Figure 7. Metal processing source for INAA-Kevex (top) and INAA-JV (bottom). This source is characterized by OC and possibly some EC and Cu.

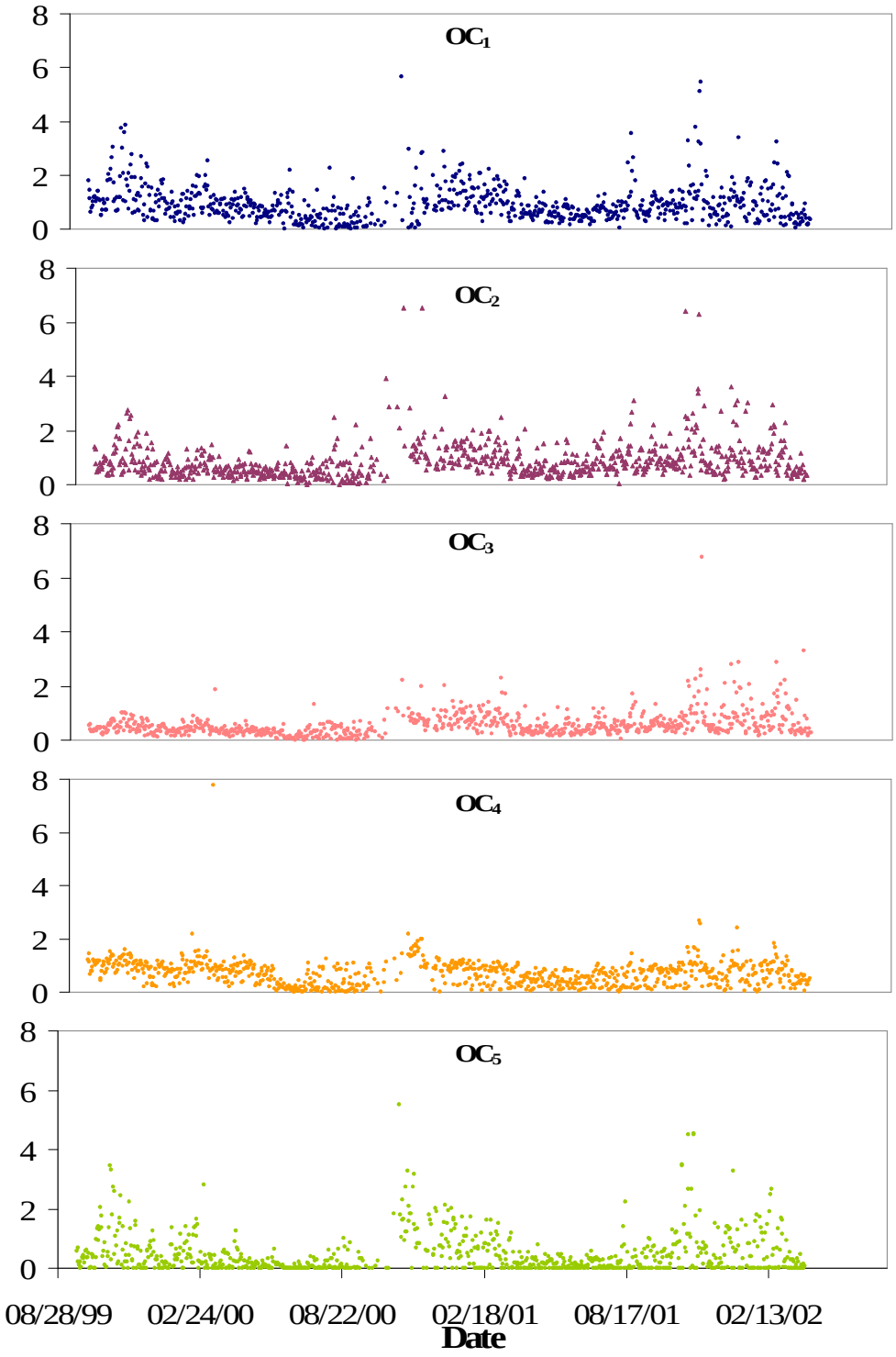


Figure 8. Organic carbon fractions over time. Note that OC_1 and OC_2 generally follow the same pattern, as do OC_3 and OC_4 with small peaks unique to each. OC_5 has its own trends, which suggest specific sources.

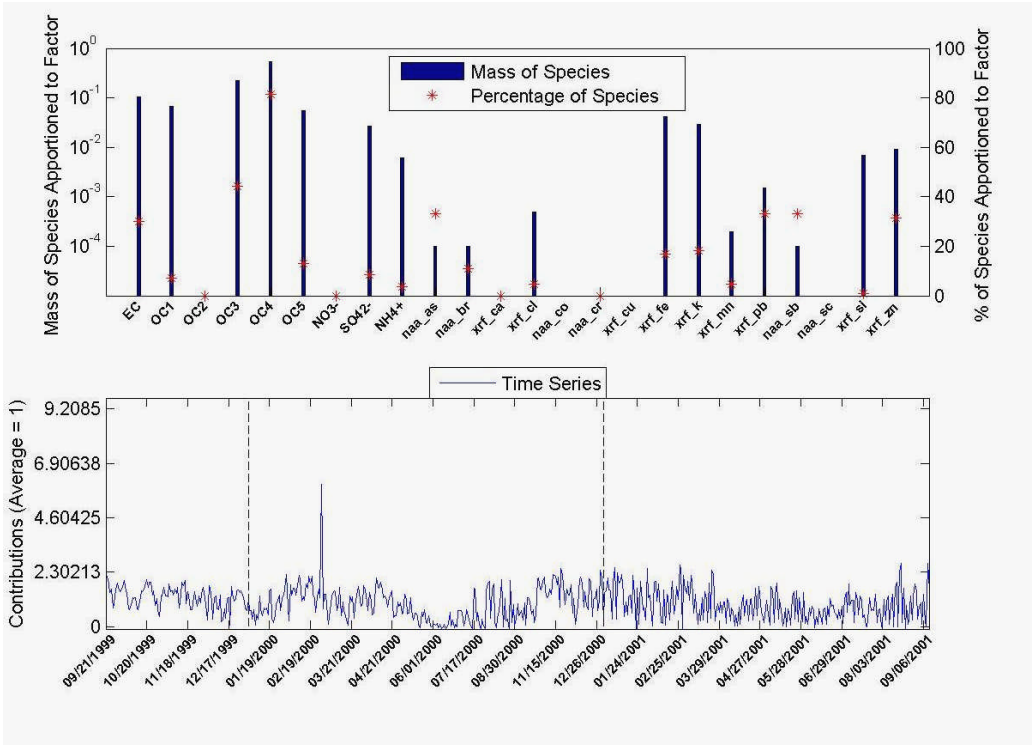


Figure 9. Vehicle exhaust source for INAA-JV with EC and OC fractions. This source is characterized by a little EC, OC_3, OC_4, some OC_5 and OC_1, and Pb, Br, and Zn.

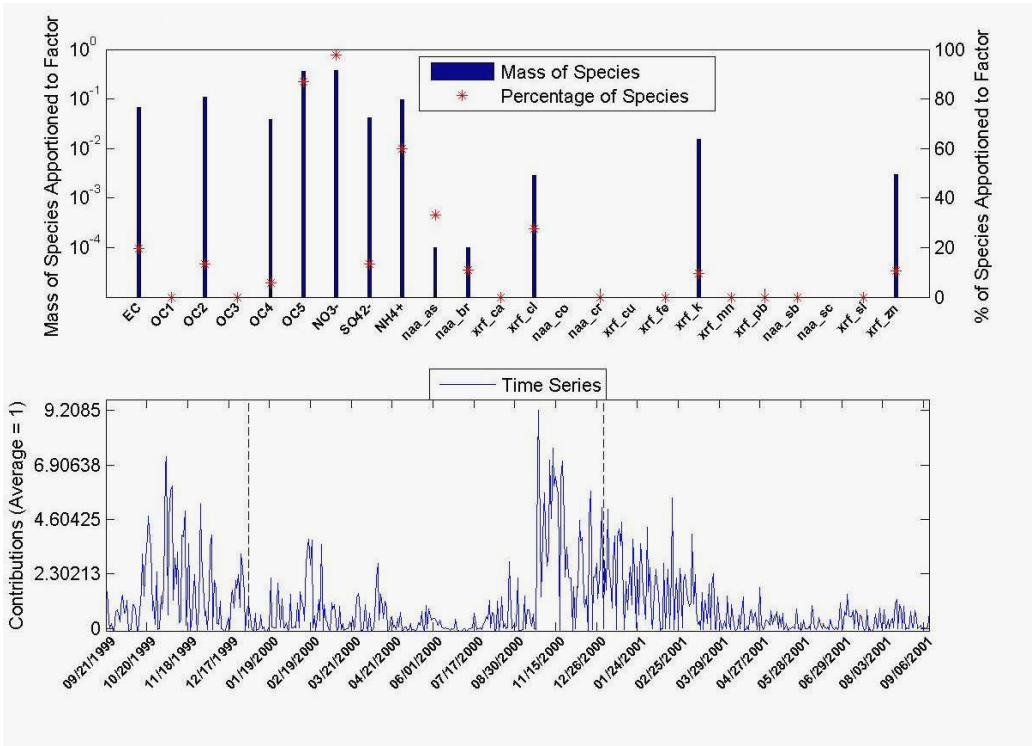


Figure 10. Nitrate source for INAA-JV with EC and OC fractions. This source is characterized by a little EC, OC_2, OC_4, and a lot of OC_5, with NO_3^-, SO_4^{2-}, NH_4^+, and K.

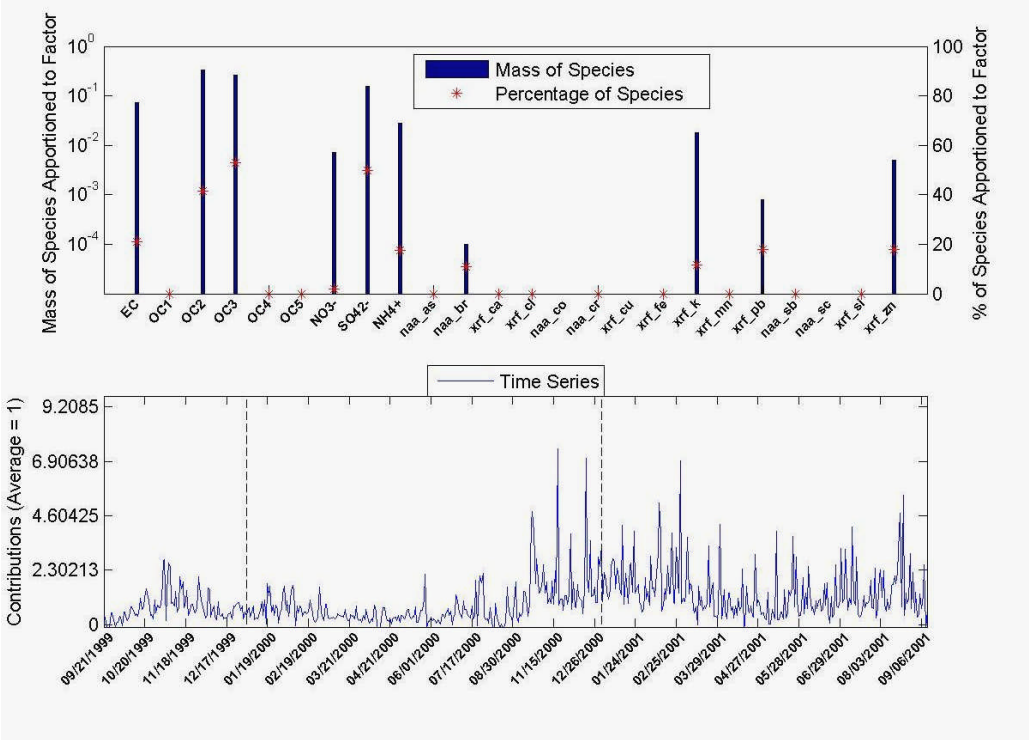


Figure 11. Biomass burning source for INAA-JV with EC and OC fractions. This source is characterized by EC, OC_2, OC_3, and SO_4^{2-}, NH_4^+, K, and some geological materials.

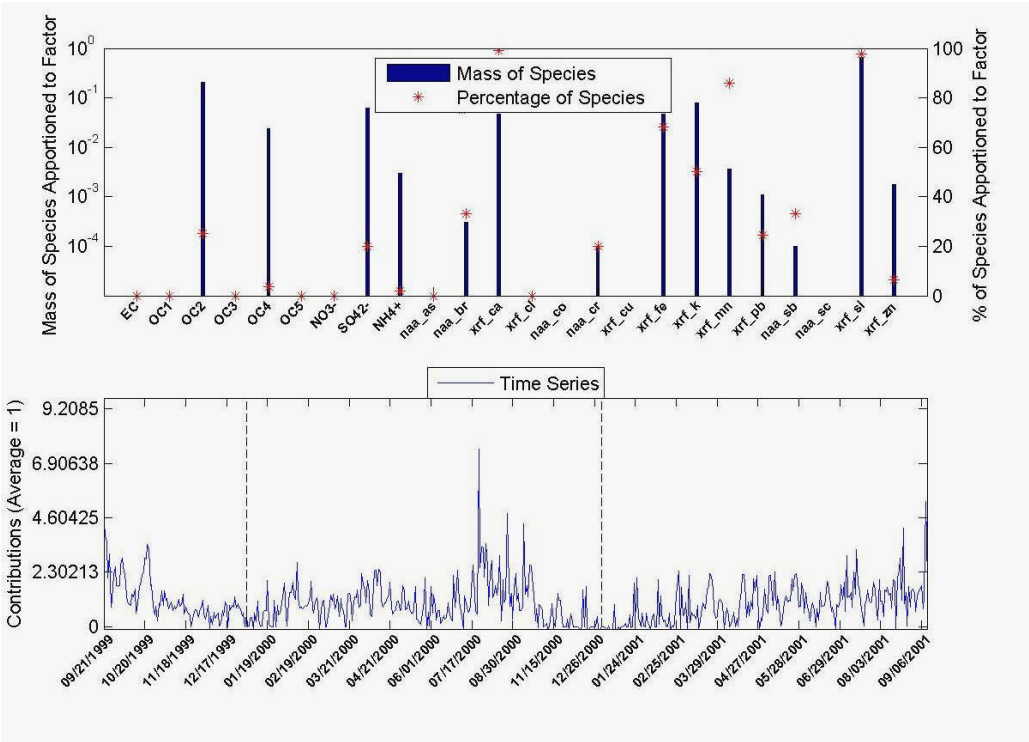


Figure 12. Airborne soil source for INAA-JV with EC and OC fractions. This source is characterized by a little OC_2 and OC_4, SO_4^{2-}, and geological materials.

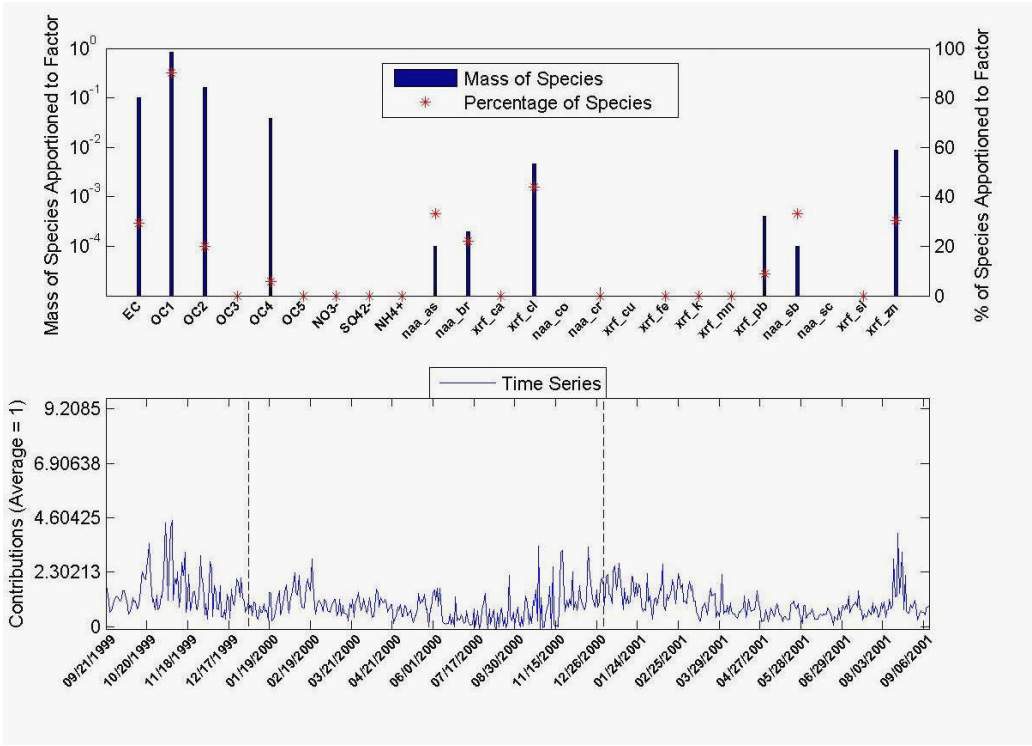


Figure 13. Cl-rich source for INAA-JV with EC and OC fractions. This source is characterized by EC, OC_1, OC_2 a little OC_4, and Cl.

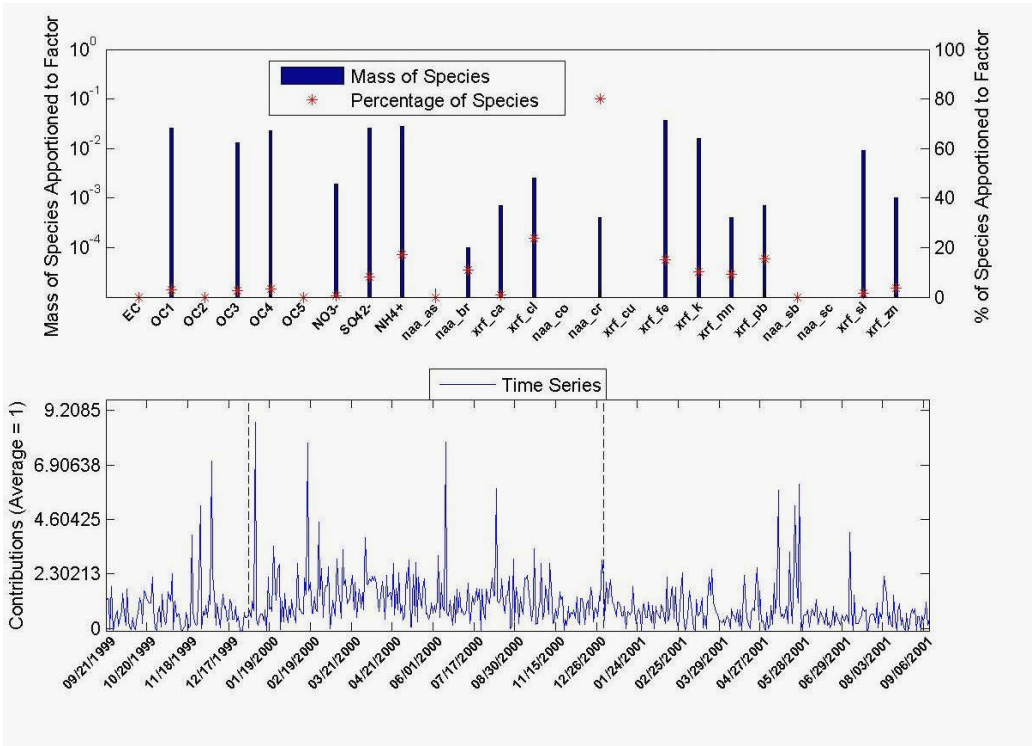


Figure 14. Cr-rich source for INAA-JV with EC and OC fractions. This source is characterized by a little OC_1, OC_3, and OC_4, Cr and geological materials.

CHAPTER 5: SUMMARY AND CONCLUSIONS

The preceding chapters stress several important facets of $PM_{2.5}$ source apportionment of a large, instrumentally-intensive dataset. Each level of specie detail provides uniquely significant features of the $PM_{2.5}$ profile. Each instrument confirms realistic attributes given by other measurements, and contributes new levels of information to the overall profile. Every measurement technique required specific input information, specifically the Jordan Valley data, and sensitivity to corrections and detection limits. Kevex was a reliable source of many trace elements, on which the bulk of the source profile features was based. The INAA method was especially useful for detecting very low quantities of trace elements. Detailed carbon information provided insight into what types of combustion are associated with the source features.

XRF, INAA, ionic, and carbon species in conjunction with PMF describe seven definite source features of pollution: motor vehicle combustion; chlorine-emitting combustion, from the municipal incinerator; biomass burning, from agriculture; airborne soil, from agriculture, traffic, and climate; metal processing, from a metal smelter; nitrate, primarily a winter time feature; and a Cr-rich source, likely from an industrial metal plant. **Table 1** summarizes all the PMF results discussed in chapters 2-4. The Kevex and Jordan Valley data treat the features somewhat differently as separate datasets; however, the same markers can be found in these results and agree with earlier work. The addition of INAA data clarified the discrepancies between the two XRF instruments and added valuable information to the profile, including the introduction of the Cr-rich source.

Several species were key in identifying source features. The crustal elements helped determine airborne soil, an especially strong feature during the late summer. Pb, Cl, and Cu separate carbon rich sources into very different and significant anthropogenic pollution sources — vehicles, municipal incineration, and metal processing. Similarly, Cr has proven to be an exclusive identifier of a new source, a specialty metal processing plant. Sb and As assist in determining mostly wintertime combustion, nitrate, and biomass burning. While many of these elements were found only in trace amounts, they can be diverse in origin and physiologically detrimental. Combining trace element analyses in source apportionment is a valuable tool in capturing the full fine particulate landscape.

The addition elemental and organic carbon details (EC and OC, respectively) confirmed prior results by adding valuable EC-OC markers. Metal processing and nitrate were found to have organic carbon, while biomass burning, vehicle exhaust, Cl-rich and Cr-rich sources were found to contain both organic and elemental carbon. There was no considerable carbon at all with airborne soil, generally in accordance with the previous trace element studies. For the major combustion sources, the organic fractions provided further clarification for carbon in the source features. OC_2 and OC_3, lighter mass groups, were present for biomass burning, as was the case for the Cl-rich source, which showed OC_1 in its profile. Vehicle exhaust and nitrate contained heavier organic fractions, OC_3 and OC_4, and OC_5, respectively. These specific mass groups give indication of what the carbon fuel sources were, and helped confirm whether sources were correctly identified by the previous trace element models.

In spite of the information each chapter has added to the overall profile, some problematic data were not resolved. The Jordan Valley Cl was subjected to several uncertainty calculations and instrumental corrections, none of which effectively separated its confusion with the vehicle exhaust. The addition of INAA and carbon data helped clarify some of its influence on the profile and identify it as its own feature; however, the Cl feature isolated some mass most likely attributed to exhaust. This limitation of the Jordan Valley does not hamper the overall effectiveness of concluding a Spokane $PM_{2.5}$ profile, but JV data should be used in concert with other trace element analyses in order to recognize this and other possible instrumental hindrances.

In future studies, this work could be enhanced by adding several complementary studies. Several parameters collected in the dataset were not used for this thesis, because of the narrow scope for this research. All National Ambient Air Quality Standards (NAAQS) criteria pollutants were measured during the study period, and this thesis focused only on $PM_{2.5}$. An appendage to this particulate study could be a look at how the other criteria pollutants track with the $PM_{2.5}$ data. The gaseous criteria pollutants – SO_2, CO, NO_2, and O_3 – could be included in conditional probability function studies in order to investigate co-existence of pollutants, and their transport downwind from sources to the receptor.

One special aspect of this criteria pollutant project would be to add a PM_{10} source study, which would investigate the features of a coarse fraction profile. There is a relationship between coarse and fine fraction particles (Hien et. al, 2004). To what extent this relationship exists for the Spokane dataset is not known. Understanding the relationship between the two size fractions could be useful in helping researchers draw

conclusions from more limited data resources such as smaller collection projects, historical data, or current PM monitoring stations. The larger fraction study would also complement the fine fraction research by adding another layer of detail about many PM constituents (such as the criteria pollutant Pb, or other toxic elements). In addition to its similarities to fine fraction PM, this companion project could also determine what unique sources there are in the coarse fraction range. While $PM_{2.5}$ is recognized as the size fraction more influential on health, PM_{10} can interfere with visibility, quality of life, and economic welfare for those in an affected area.

Another possibility includes incorporating the latest carbon measurement methods into the dataset for a third-tier carbon comparison. Thermal/Optical Reflectance is similar to TOT, and is regarded favorably in the air quality field (El-Zanan et. al, 2005; Ye et. al, 2003). While this study included two types of instruments for total carbon and the elemental/organic split, it did not address any instrument shifts in the organic fraction computations. The introduction of a newer method could be useful in overcoming instrumental limitations in fully resolving the elemental and organic carbon data.

Daily weather conditions can also shed light onto why profiles may fluctuate as they do. This thesis focuses only on the source profile as seen at the receptor site. One potentially large modifier of the ambient PM is the weather it encounters between emission and collection. Consideration for daily weather may clarify the relationship between emission and potential for exposure. Other PMF air quality studies have incorporated weather parameters to help identify sources of pollution and to clarify pollutant behavior (Hien et. al, 2002; Juntto and Paatero, 1994; Paatero and Hopke,

2002). Given such a wealth of information in this dataset, the potential knowledge gained from adding these parameters is significant.

Research into $PM_{2.5}$ for Spokane has successfully disclosed several details about particulate air quality. The results from this body of work indicate several features in the particulate profile, with unique details regarding vast instrumental methods employed, and identify both temporal shifts and long-term trends in pollution. This study, while complete in its own form, is but one piece of the puzzle that is Eastern Washington air quality.

References

El-Zanan, H.S., Lowenthal, D.H., Zielinska, B., Chow, J.C., Kumar, N., 2005. Determination of the organic aerosol mass to organic carbon ratio in IMPROVE samples. Chemosphere 60, 485-496.

Hien, P.D., Bac, V.T., Tham, H.C., Nhan, D.D., Vinh, L.D., 2002. Influence of meteorological conditions on $PM_{2.5}$ and $PM_{2.5-10}$ concentrations during the monsoon season in Hanoi, Vietnam. Atmospheric Environment 36, 3473-3484.

Hien, P.D., Bae, V.T., Thinh, N.T.H, 2004. PMF receptor modeling of fine and coarse PM_{10} in air masses governing monsoon conditions in Hanoi, northern Vietnam. Atmospheric Environment 38, 189-201.

Juntto, S., Paatero, P., 1994. Analysis of Daily Precipitation Data by Positive Matrix Factorization. Environmetrics 5, 127-144.

Paatero, P., Hopke, P., 2002. Utilizing wind direction and wind speed as independent variables in multilinear receptor modeling studies. Chemometrics and Intelligent Laboratory Systems 60, 25-41.

Ye, B., Ji, X., Yang, H., Yao, X., Chan, C.K., Cadle, S.H., Chan, T., Mulawa, P.A., 2003. Concentration and chemical composition of $PM_{2.5}$ in Shanghai for a 1-yr period. Atmospheric Environment 37, 499-510.

$PM_{2.5}$ Source	XRF Data		INAA-XRF Data		INAA-XRF-EC-OC Data		INAA-XRF-EC-OC1-5 Data
	Kevex	JV	INAA-Kevex	INAA-JV	INAA-Kevex	INAA-JV	INAA-JV
Airborne Soil	1.06 / 0.53 / 0.49 9.70 / 5.00 / 4.60	1.71 / 1.07 19.9 / 10.5	0.74 / 0.16 7.80 / 1.50	1.63 / 1.63 20.0 / 15.4	0.00 0.00	1.51 15.5	1.54 15.9
Nitrate	2.96 / 2.31 / 1.81 27.0 / 21.5 / 17.0	1.78 / 1.66 18.5 / 16.4	0.83 / 2.79 8.70 / 25.1	0.81 / 2.55 9.90 / 24.3	4.71 45.9	2.32 23.9	2.02 20.7
Cl-rich	0.91 / 0.66 / 0.12 8.20 / 6.20 / 1.00	1.19 / 4.70 13.9 / 46.1	0.75 / 0.16 7.80 / 1.40	3.35 / 0.55 41.1 / 5.30	0.19 1.90	0.59 5.80	1.29 13.3
Metal processing	0.55 / 0.95 / 0.44 5.00 / 8.90 / 4.20	0.50 / 0.40 5.82 / 3.92	0.77 / 0.37 8.10 / 3.40	0.25 / 0.25 3.10 / 2.40	0.42 4.10	0.31 3.20	/ /
Vegetative burning	3.71 / 2.66 / 5.30 33.8 / 24.9 / 49.8	1.04 / 1.36 12.1 / 13.3	2.29 / 5.18 24.0 / 46.6	0.34 / 4.67 4.20 / 44.3	3.36 32.7	4.45 45.8	2.26 23.3
Vehicle exhaust	1.81 / 3.60 / 2.48 16.4 / 33.7 / 23.3	2.37 / 0.98 27.6 / 9.60	4.02 / 2.38 42.1 / 21.4	1.75 / 0.00 21.5 / 0.00	1.50 14.6	0.46 4.70	1.96 20.1
Cr-rich	/	/	0.16 / 0.08 1.60 / 0.70	0.01 / 0.87 0.20 / 8.30	0.08 0.80	0.07 0.70	0.65 6.70
Source Mass Total	11.0 / 10.7 / 10.6	8.59 / 10.2	9.56 / 11.1	8.16 / 10.5	10.3	9.71	9.71

Table 1. Mean contributions and bootstrap confidence intervals for source features, for each PMF bootstrap subset. Leading numbers are mean $PM_{2.5}$ mass for each feature, in µg/m3; 2nd row numbers are per cent contributions to the total source mass explained.

APPENDIX

The appendix contains supplemental information not required for submission of a research proposal, but is deemed relevant for consideration of the research. This additional information will help clarify research details for those who have little or no prior knowledge of these pollution modeling tools or the unique aspects of the dataset.

Summary statistics were calculated during the quality filtering and are included in **Table A1**. The statistics summarize the arithmetic and geometric means, as well as the frequency of each specie below its respective detection level, which were subject to modification before being modeled with PMF. The Kevex and Jordan Valley (JV) instruments have similar levels of abundance for the species presented in this study. For Kevex, chlorine, manganese, bromine, and lead are the species with the lowest abundance, all of which are present above detection for the majority of the study. For Jordan Valley, bromine and aluminum have relatively poor abundances above detection relative to other species; in the case of aluminum, the element is not present above detection the majority of the time. In spite of these limitations, ample data is available to characterize sources of pollution in Spokane and understand the changes in analysis due to instrument transition.

For the XRF data, minimum detection limits (MDLs) were taken from previously published values or determined in the laboratory. The MDL values for the Kevex species given in Appendix **Table A2** are taken from Kellogg (1994), which discusses the Kevex instrument and its characteristics. In the case of the JV XRF information, limits of detection were not provided by the instrument manufacturer, and were therefore

determined by test blank data. Test blanks were analyzed along with standards and samples to ensure quality results. For each specie, an average of the blank concentration values was used to represent a lower detection of the instrument. A summary of these calculated detection limits are given in **Table A2**. In the case of two or more targets used to determine blanks, the larger of the two detection limits was used, to take advantage of the effect on uncertainty. In each case of a single specie having multiple MDLs, the MDLs were similar to one another, further supporting this method.

One quality test performed was a reconstructed $PM_{2.5}$ mass, which was compared to the measured PM mass. In a reconstruction test, the trace elements and the particulate species are summed according to the following equations and compared to the VAPS and TEOM mass quantities to point out any major discrepancies. The basic equation for fine fraction PM is as follows:

$$PM_R = 1.4OC + EC + NO_3 + (NH_4)_2SO_4 + \sum XRF \qquad (1)$$

where *OC* is the organic carbon contribution and is multiplied by a factor of 1.4 to account for the average mass ratio of bound hydrogen and oxygen to the carbon, *EC* is the elemental carbon concentration, NO_3^- is the nitrate, $(NH_4)_2SO_4$ is the particle form of ammonium sulfate, and *XRF elements* are the trace elements on the filters, analyzed by X-ray Fluorescence whose mass contribution is given by:

$$\sum XRF = 1.89Al+2.14Si+1.21K+1.4Ca+1.67Ti+1.29Mn+1.43Fe+\sum other \qquad (2)$$

where the multipliers are used to account for the bound oxygen for the common soil components.

This series of equations takes into account all particulate pollution analyzed in this study and most of the particulate matter that would be captured on a filter. Within this reconstruction technique, there were two calculations of $(NH_4)_2SO_4$ made. The first calculation used the sulfate data, and mass was computed via the ratio of sulfate to ammonium sulfate. The second calculated the mass ratio of ammonium sulfate to sulfur, using XRF-derived sulfur.

$$(NH_4)_2SO_4 = SO_{4Quartz\ filter} * ratio\ of\ ((NH_4)_2SO_4:S0_4) \quad (3a)$$

$$(NH_4)_2SO_4 = XRF_S * ratio\ of\ ((NH_4)_2SO_4:S) \quad (3b)$$

In either case, it is assumed that all of the sulfur is tied up in the ammonium sulfate, so in *Equation 3*, the XRF sulfur is intentionally not added to the overall mass. *Equations 1-3b* are collectively taken from previous work to approximate soil contributions to PM and are discussed in more detail in Gray et al. (1986), Chow et al. (1994), and Norris (1998).

Poor or suspicious data can stem from instrumental problems, filter storage, and filter contamination. Instrumental problems can refer to failure of sensors or data loggers, overloading the samplers, or human error. In cases of these such interferences, whole samples are considered suspicious because the total PM mass is questionable. Filter storage can lead to loss of volatilization of compounds. Handling can lead to loss or gain of mass, and skew one or more, or all, the specie concentrations. For example, skin

contact of the exposed filter surface can transfer salts from sweat or oils onto the filter. Poor quality lab tools can contaminate filters with metals. Laboratory quality checks reveal small incidences of flagged data of the following: Na, Ba, Br, Cl, Cs, Ni, Si, Sr, W, Br, Zr, Ca, K, Fe, Rb. The flags were either due to specific contamination or very high values with no justifiable cause. In spite of all the potential for suspect data, these instances encompass less than 4% of the overall dataset.

The reconstructed data groups were plotted against each other to graphically assess the level of agreement between modeled and actual PM. Included in this check were the two reconstructions set against one another, each reconstruction with the $VAPS_{2.5}$ data, and $VAPS_{2.5}$ data with $TEOM_{2.5}$. The two sulfur reconstruction methods are plotted against each other, to display any obvious outliers in either the XRF sulfur or sulfate data. Each reconstruction is also plotted against the actual VAPS $PM_{2.5}$ mass, to pinpoint any cases where either the VAPS may be wrong or the reconstructions may not be valid. In all cases, the data are separated into yearlong groups. For each plot, suspicious data were logged for further scrutiny. Samples with pre-existing flags were either taken out or their uncertainties increased. Those suspect points with no prior flags were considered individually, and either eliminated or flagged in case of influence on the model results.

Graphical analyses of the reconstructions indicate strong data quality for the most of the study. The following graphs (**Figure 1**) are the data quality plots for the first year of study, 1995, and are representative of the overall reconstruction analysis. The data are generally very good, and the questionable points are usually very deviant from the norm. Results indicate solid agreement with the reconstruction and measured massed. The

majority of the outlier samples have corresponding flags with their sampling or analysis methods. In total, roughly 4% of samples tested had species within them that were recommended to be altered or taken out entirely. The level of quality data also increases over time; that is, much of the discord in the data occurs early in the study, suggesting that part of the variability is due to operators learning to use the equipment properly and/or correct calibration and maintenance methods.

Another series of sensitivity tests were performed for this research, to ensure the source features were accurate, and not superficially imposed by the models. For the XRF data, before thesis modeling began, a replication was performed on the work by Kim et. al (2003), which included ionic, total carbon, and Kevex XRF data from the first three years of the dataset. This reconstructed PMF study was done in order to understand how the PMF model responded to the early years of the dataset and that there was correspondence in assumptions in the kin studies. PMF was used for this analysis; in the beginning, the exact same settings were used, and were gradually relaxed to understand the robustness of the model. The resulting profiles were similar but not identical, attributed to the random nature of the model, and small differences in data filtering.

The Kim et. al reconstruction was expanded to include all Kevex data years (1995-1999) and deletion of selected elements. This premise behind this sensitivity test was to confirm the robustness of the Kevex data. Jordan Valley XRF data had known problems; the instrument was especially bad for elements Al, Mn, and Pb. Because Kevex was used as the standard for correction, it was important to understand the unique influence these problem species had on the PMF model, in a situation where the elements were deemed “good.” Al, Pb, and Mn were individually removed and then in

combination, and the resultant dataset modeled by PMF to determine whether their exclusion would hinder the profile. In sources where an element was a key feature, such as Al and Mn for airborne soil, or Pb for vehicle exhaust, those features were still identifiable, and other source features were not influenced. Based on the PMF model outcome and the reported instrumental issues, it was decided that Al would be eliminated from the dataset, and Pb and Mn would stay.

One post-analysis tool added to PMF version 1.1 is the bootstrapping program. This program is intended to provide some measure of how well the designed model predicts the behavior of the data. For each chapter, data were broken into groups of roughly 500 samples each. The $PM_{2.5}$ masses were added to data and down weighted by a factor of thirty (in order to show up in the profile without affecting the profile structure). A large number of runs are selected in the bootstrap, random starting points with which the bootstrap program attempts to match runs to the original output. The variation resulting from matching these runs to the original are plotted in **Figures A2-A21**. These variations indicate the uncertainty associated with each source feature, based on the goodness of the model.

Specie	Cases (N =2663)	Selected (%)	Replaced (%)	Geometric Mean (µg/m^3)	Arithmetic Mean (µg/m^3)
TC	2370	89	11	6.581	8.565
NO3	2183	82	18	0.486	0.698
SO4	2546	96	4	0.696	0.856
NH4	1592	60	40	0.349	0.416
Kevex Instrument					
Specie	**Cases (N =1515)**	**Selected (%)**	**Replaced (%)**	**Geometric Mean (µg/m^3)**	**Arithmetic Mean (µg/m^3)**
Al	1243	82	18	0.102	0.145
Si	1477	97	3	0.166	0.290
Cl	927	61	39	0.016	0.025
K	1506	99	1	0.062	0.080
Ca	1384	91	9	0.041	0.060
Mn	1184	78	22	0.003	0.004
Fe	1515	100	0	0.087	0.125
Cu	1451	96	4	0.011	0.021
Zn	1473	97	3	0.008	0.011
Br	1164	77	23	0.002	0.002
Pb	1046	69	31	0.004	0.005
Jordan Valley Instrument					
Specie	**Cases (N =1147)**	**Selected (%)**	**Replaced (%)**	**Geometric Mean (µg/m^3)**	**Arithmetic Mean (µg/m^3)**
Al	419	37	63	0.261	0.917
Si	1036	90	10	0.501	0.757
Cl	976	85	15	0.059	0.074
K	1147	100	0	0.118	0.149
Ca	1089	95	5	0.050	0.087
Mn	1055	92	8	0.007	0.009
Fe	1144	99	1	0.179	0.244
Cu	1022	89	11	0.007	0.007
Zn	1146	99	1	0.018	0.017
Br	785	68	32	0.003	0.004
Pb	1047	91	9	0.005	0.006

Table A1. Summary Statistics for species used in Kevex and Jordan Valley PMF datasets.

Element	Kevex MDL	JV MDL	Element	Kevex MDL	JV MDL
Na	0.0056	.5440	**Se**	0.0008	.0002
Mg	0.0034	.0210	**Br**	0.0006	.0007
Al	0.0184	.0442	**Rb**	0.0008	.0004
Si	0.0084	.0275	**Sr**	0.0012	.0052
P	0.0028	.0121	**Y**	0.0012	n/a
S	0.0028	.0080	**Zr**	0.0012	.0037
Cl	0.0050	.0161	**Mo**	0.0016	n/a
K	0.0066	.0025	**Rh**	0.0270	n/a
Ca	0.0094	.0006	**Pd**	0.0230	n/a
Sc	0.0016	n/a	**Ag**	0.0210	n/a
Ti	0.0177	.0220	**Cd**	0.0230	.0043
V	0.0056	.0028	**Sn**	0.0320	n/a
Cr	0.0080	.0319	**Sb**	0.0330	.0010
Mn	0.0008	.0013	**Te**	0.0276	n/a
Fe	0.0008	.0069	**I**	0.0372	n/a
Co	0.0004	.0002	**Cs**	0.0512	n/a
Ni	0.0006	.0001	**Ba**	0.0542	.0248
Cu	0.0008	.0001	**La**	0.0740	n/a
Zn	0.0010	.0004	**W**	0.0036	n/a
Ga	0.0016	n/a	**Au**	0.0018	n/a
Ge	0.0012	n/a	**Hg**	0.0016	n/a
As	0.0008	.0001	**Pb**	0.0016	.0019

Table A2. Instrumental limits of detection ($\mu g/m^3$) for the Kevex and Jordan Valley XRF.

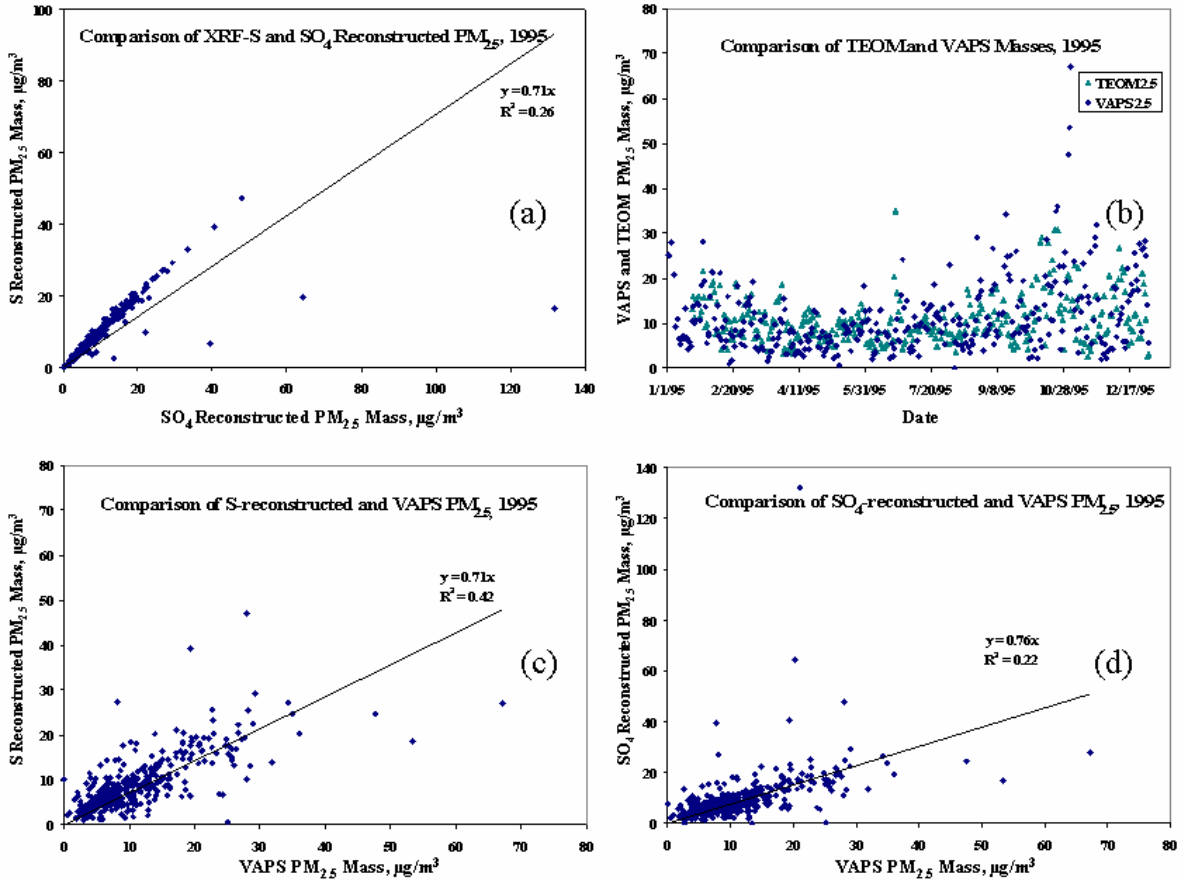


Figure A1. $PM_{2.5}$ Reconstructed Masses using XRF, ion, and total carbon masses. Figure (a) represents a comparison of the sulfur versus ammonium sulfate masses; (b) represents the comparison of VAPS and TEOM measured $PM_{2.5}$ masses; (c) represents the comparison of VAPS measured $PM_{2.5}$ mass to sulfur-derived reconstructed mass; and (d) represents the comparison of VAPS measured $PM_{2.5}$ mass to ammonium sulfate-derived reconstructed mass.

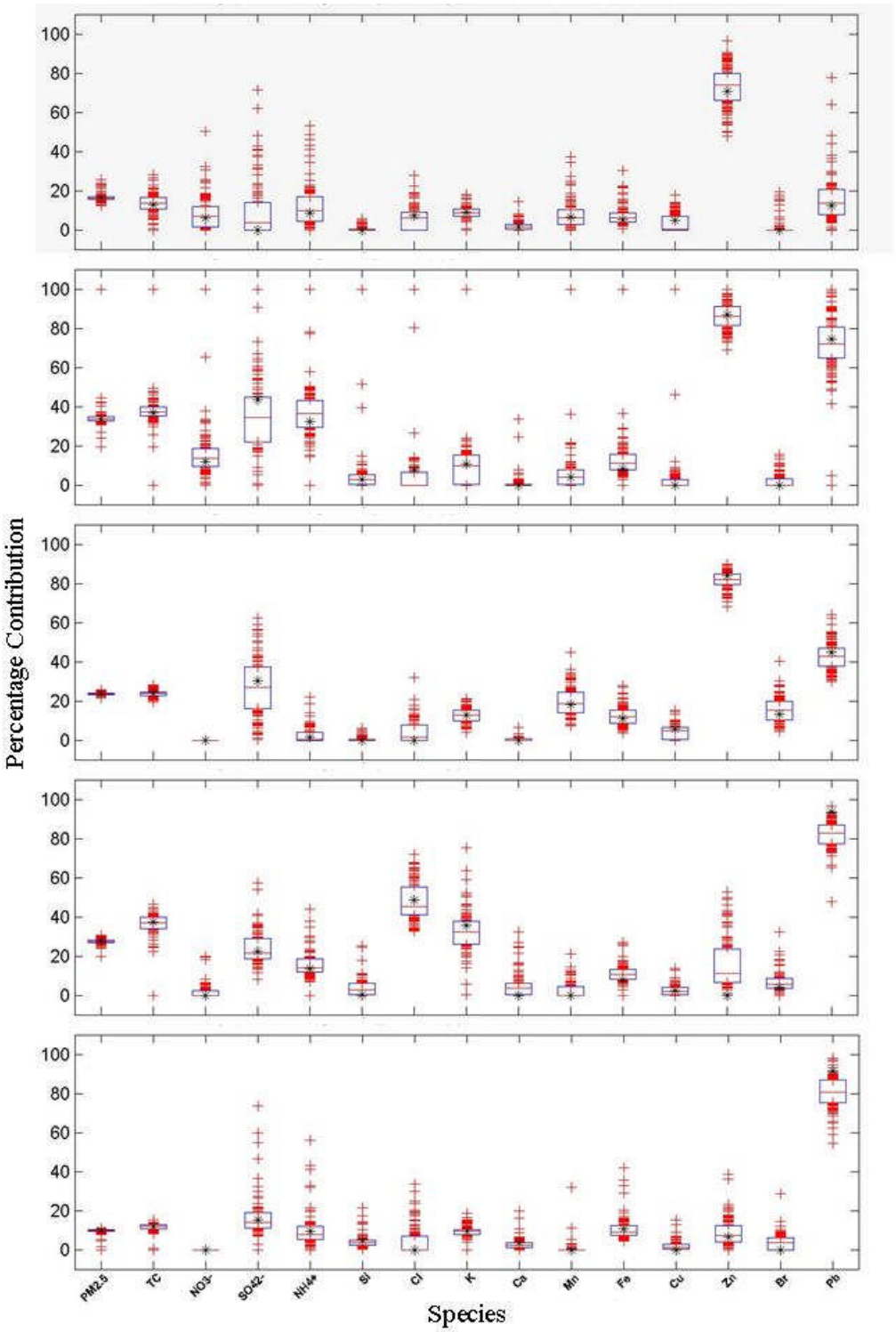


Figure A2. Vehicle Exhaust bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

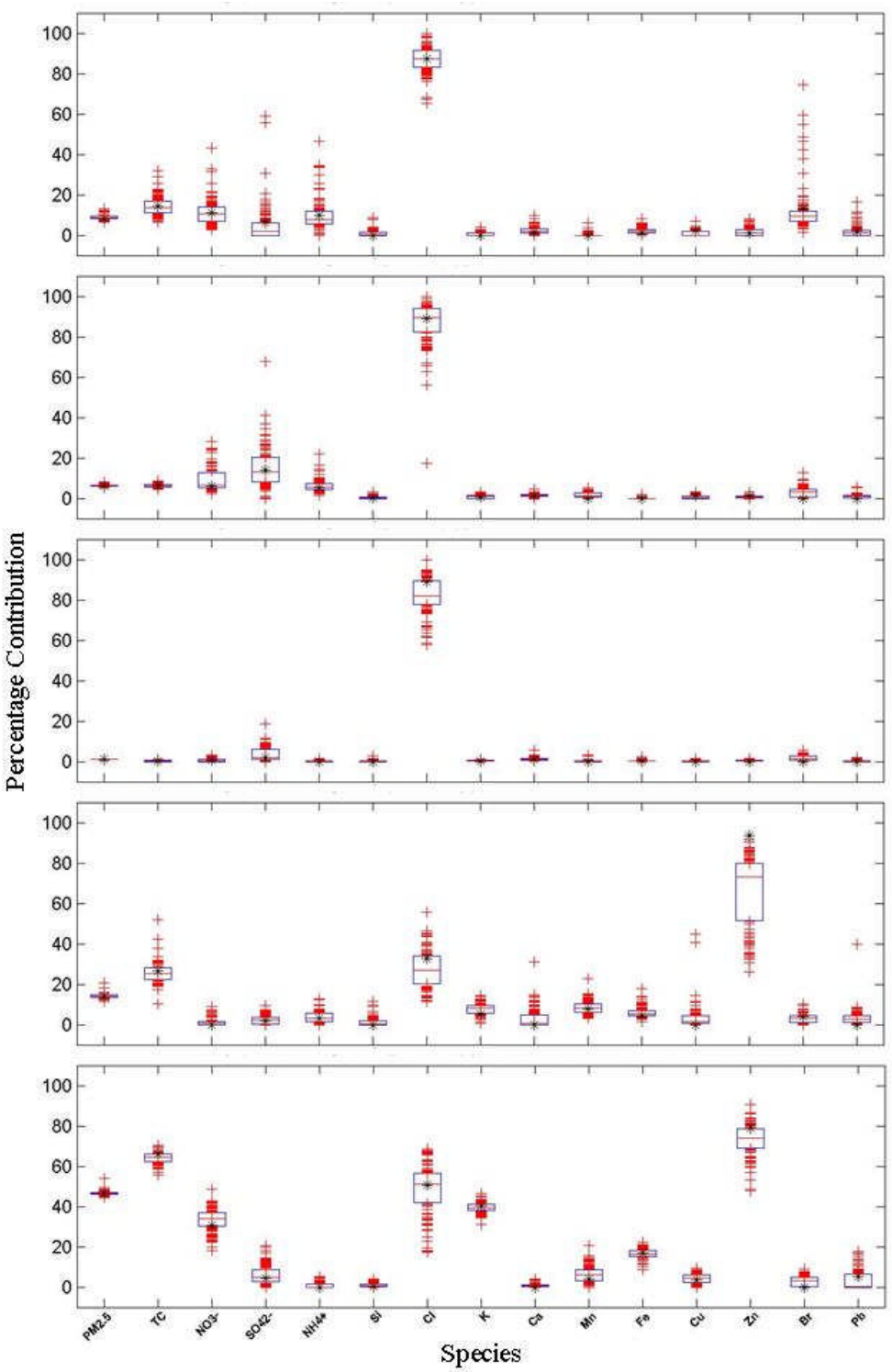


Figure A3. Chlorine-rich feature bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

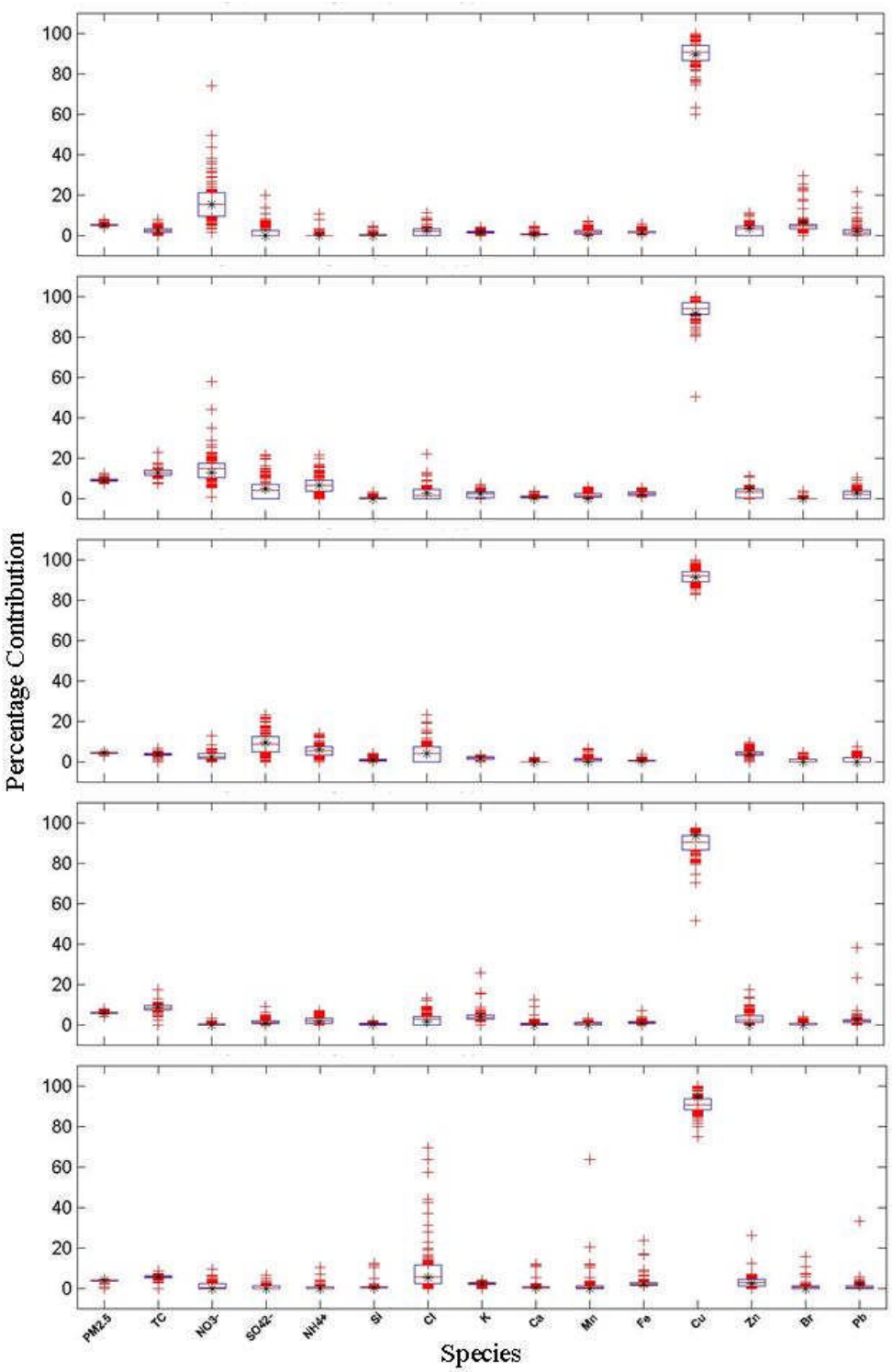


Figure A4. Metal processing bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

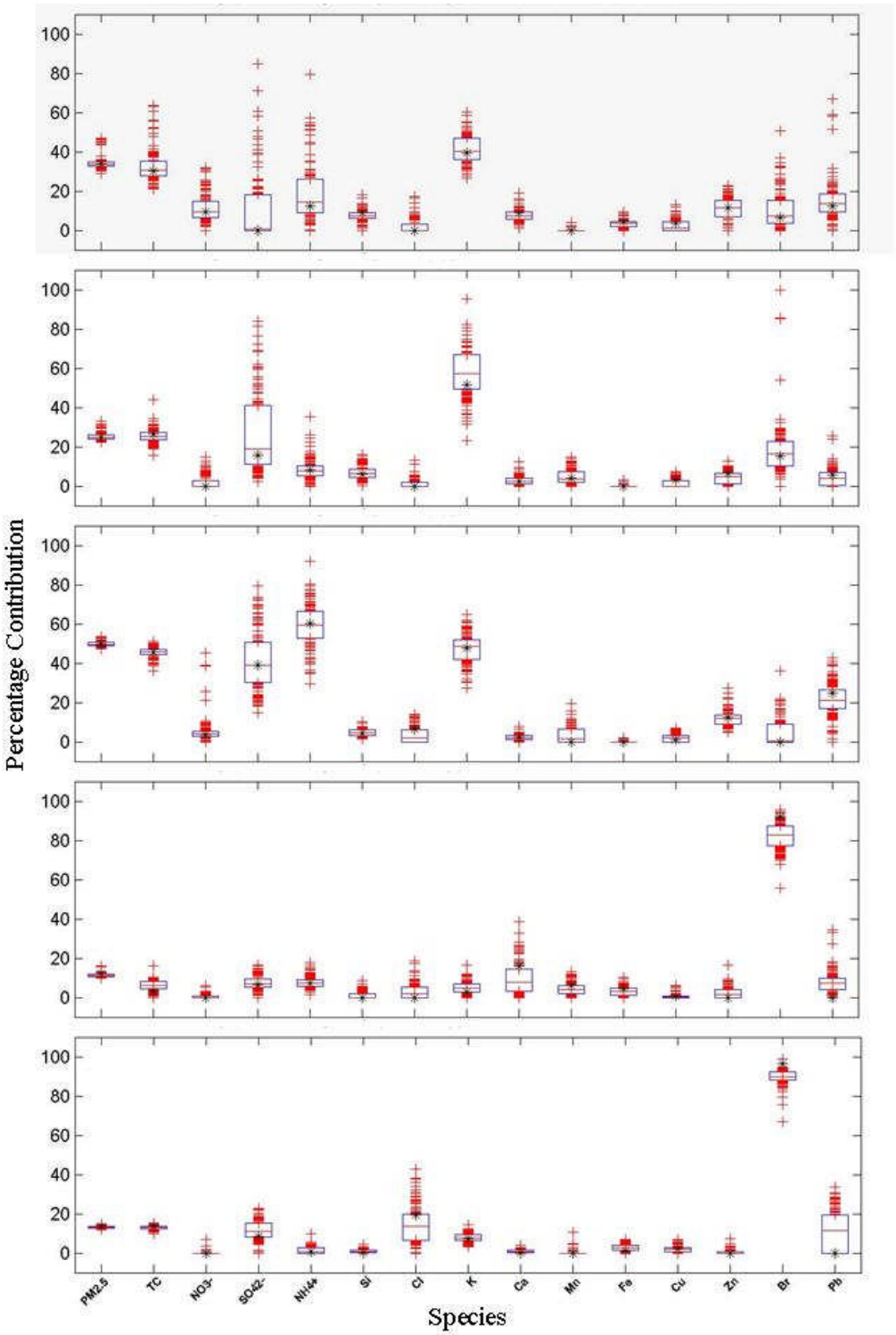


Figure A5. Biomass burning bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

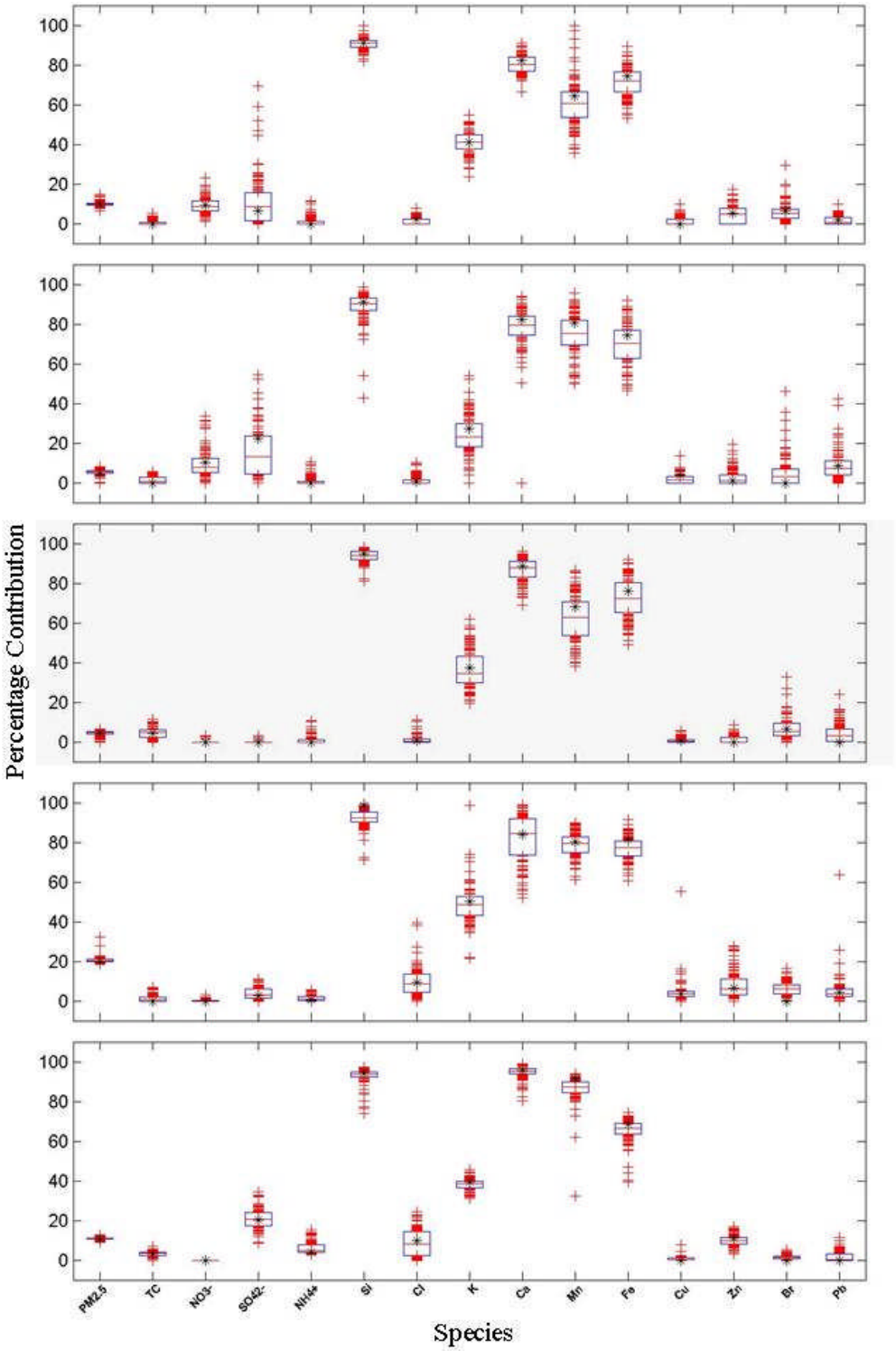


Figure A6. Airborne soil bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

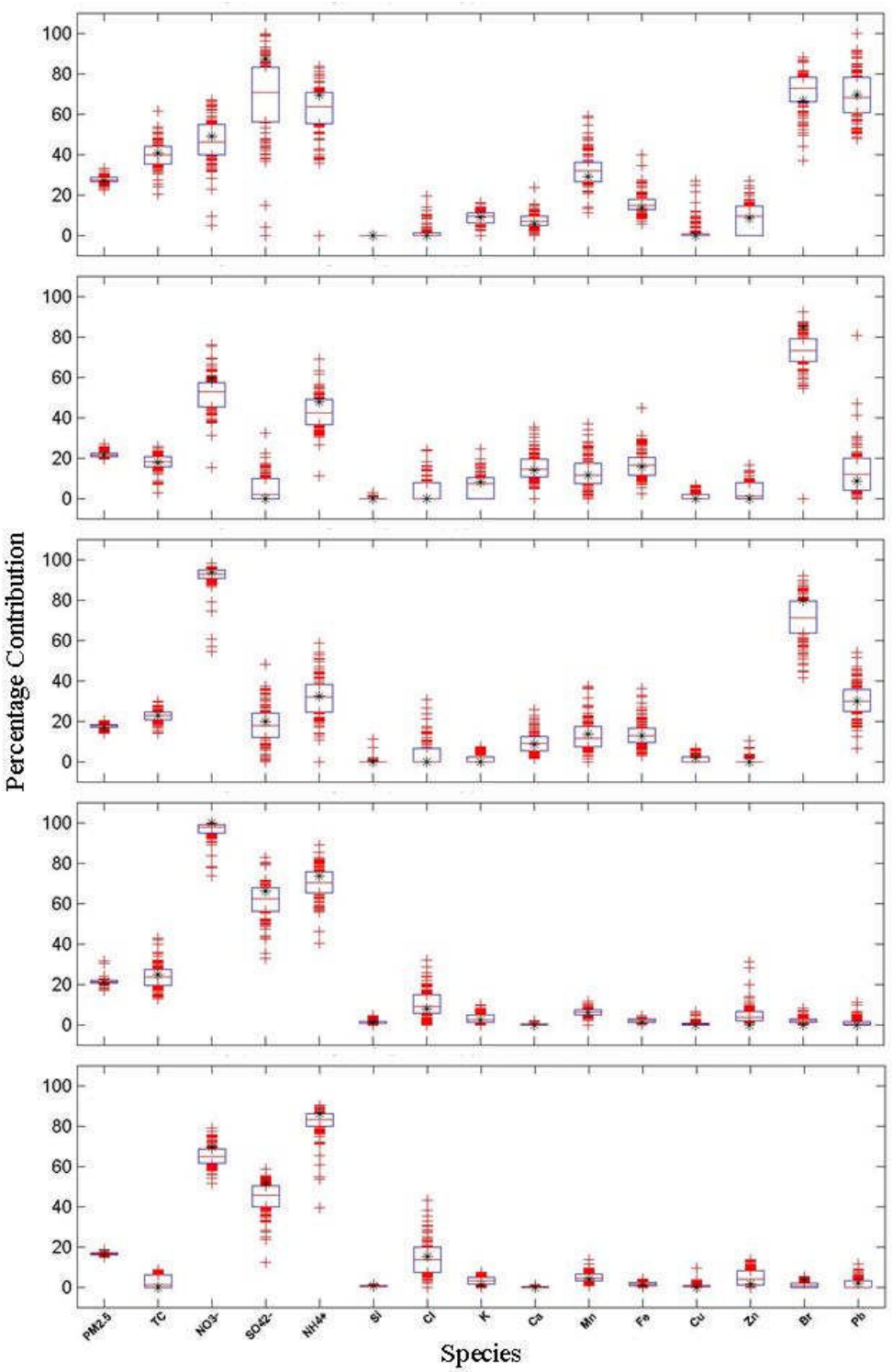


Figure A7. Nitrate bootstrap uncertainties for XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

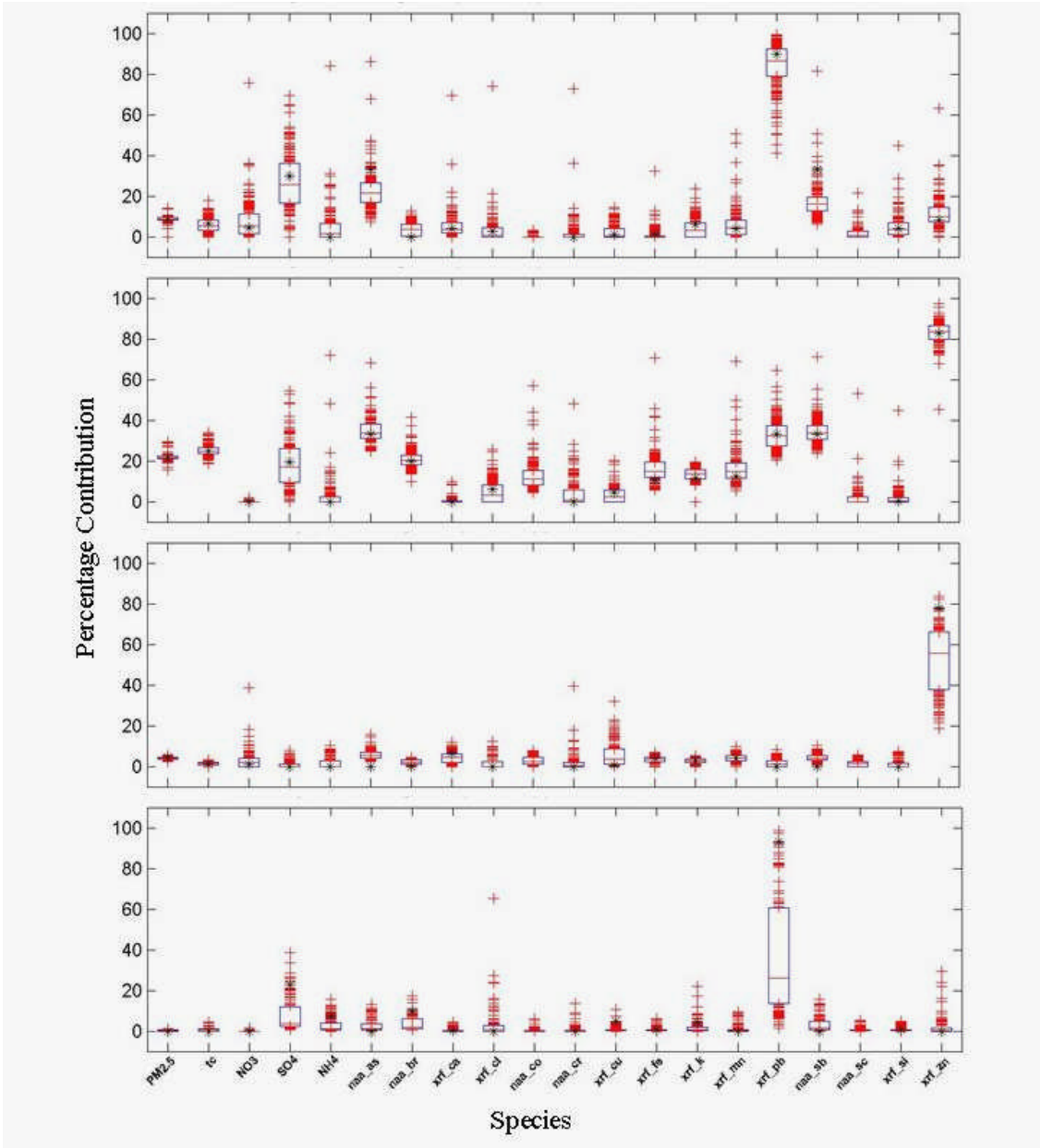


Figure A8. Vehicle exhaust bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

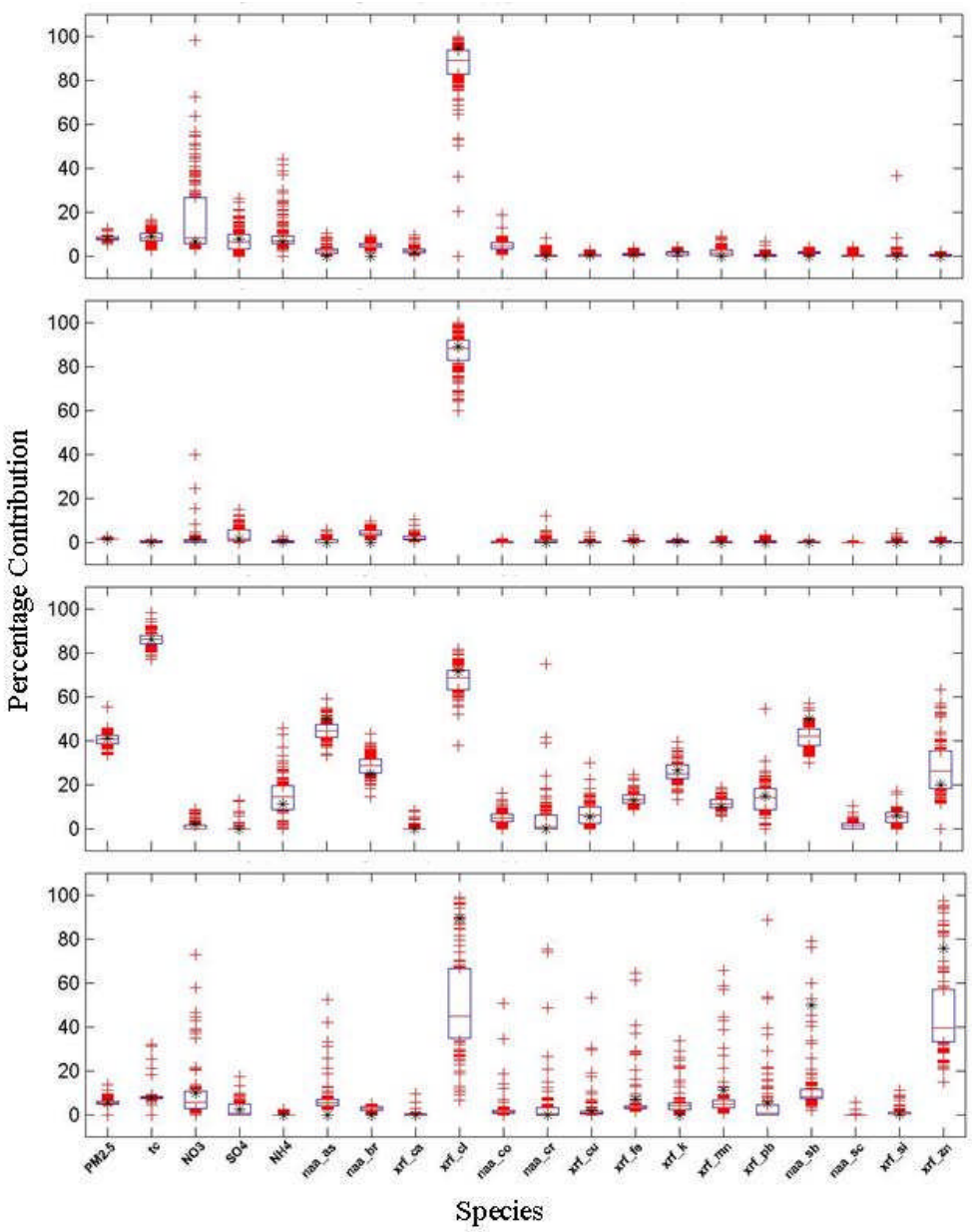


Figure A9. Chlorine-rich feature bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

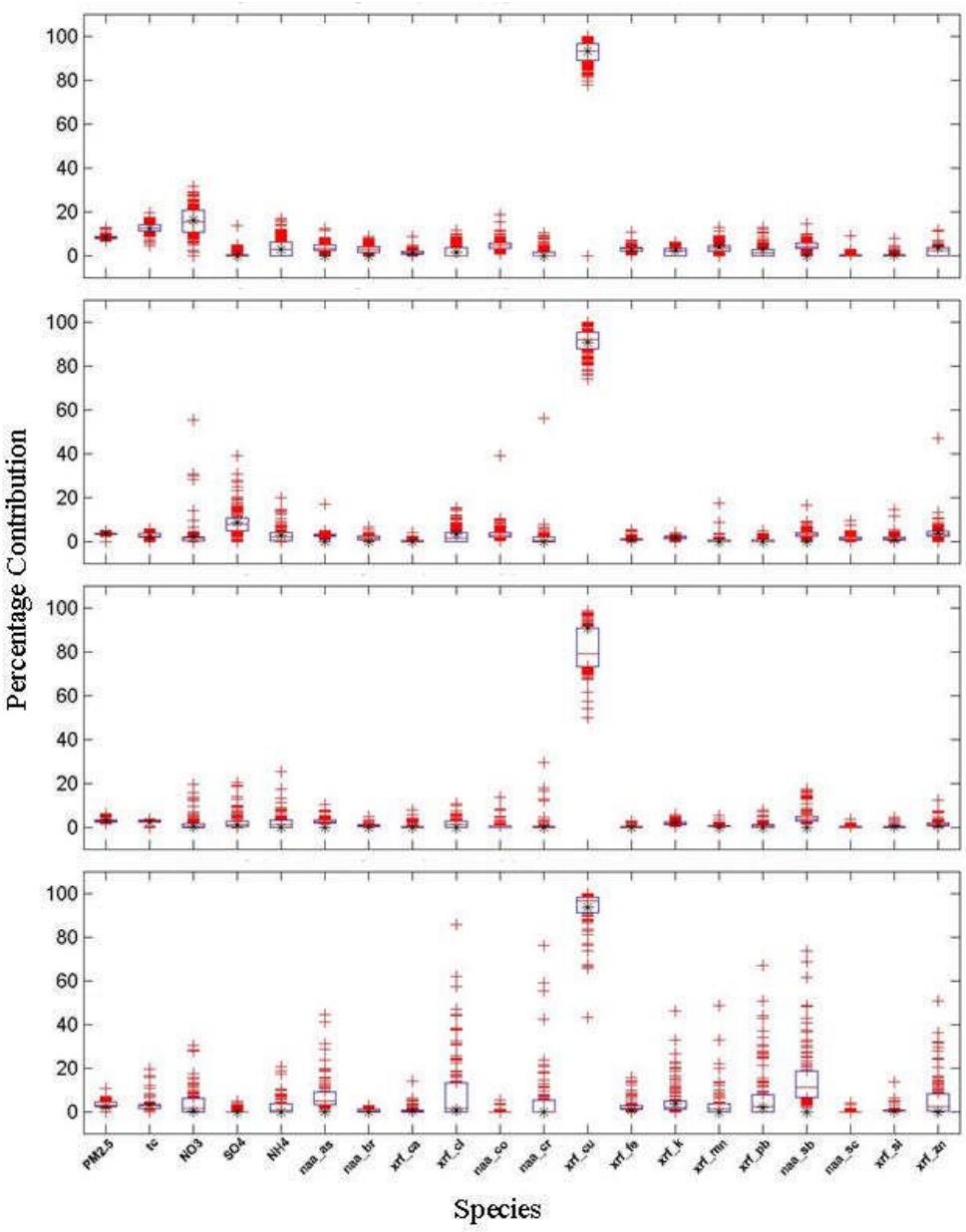


Figure A10. Metal processing bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

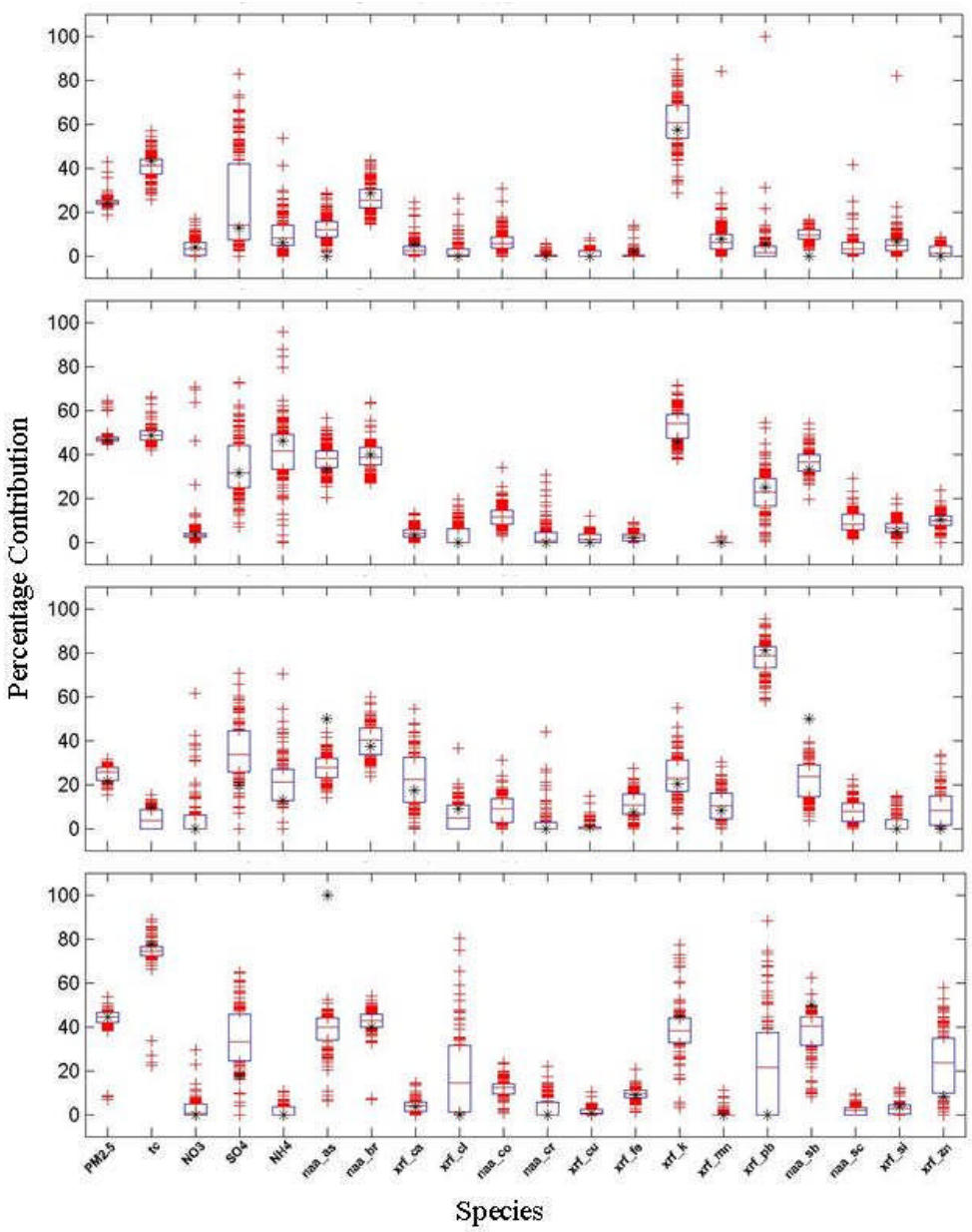


Figure A11. Biomass burning bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

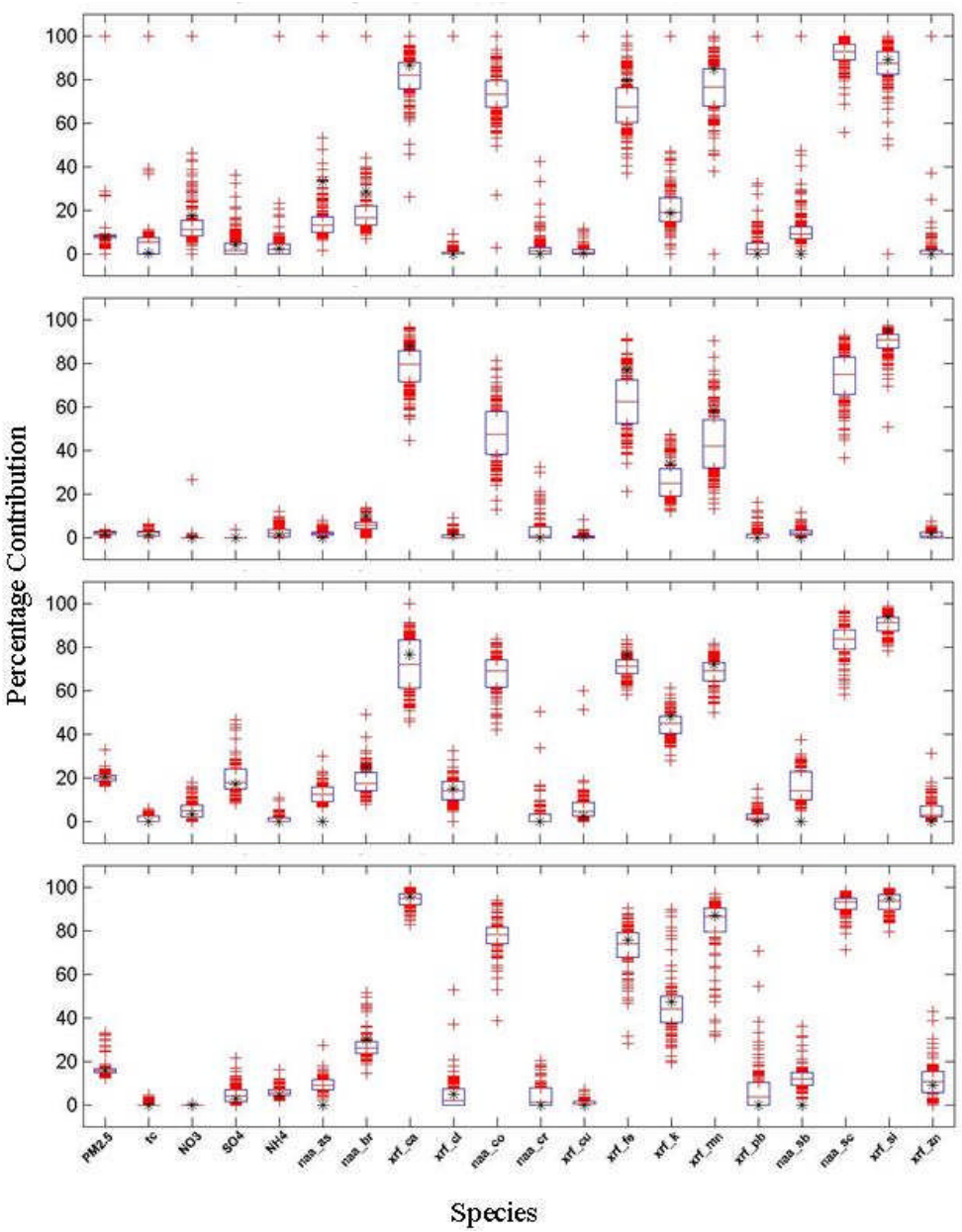


Figure A12. Airborne soil bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

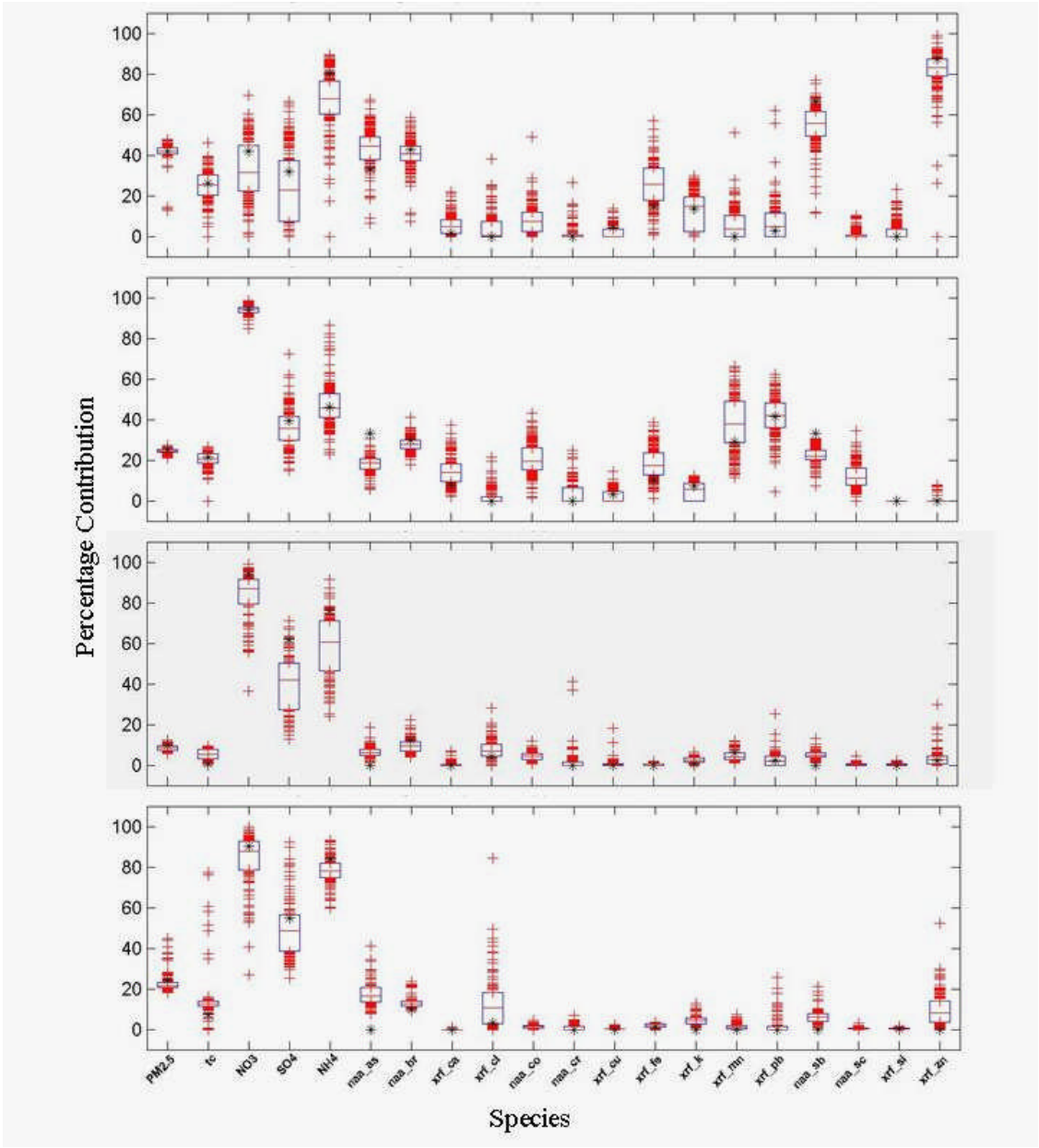


Figure A13. Nitrate bootstrap uncertainties for INAA-XRF data. Plots represent two Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

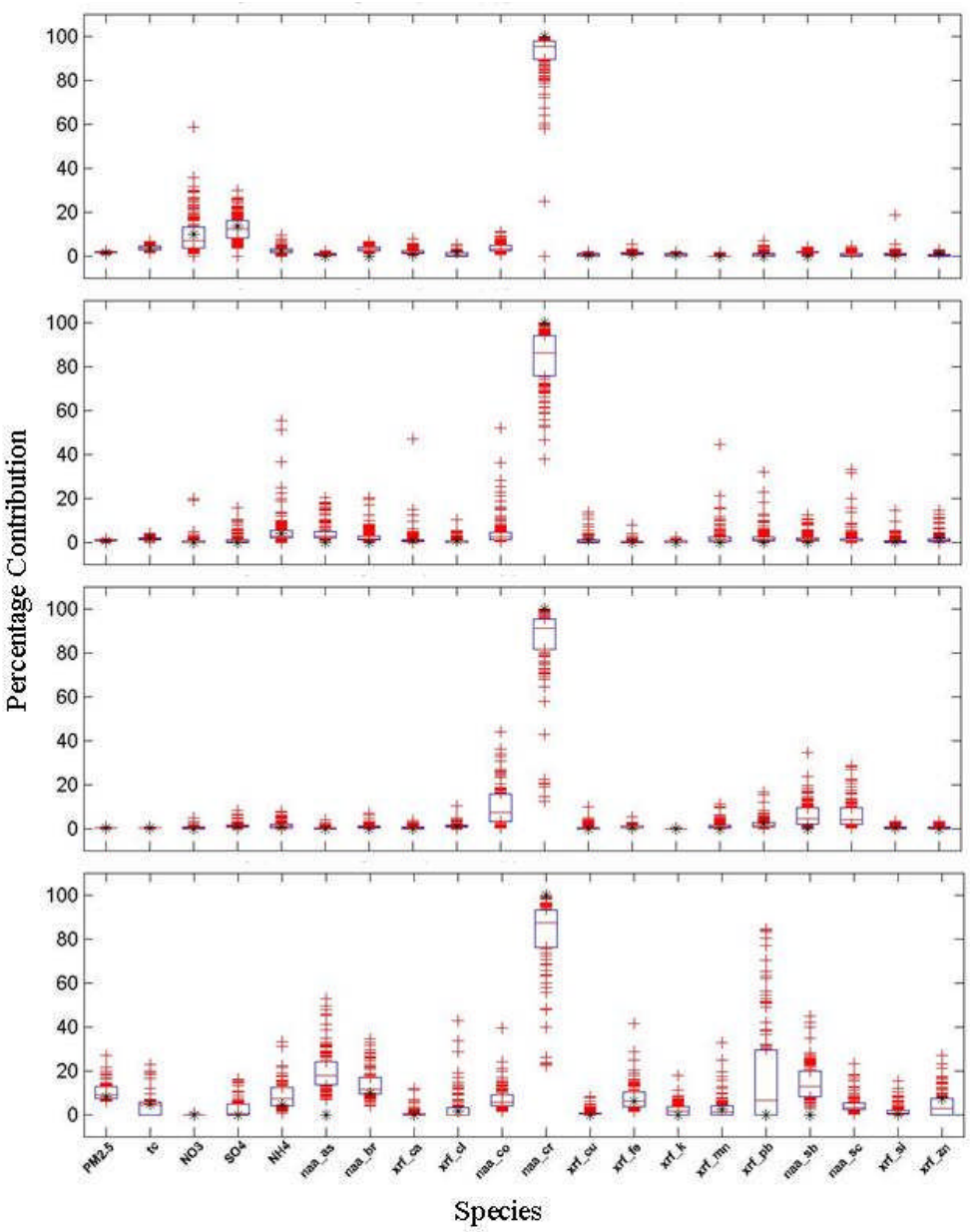


Figure A14. Chromium-rich feature bootstrap uncertainties for INAA-XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

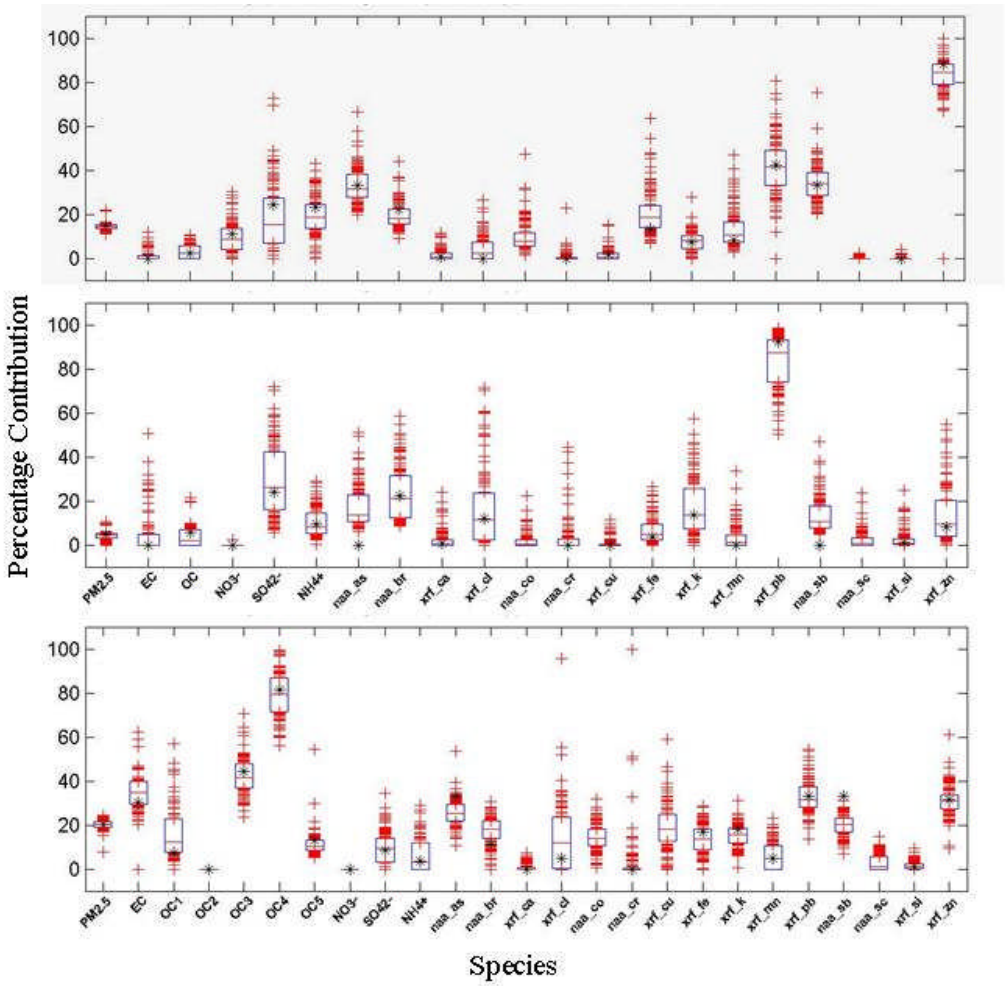


Figure A15. Vehicle exhaust bootstrap uncertainties for INAA-XRF-Carbon data. Plots represent INAA-Kevex-EC-OC, INAA-JV-EC-OC, and INAA-JV-EC-OC fractions, broken up in approximately five hundred points each.

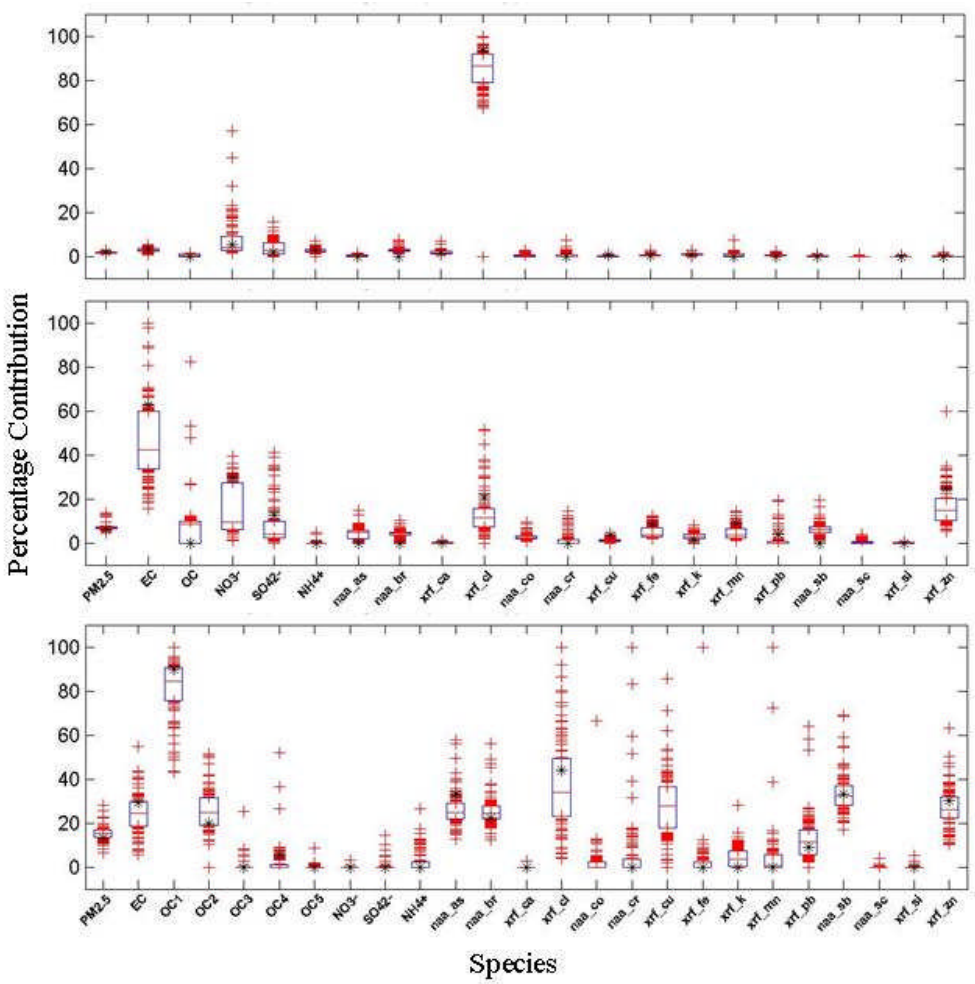


Figure A16. Cl-rich feature bootstrap uncertainties for INAA-XRF data. Plots represent INAA-Kevex-EC-OC, INAA-JV-EC-OC, and INAA-JV-EC-OC fractions, broken up in approximately five hundred points each.

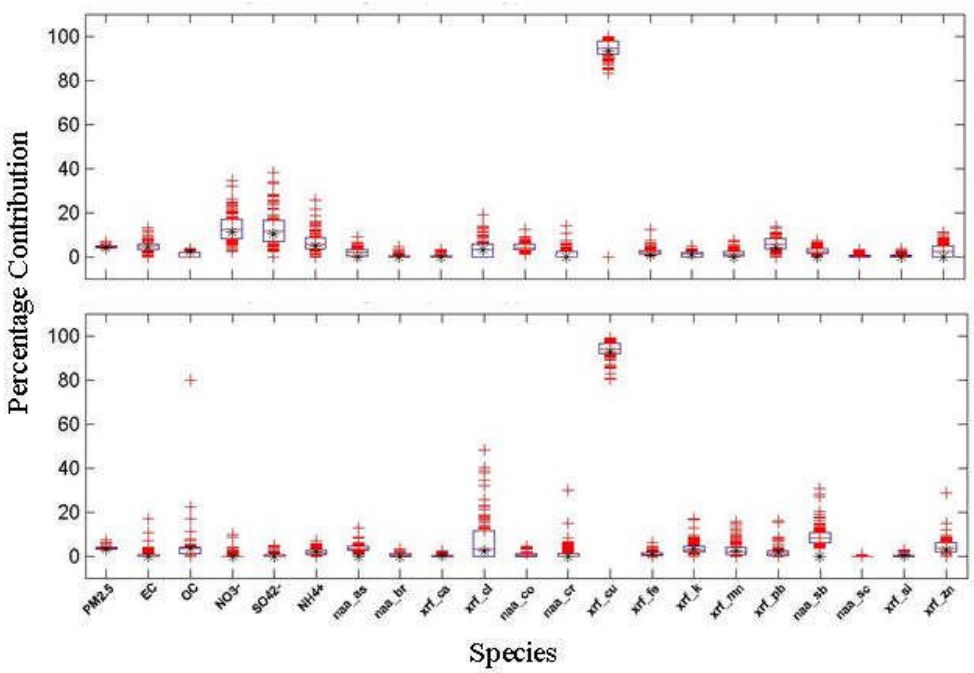


Figure A17. Metal processing bootstrap uncertainties for INAA-XRF data. Plots represent INAA-Kevex-EC-OC and INAA-JV-EC-OC, broken up in approximately five hundred points each. The OC fractions data did not produce a metal processing source.

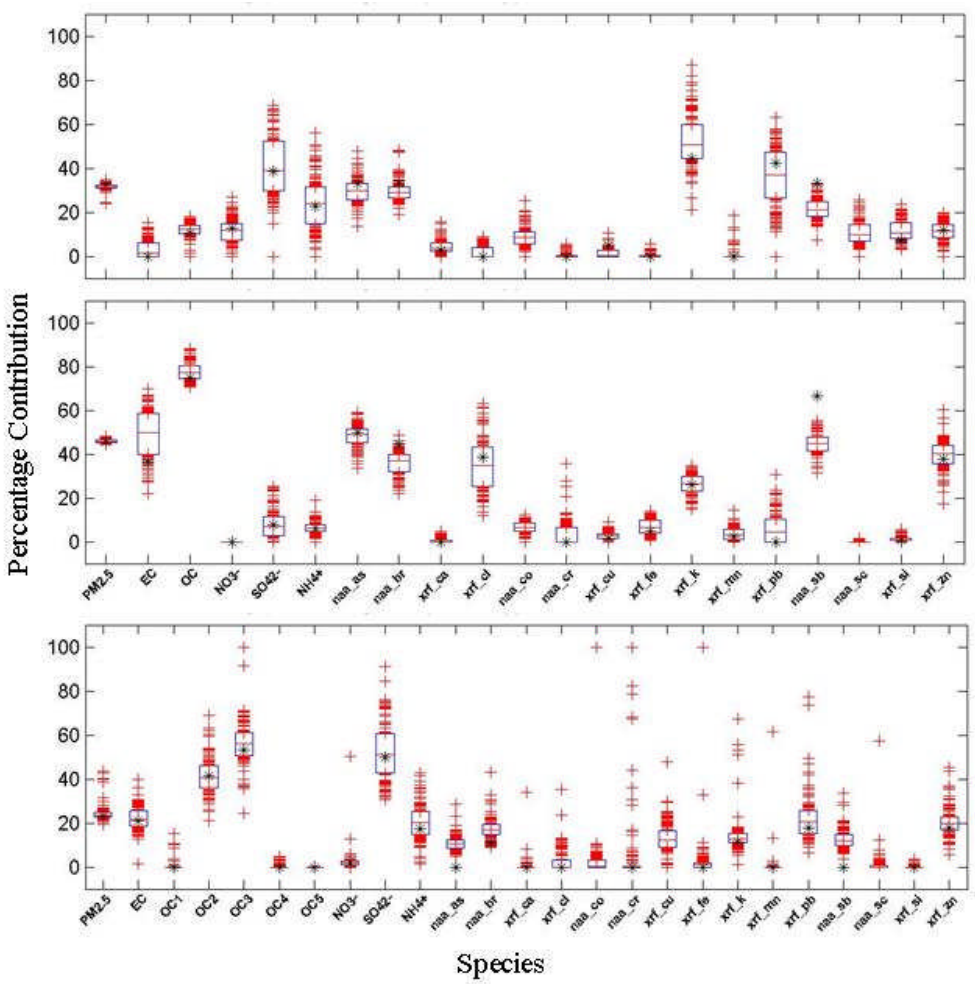


Figure A18. Biomass burning bootstrap uncertainties for INAA-XRF data. Plots represent INAA-Kevex-EC-OC, INAA-JV-EC-OC, and INAA-JV-EC-OC fractions, broken up in approximately five hundred points each.

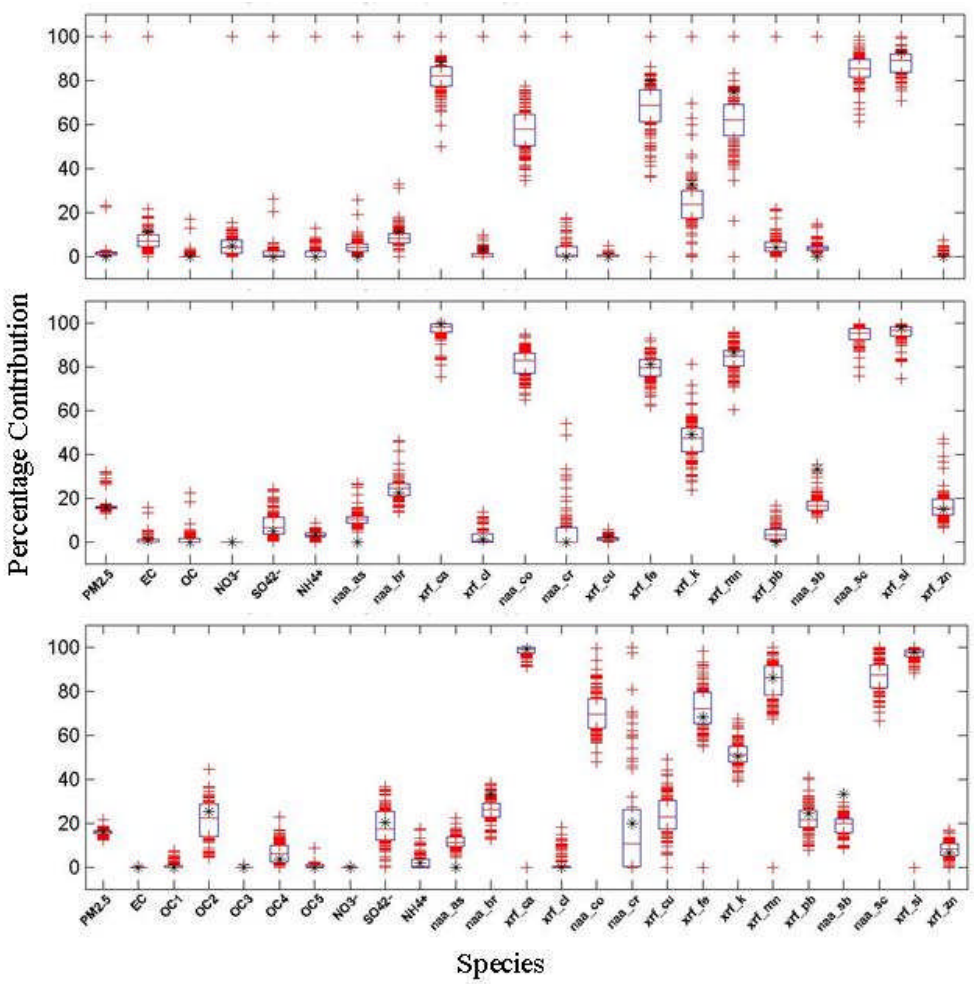


Figure A19. Airborne soil bootstrap uncertainties for INAA-XRF data. Plots represent INAA-Kevex-EC-OC, INAA-JV-EC-OC, and INAA-JV-EC-OC fractions, broken up in approximately five hundred points each.

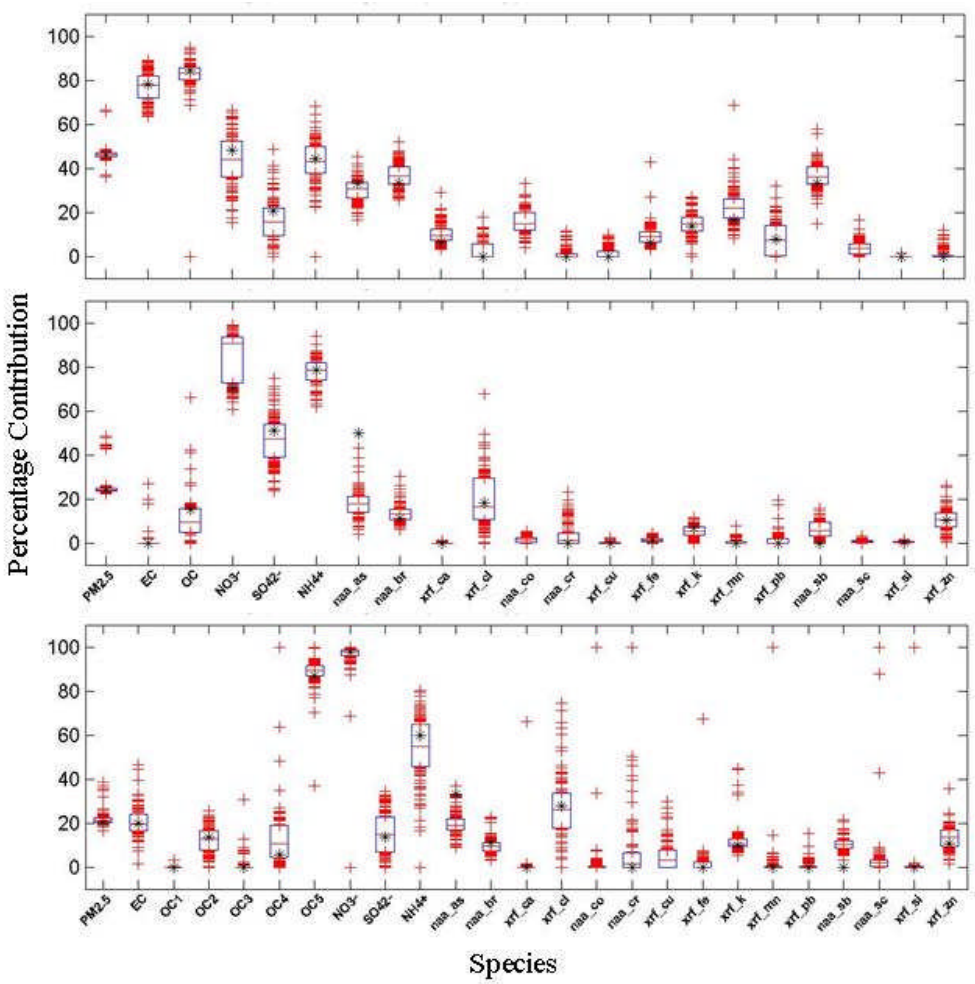


Figure A20. Nitrate feature bootstrap uncertainties for INAA-XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.

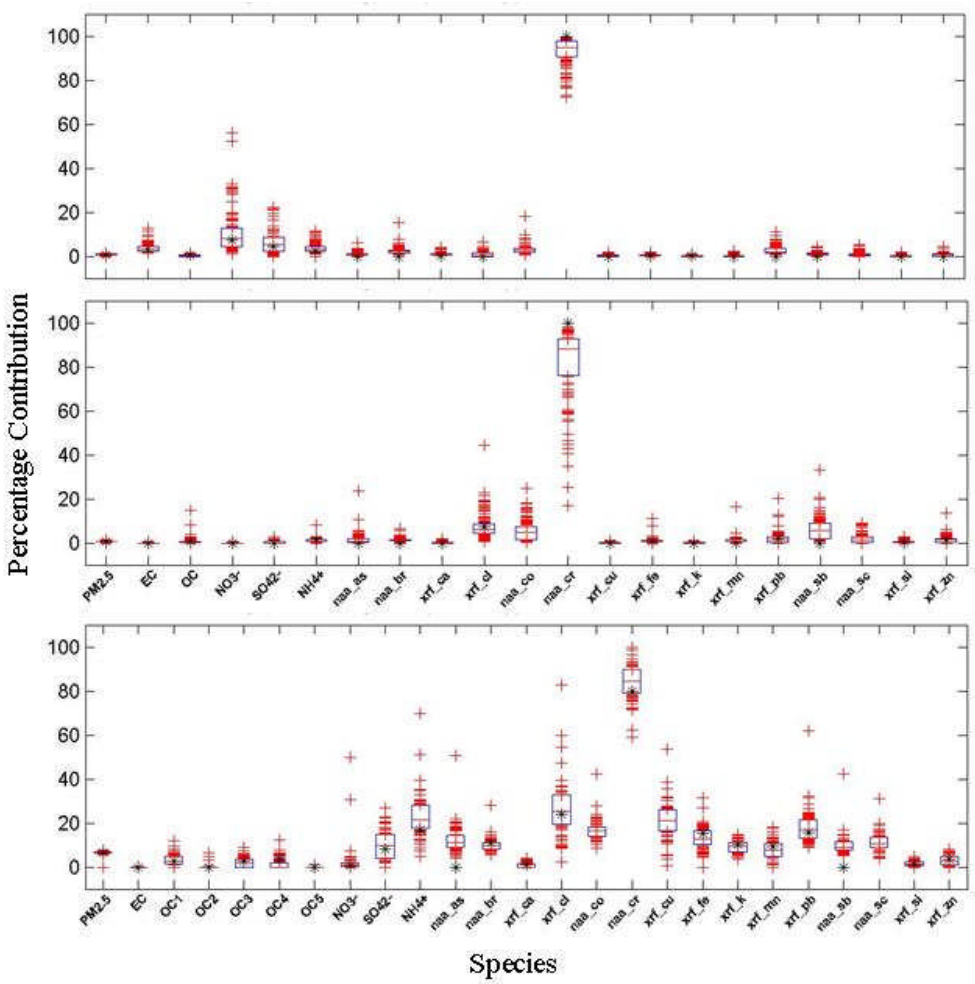


Figure A21. Chromium-rich feature bootstrap uncertainties for INAA-XRF data. Plots represent three Kevex and two Jordan Valley data groups, broken up in approximately five hundred points each.