2016

Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimkujik National Park, Nova Scotia

Yanyin Liao
University of Windsor

Recommended Citation
Liao, Yanyin, "Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimkujik National Park, Nova Scotia" (2016). *Electronic Theses and Dissertations.* Paper 5840.

Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimkujik National Park, Nova Scotia

By

Yanyin Liao

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of **Civil and Environmental Engineering**
in Partial Fulfillment of the Requirements for
the Degree of **Master of Applied Science**
at the University of Windsor

Windsor, Ontario, Canada

2016

Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimkujik National Park, Nova Scotia

by

Yanyin Liao

APPROVED BY:

Dr. D. Haffner
Great Lakes Institute for Environmental Research

Dr. R. Seth
Department of Civil and Environmental Engineering

Dr. X. Xu, Advisor
Department of Civil and Environmental Engineering

September 21, 2016

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The sources and processes affecting ambient speciated mercury concentrations including gaseous elemental mercury (GEM), gaseous oxidized mercury (GOM) and particulate bound mercury (PBM) at Kejimukujik National Park were identified using positive matrix factorization (PMF) model and principal component analysis (PCA). Four factors, Combustion Emission, Industrial Source, Photochemistry and Re-emission of Hg, and Sea Salt, were identified in both 2009 and 2010 by PMF model. The factors Photochemistry and Sea Salt were found to have the largest and smallest impact on ambient speciated mercury concentrations using PMF model, respectively. The components derived from PCA using the same dataset were largely consistent with the factors identified by PMF. A shift of factor impact on mercury concentrations between 2009 and 2010 was observed using both methods. An additional PCA component Gas-particle Partitioning of Hg was identified in 2009 according to the negative relation between GOM and PBM. After including meteorological parameters in the input of PCA, mercury wet deposition, a new factor, was identified in both years.

The reproduction of observed GEM concentrations by PMF model was the best among all three mercury forms followed by PBM and GOM. The sensitivity of PMF model to the different treatment to improve the data quality were tested. Imputations and combining or excluding GOM and PBM were found to have no obvious improvement on the model performances. However, increasing the low GOM and PBM concentrations by a scaling factor were effective in improving the model performances. Different treatments of input data had little impacts on factor profiles but factor contributions to Hg were affected to some extent.

DEDICATION

To my parents and girlfriend.

ACKNOWLEDGEMENTS

I would like to thank many people who helped me with the thesis. I would like to express my gratitude to my supervisor, Dr. Iris Xiaohong Xu, for her encouragement and support during my study in University of Windsor. I would like to thank Dr. Rajesh Seth, Dr. Douglas Haffner, and Ms. Irene Cheng, my committee members, for their valuable comments and suggestions on my thesis. I also sincerely appreciate the help from Dr. Leiming Zhang at Environment Canada.

I sincerely thank Xiaobin Wang and Tianchu Zhang for helping me learning about the methods used in this study. I am thankful to Environment Canada and Natural Sciences and Engineering Research Council of Canada for providing the funding of this project. I am also thankful to Environment Canada for the data used in this study. I acknowledge US EPA for the PMF model used in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Mercury is an odorless heavy metal with silver white color. Mercury is the only metal remaining liquid under standard state. Due to high vapor pressure, mercury is more volatile than other metals. The physical and chemical properties of mercury are listed in Table 1.1.

Table 1.1: Physical and chemical properties of mercury (Environment Canada, 2013a; Gaffney & Marley, 2014).

Parameters	Values
Atomic number	80
Atomic mass	200.59g/mole
Specific gravity	13.5
Melting point	-38.9℃
Boiling point	357.3℃
Valence states	0, +1, +2
Vapor pressure	0.261Pa (25℃)
Solubility (in water)	59μg/L (25℃)

Mercury air emission sources include natural sources, anthropogenic sources, and re-emission of mercury (United Nations Environmental Programme (UNEP), 2013). Natural sources include mercury volatilization from ocean, volcanic eruption, geothermal activities, and weathering of Hg-containing minerals. The atmospheric mercury released by the volatilization from ocean contributes most (70%) to the total emission of the natural sources (Gaffney & Marley, 2014). Artisanal and small-scale gold mining operations, mining, and smelting activities, and coal combustion processes are the three major anthropogenic mercury sources (UNEP, 2013). The re-emission of previously deposited mercury is also an important source. The Hg re-emission from ocean, soil, biomass burning, and lakes contribute 47%, 37%, 14%, and 2% to the total global re-emission, respectively (Pirrone et al., 2010).

In the atmosphere, mercury consists of three mercury forms which are operationally defined as gaseous elemental mercury (GEM), gaseous oxidized mercury (GOM), and particulate bound mercury (PBM). More than 83 and 92% of the gaseous atmospheric mercury exist as GEM in continental and marine air, respectively (Slemr et al., 1985). GEM can be oxidized to Hg_2^{2+} and Hg^{2+} by oxidizing agents such as O_3 and Br_2 (Clever et al., 1985). According to the standard potentials, only the oxidizing agents with an oxidizing potential between -0.80V and -0.85V can oxidize the elemental mercury to Hg_2^{2+} (Vanderzee & Swanson, 1974). However, these kind of agents do not exist naturally in our environment. When the oxidized mercury is in the gaseous state, they are named as GOM (e.g. HgO, $HgBr_2$ and $HgCl_2$). Approximately 3% of the total gaseous mercury are GOM (Lindberg & Stratton, 1998). When the oxidized mercury is associated with particulate matter, they are called PBM. Approximately 0.2-0.9% of atmospheric mercury are PBM (Slemr et al., 1985).

GEM could travel a long time in the atmosphere before deposition due to its low reactivity and low water solubility (Table 1.1) (Gaffney & Marley, 2014). The life time of GEM in atmosphere is approximately 1-2 years (Lindqvist and Rodhe, 1985). In contrast, GOM and PBM are removed faster in the atmosphere than GEM because GOM has higher water solubility and PBM has higher deposition velocity compared to GEM

(Gaffney & Marley, 2014). The lifetime of GOM and PBM is from several days to a few weeks (Lindqvist and Rodhe, 1985). After mercury's dry and wet deposition to soil, water, and vegetation surfaces, some of GOM and PBM are reduced to elemental mercury and then re-emitted to the atmosphere (UNEP, 2008). Some of the deposited mercury were transformed to organic mercury.

All forms of mercury are harmful, especially for the organic mercury. Exposures to mercury will cause central nervous system problems and kidney problems in human (Environment Canada (EC), 2013a). People could die by contacting several drops of dimethylmercury (Me_2Hg) (Nierenberg et al., 1998). The main pathway of mercury intake in humans is the consumption of methylmercury (MeHg) contaminated food, especially fish, due to the bioaccumulation of MeHg (EC, 2013b). The bioaccumulation factor for piscivorous fish is 10^7 (EC, 2013b). In other words, the MeHg concentration in a piscivorous fish is 10 million times of the concentration in water.

The bioaccumulation and the health concern of Hg raised the public attention to Hg pollution. A better understanding of the sources/processes affecting mercury concentrations will help to control the risks of mercury pollution. Receptor models are usually used to identify the processes or sources affecting the ambient mercury concentrations at sampling sites. Positive matrix factorization (PMF) model (US EPA, 2014a) and principal component analysis (PCA) (Thurston and Spengler, 1985a) are two common methods used in Hg source apportionment studies.

To date, only three studies about speciated mercury source apportionment using PMF model were conducted (Liu et al., 2003; Cheng et al., 2009; Wang et al., 2013). Only one study (Cheng et al., 2009) used both PCA and PMF model in mercury source apportionment. However, a thorough comparison between PMF and PCA results was not conducted in that study (Cheng et al., 2009). It is necessary to compare PMF and PCA results because it helps to verify the major factors affecting ambient mercury concentrations at the receptor site.

The overall objective of this study is to identify the factors affecting ambient mercury concentrations at receptor site using both PMF and PCA. The specific objectives are to:

- Identify and compare the factors affecting ambient mercury concentrations using PCA and PMF model;
- Summarize the similarities and differences of the factors and components between 2009 and 2010;
- Summarize the similarities and the differences between the results derived from PCA and PMF model;
- Evaluate the PMF model performances and determine which mercury forms are better reproduced and in which year the speciated mercury are better reproduced;
- Analyze the impacts of different treatments of missing values and the low concentrations of speciated mercury on PMF results;

- Analyze the impacts of including the meteorological parameters on PCA results.

CHAPTER 2

LITERATURE REVIEW

2.1 Methylation of mercury

Methylmercury (MeHg) and dimethylmercury (Me_2Hg) are two main organic Hg existing naturally (Ullrich et al., 2001). MeHg is mainly found in freshwater while Me_2Hg is more likely to exist in deep ocean water (Ullrich et al., 2001). They are formed by the activities of microorganisms in the aquatic environment after the wet or dry deposition of inorganic Hg (Compeau & Bartha, 1985). Because MeHg is not readily eliminated from the organisms, MeHg accumulates in the food chain from bacteria to plankton, through herbivorous fish to piscivorous fish (Wiener et al., 2002). Me_2Hg is not likely to accumulate in the food chain because it is readily removed from water by evaporation due to its high volatility and poor solubility (Talmi & Mesmer, 1975; Morel et al., 1998).

Mercury methylation rates are influenced by a lot of factors such as biological activity, nutrient availability, pH, temperature, redox potential, and the presence of inorganic and organic complexing agents (Ullrich et al., 2001). It was found that the rise of the temperature could accelerate the evaporation of elemental mercury as well as the methylation rates of the mercury (Pack et al., 2014). Therefore, global warming could increase the risk of mercury pollution.

2.2 Receptor Models

Although many factors could affect the methylation of mercury as mentioned in section 2.1, the amount of mercury available for methylation is the most important factor. The amount of the mercury available for methylation is affected by the sources emitting the mercury. In order to control the risk of the mercury pollution due to its health concern, receptor models could be used to identify the sources and processes affecting the ambient mercury concentrations. Receptor models apportion the observed ambient pollutants concentrations at a receptor site to different factors (Henry et al., 1984). The factors are assigned to specific sources based on the knowledge about the sources and the receptor site. The fundamental of receptor model is chemical mass balance equation (Equation 1, Watson et al., 2008):

$$C_{it} = \sum_{j=1}^{k} F_{ij} S_{jt} + E_{it} \qquad (1)$$

where:

k = total number of factors;

C_{it} = concentration of the i^{th} species in t^{th} sample;

F_{ij} = mass fraction of the i^{th} species in j^{th} factor;

S_{jt} = contribution of the j^{th} factor to the t^{th} sample;

E_{it} = difference of the calculated concentration and observed concentration of i^{th} species in t^{th} sample.

The outputs of different receptor models were different. The outputs of the following three methods are discussed here: Chemical Mass Balance (CMB) model (Thurston and Spengler, 1985), PMF model (US EPA 2014), and PCA (US EPA, 2004). Factor contribution is the only output for CMB model (US EPA 2004) while both factor profiles and factor contributions are provided by PMF model (US EPA 2014). PCA only provides qualitative factor components as outputs. Detailed descriptions about the advantages and disadvantages of CMB, PMF and PCA could be found in Section 3.3.

2.2.1 PMF model

PMF solves the mass balance equation (Eqn 1) by decomposing the matrix of sample data (C_{it} in Eqn 1) into the matrix of factor contributions (S_{jt} in Eqn 1) and the matrix of factor profiles (F_{ij} in Eqn 1) with the constraint that both factor contributions and factor profiles are required to contain no negative values. The goal of PMF in air quality study is to get the factor contributions, factor profiles, and the number of factors which have the most reasonable physical meaning (Paatero et al., 2014). After the number of factors is selected, the reasonable factor contributions and factor profiles could be found when the objective function Q (Eqn 2) reaches the minimum (US EPA, 2014a).

$$Q = \sum_{i=1}^{m}\sum_{t=1}^{n}\frac{C_{it} - \sum_{j=1}^{k} F_{ij}S_{jt}}{u_{it}} \quad (2)$$

where:

m= total number of the included species,

n= total number of the included samples,

u_{it}= uncertainty of the ith species in the tth sample

Ambient concentration data and their uncertainties are two input files required by PMF model. According to user guide from US EPA (2014), observation-based and equation-based uncertainties are acceptable by PMF model. The observation-based uncertainties reflect the errors caused by sampling and measuring process. They are usually provided with the ambient concentrations by the laboratory or the reporting agency. The equation-based uncertainties are usually calculated using concentrations, error fractions and sometimes method detection limits (MDL). Equation-based uncertainties allow the users to have modicum control over the deviation in PMF solutions (Reff et al., 2007). Equation (3) is recommended by US EPA (2014) to calculate the equation-based uncertainties. The error fraction in equation (3) is assumed by the user according to the measurement and knowledge about species calculated.

$$\begin{gathered}\text{Uncertaity} = \frac{5}{6} \times MDL, when\ concentration \leq\ (MDL) \\ \text{Uncertainty} = \sqrt{(Error\ Fraction \times concentration)^2 + (0.5 \times MDL)^2}, \\ when\ concentration > MDL\end{gathered} \tag{3}$$

Before running the PMF model, the chemical species need to be categorized as “strong”, “weak”, or “bad” based on the knowledge about sources, sampling process, and analytical uncertainties. “Weak” species are downweighted by tripling the provided uncertainties while “bad” species are removed from the calculation (US EPA, 2014a).

Markers are the species used to indicate specific pollution sources (Hastings & Gross, 2012). For example, SO_2 is the marker of coal combustion process (Liu et al., 2007). When SO_2 is not included in the input files, it is hard to identify the coal combustion process. Therefore, marker species should not be categorized as “bad”. According to PMF user guide (US EPA, 2014a), when the information of the dataset is not enough to decide the category of the species, the signal to noise (S/N) ratio could be used to do the categorization. The S/N ratio indicates “whether the variability in the measurements is real or within the noise of the data” (US EPA, 2014a). The species with an S/N ratio less than 0.5 should be categorized as “bad” and the species with an S/N ratio between 0.5 and 1.0 should be categorized as “weak”. Users could specify a total variable which would be used in the post-processing of the results such as the percentage of the total mass in each of the factors. The total variable is an artificial variable which is the total mass of the same type of species. For example, when input of the PMF model are $PM_{2.5}$ components, $PM_{2.5}$ mass should be calculated and selected as the total variable (US EPA, 2014a). It should be noted that the total variable is usually categorized as “weak” because the total variable should not have a large influence on the results.

Other tools such as concentration scatter plot and concentration time series provided by PMF model could help to analyze the input data before running the model. The concentration scatter plot shows the relationship between two user-specified species. A correlation between two species indicates they are emitted from a similar type of source or a nearby location (US EPA, 2014a). The concentration time series could be used to examine whether there are unusual events during the sampling period. The samples affected by the unusual events need to be excluded from the dataset (US EPA, 2014a).

Number of runs, number of factors, and seed number are the three parameters to be determined before a base model run. The number of runs is recommended to be 20 because this allows an evaluation of the results stability (US EPA, 2014a). The seed number is the start point for each iteration. Either a random start or a fixed start point is accepted by PMF model. PMF model is able to reproduce the base run results by using the same seed number (US EPA, 2014a).

The number of factors needs to be chosen according to the knowledge about the dataset. When the background information is not enough to determine the number of the factors, several methods could be used to determine the range of the number of the factors. The maximum individual column mean (IM) and the maximum individual column standard deviation (IS) of the scaled residual matrix could be used to determine the range of the number of factors (Lee et al., 1999). When the number of the factors increases to a critical value, IM and IS will experience a drastic drop. The optimal number of factors should be no less than the critical value. The change of the Q values (dQ) also provides useful information on deciding the number of factors (Hopke, 2000; Viana et al. 2008a; Brown et al., 2015). When dQ becomes small as the number of factor increases, there might be too many factors (Hopke, 2000; Brown et al., 2015). Runs with different numbers of factors in the range determined by IM, IS, and dQ need to be conducted. The interpretability of each result should be checked. The final solution should be a compromise of the indexes and the interpretability (Cesari et al., 2016).

Factor profiles and contributions are the outputs of PMF model. Three kinds of factor profiles are provided by PMF model: (1) concentrations of each compound in each factor, (2) percentage of a compound's total mass for each compound in each factor, and (3) percentage of total mass within the factor for each compound in each factor. In factor profiles (3), the sum of a compound's concentrations in each factor should equal to the predicted concentration of that compound. On the factor profile screen, there are two graphs about the factor profiles and factor contributions. One shows the bar plot of factor profiles and the other one shows the time series of the factor contributions. The mass fraction of each species contributing to each of the factors is showed as stacked bar plot in Factor Fingerprint Screen. This plot could be used to show the species distributions (US EPA, 2014a). The Factor Contribution Screen displays two graphs, a pie chart showing the contributions of each factor to each species and the time series of factor contribution of all factors to each sample.

Among three factor profiles, the factor profiles (3) (page 12) is similar to the source profiles while the other two factor profiles are different from the source profiles. Source profiles are the mass fractions of each species in a source (US EPA, 2004). For example, the iron sinter emission consists of 14.8% acetylene, 3% ethane, 5.9% ethylene, 73.3%

methane, and 3% propylene (US EPA, 2014b). Therefore, the factors could be interpreted based on the comparison between the factor profiles (3) and the source profiles from the literature or database (Leuchner et al., 2015). However, it is hard to do a quantitative comparison between the factor profiles (3) and the source profiles directly because the composition of the sources changed during the transport from the sources to the receptors due to mixing and chemistry (Leuchner et al., 2015). In addition, the source profiles in the literature and database may be outdated because the control technologies and strategies are upgraded to meet the emission limits (Reff et al., 2009). Therefore, it is difficult to interpret the factors based on a direct quantitative comparison between source profiles and factor profiles. A qualitative comparison between the major variables (i.e. variables with a mass fraction of the total species mass larger than a user-specified value) of each factor in factor profiles (2) and the markers of the sources from the outcomes of other studies could be conducted to identify the possible source.

Model performances are another output of the PMF model. Q(robust), Q(true), and convergence of the results are three basic performance measurements among the performance indexes. Q(true) is a goodness-of-fit parameter calculated including all point while Q(robust) is calculated excluding the samples with an uncertainty-scaled residual greater than 4 (US EPA, 2014a). When the Q(robust) values in several runs vary in a large range, the stability of the result is poor (US EPA, 2014a). When differences between Q(true) and Q(robust) values are small in each of the 20 runs, the uncertainties of the input data might be too high (US EPA, 2014a). An in-convergence result indicates that the minimum Q value is not found (US EPA, 2014a). Too-low uncertainties, incorrect uncertainties or inappropriate input parameters could lead to an in-convergence result (US EPA, 2014a). These uncertainties need to be checked and input parameters need to be adjusted before the model is rerun. The converged result with the lowest Q(robust) value is highlighted and used for further analysis.

Other performances indexes such as uncertainty scaled residuals, Obs/Pred scatter plot and Obs/Pred time series could be used to evaluate the model ability to reproduce the observations. The uncertainty scaled residuals are calculated by dividing the residuals by error estimates of the observed data. When the scaled residual distribution is skewed or bimodal, the observed concentrations are poorly reproduced by the model (Polissar et al., 1998). When all the scaled residuals are between +3 and -3 and the stacked bar plot (histogram) of the scaled residuals is normally distributed, the observed concentrations are reproduced well (US EPA, 2014a). The Obs/Pred Scatter Plot screen contains a Base Run Statistics table and an Obs/Pred Scatter Plot for each species. The Obs/Pred Scatter Plot contains a one-to-one line and a regression line of the observed and predicted concentrations for each species. When there is no large bias between these two lines and the determination coefficients (R^2) of the regression line is close to one, the model reproduces the observed concentrations well. When the R^2 value is low, the model performance on reproducing the observations is poor no matter how close to 1 the slope of regression line is. The Observed/predicted Time Series is a line chart of the observed and predicted concentrations versus time. When the observed concentrations and the predicted concentrations do not track each other, the reproduction of the observed concentrations is poor (US EPA, 2014a). G-space plot is a scatter plot of the factor contributions for two user-selected factors. When the G-space plot has a clear edge and it is not aligned with the axis, the two factors are correlated (US EPA, 2014a).

2.2.2 PCA

PCA is a dimension reduction process. It converts a number of interrelated variables into a smaller set of independent variables. The new independent variables are called principal components (PCs). Each variable is a linear combination of the PCs (Jackson, 1991).

The input of the PCA could include both ambient concentration data and meteorological parameters. The units of the input data are not required to be the same but the interval time of the observations must be the same.

Usually, the first step before running the PCA is to standardize the data because standardization makes both large and small magnitude variables have the same opportunity to influence the analysis. When the standardization process is not done before running PCA, the variables with small magnitudes may be ignored during the analysis (Thurston & Spengler, 1985a). The raw data are standardized to dimensionless data using

Equation (4):

$$Z_{it} = \frac{C_{it} - C_i}{\sigma_i} \quad (4)$$

where:

C_i=the mean value for the ith species over all observations;

σ_i= the standard deviation for the ith species over all observations;

Z_{it}=standardized value of the ith species in the tth sample.

The following discussion is based on the statistical software SPSS 22.0 (IBM Corp., 2013). The component matrix, the score matrix, and the total variance explained matrix are the outputs of the PCA. The columns of the component matrix represent the principal components and the rows represent the factor loadings of the variables. The score matrix represents the input matrix in the new principal component space. The data in score matrix are the projections of the original data on each of the main component vectors. The total variance explained matrix returns variance explained by each factor and the cumulative variance explained by several factors in percentage. The percent variance explained by the component is calculated by dividing the eigenvalue of the component by the sum of eigenvalues of all components. The principal components are placed in descending order according to the variance they explained.

The number of components to retain for further interpretation needs to be determined. The Kaiser Criterion, the percentage variance explained by components, and the scree plot are the three most commonly used criteria to retain the principle components. The Kaiser Criterion retains the components with eigenvalues larger than 1 because the average value of the eigenvalues derived from the correlation matrix of the input data is 1 (Cangelosi & Goriely, 2007). The components could also be retained according to their stacked percentage of explained variance. The components with a stacked percentage of explained variance larger than a user-determined value should be retained (Jolliffe, 2002). The user-determined cut-off value was usually between 70% and 90% (Jolliffe, 2002). The scree plot is a line plot of the eigenvalues versus the number of components. The components before the inflection point in the slope should be retained (Cangelosi & Goriely, 2007).

Because the unrotated components usually could be assigned to several sources (Thurston & Spengler, 1985a), rotations should be adapted to make the factors more meaningful in practice after determining the number of factors to keep. The rotation methods could be sorted into two categories (Thurston & Spengler, 1985a): Orthogonal rotations and Oblique rotations. Orthogonal rotations assume that the factors are not correlated while Oblique rotations assume that the factors are correlated (Brown, 2009). The Varimax rotation, an orthogonal rotation, is usually used in source apportionment because the sources are assumed to be independent (Thurston & Spengler, 1985a).

The component loadings in PCA output indicate the correlation between the species and the principal component (Beaumont, 2012). The source profiles are the mass fractions of each compound contained in a source (US EPA, 2004). The component loadings in PCA and the source profiles have very different meaning. Therefore, a comparison between PCA outputs and the source profiles cannot be conducted. However, a qualitative comparison between major variables of the principal components and the outcomes from other studies allows to interpret the PCA results. The variables with the loading larger than the user-specified cut-off value are considered as major variables. In addition, the signs of the major variables should also be considered. The loadings with the same sign indicates the value of the variables increase or decrease together while the loadings with different signs indicates that one variable increases as the others decrease. The component is named to the sources and processes containing the same major variables with the same sign in the outcomes of other studies.

The interpretation of the principal component in a speciated mercury source apportionment study using PCA (Huang et al., 2010) is used as an example to illustrate the meaning of the factor loadings and the interpretation of PCA results. The PCA results in Huang et al. (2010)'s study are listed in Table 2.1. According to Huang et al. (2010), the variables with a loading larger than 0.50 was thought as the major variable of the factor. In PC1, the major variables are Hg_0, Temperature, and Melting (Table 2.1). The negative loadings on Melting variable and Temperatures and The positive loading of Hg_0 in this factor suggested that the Hg_0 concentrations increase as the snow melts. Significant Hg_0 fluxes during snowmelt was observed in other studies (e.g. Ferrari et al., 2008; Choi and Holsen, 2009). The presence of Hg_0 in this factor was consistent with the ice melting process. Therefore, PC1 was named as snow melting factor. The rest three factors in this study were interpreted by the authors using the same method (Huang et al., 2010).

Table 2.1: An example of the speciated mercury source apportionment result by PCA (blank indicates factor loading lower than 0.1; bold indicates important variables with factor loadings larger than 0.5) (Huang et al., 2010)

	PC1	PC2	PC3	PC4
Hg_0	**0.70**	0.14	0.36	-0.25
RGM	0.12	**0.57**		**0.58**
Hg_p		0.15	0.2	**0.89**
SO_2		**0.92**		0.12
O_3			**-0.81**	-0.2
CO	-0.38		**0.56**	0.47
$PM_{2.5}$	0.22	0.24	0.65	0.26
Temperature	**0.93**			
Melting	**0.91**			
Barometric pressure	-0.45		0.31	
Wind speed	0.26	-0.1	**-0.73**	
Relative humidity		-0.1	**0.55**	

2.2.3 Treatment of missing values

Some receptor models are not able to deal with missing values in the input data. Therefore, missing values must be treated before running the models. Exclusions and imputations of the missing values are two common methods to deal with the missing values.

Exclusion of the missing values includes listwise deletion and pairwise deletion. The listwise deletion excludes all the cases having one or more missing values. The listwise deletion provides a full data matrix, but it may cause a large reduction of the dataset when one species has many missing values. The listwise deletion may also bias the results because the listwise deletion benefits the cases with high concentrations (Huang et al., 1999). The pairwise deletion only deletes the information when the analysis needs it. The pairwise deletion maximizes the use of the data and this method is usually used in the covariance calculation. However, the covariance matrix may not be positive definite using pairwise deletion and would result in negative eigenvalues (IBM Corp., 2014). To adapt pairwise deletion, the missing values in the dataset are required to be missing at random (IBM Corp., 2014).

Imputation is to replace the missing values with a predicted value based on the information of the dataset. Median and mean concentrations are the two most commonly used values in imputation because they represent the central tendency of the variable concentrations (Acuña & Rodriguez, 2004). However, the mean value could be affected by the outliers in the dataset while the median value is less affected by the amount of the outliers (Acuña & Rodriguez, 2004). Median imputation is also recommended when the dataset is skewed (Acuña & Rodriguez, 2004). Other than the median, the geometric mean is also less affected by the extreme values and could represent the central tendency of the data set in environmental studies (Blackwood, 1992; Parkin & Robinson, 1992; Pekey et al., 2004). Although imputation provides a complete input matrix, imputation could be useless or even harmful when the imputation values and the non-missing species have similar information (Acuña & Rodriguez, 2004). Too many imputations in the dataset could also reduce the variance of the dataset and affect the further analyzes (Acuña & Rodriguez, 2004).

Only listwise exclusion or imputation could be adapted by PMF model to deal with missing values because the calculation of PMF model was completed at one time which requires a complete dataset. However, the calculations of PCA do not require to include all variables at one time. Therefore, all methods mentioned in this section including pairwise exclusion, listwise exclusion and imputation could be accepted by PCA. When an inter-comparison between the receptor models need to be done, the treatment of the missing values of the different receptor models should be the same.

2.3 Review of studies about source apportionment of speciated mercury

10 papers (Lynam & Keeler, 2006; Swartzendruber et al., 2006; Gao, 2007; Liu et al., 2007; Li et al., 2008; Cheng et al., 2009; Huang et al., 2010; Cheng et al., 2012; Cheng et al., 2013; Ren et al., 2014) about source apportionment of speciated mercury (GEM, GOM, and PBM) were reviewed**.**

2.3.1 Methods used

All papers used PCA to find the major sources/processes affecting the ambient mercury concentrations while only one paper (Cheng et al., 2009) used both PCA and PMF model to identify the sources affecting the ambient mercury concentrations. The study (Cheng et al., 2009) used PMF model had 7 compounds available in the input data. Only the types of sources (e.g. combustion source) were identified in the study (Cheng et al., 2009) due to a lack of makers in the input data. Therefore, the uncertainties of the PMF results were high when the variables included were not enough. It could be noted that 8 out of 10 studies included less than 8 chemical species in the input file (Table 2.2).

Therefore, it is not unexpected that most (9 out of 10) of the studies did not use PMF. PCA was able to identify the sources affecting the ambient speciated mercury in all 10 studies with fewer compounds. The inclusion of the meteorological parameters in the PCA helped to identify the meteorological processes.

Table 2.2: Variables used in 10 reviewed Hg studies.

Paper	SO_2	O_3	CO	NO_x	NO	NO_2	$PM_{2.5}$	Ca^{2+}	Mg^{2+}	K^+	Na^+	Cl^-	NH_4^+	NO_3^-	SO_4^{2-}	HNO_3	BrO	black carbon	PM_{10}	Total
Swartzendruber et al., 2006		√	√																	2
Liu et al., 2007	√	√		√																3
Huang et al., 2010;	√	√	√				√													4
Lynam and Keeler, 2006	√	√	√	√			√													5
Gao, 2007	√	√	√	√			√													5
Ren et al., 2014	√	√	√		√												√	√		6
Cheng et al., 2009	√	√	√	√	√	√	√													7
Li et al., 2008	√	√	√	√	√	√													√	7
Cheng et al., 2012;	√	√						√	√	√	√	√	√	√	√	√				11
Cheng et al., 2013	√	√					√	√	√	√	√	√	√	√	√	√				12

2.3.2 Identified sources and processes affecting speciated mercury

The factors and processes identified in the 10 reviewed papers are categorized to anthropogenic sources, natural sources and re-emission, and atmosphere processes in Table 2.3. Natural sources and the re-emission of mercury are merged into one category because the re-emission of mercury is hard to distinguish from the natural sources (UNEP, 2008).

Table 2.3: Category and frequency of factors identified in 10 reviewed Hg studies.

Category	Factors identified	Frequency in 10 papers
Anthropogenic sources	Combustion	8
	Long range transport of industrial emission	2
	Industrial process	2
	Mobile sources	2
	Sewage treatment	1
Natural sources and re-emission	Road salt	1
	Crustal/soil sources	1
	Snow melting	1
	Mercury evasion from ocean	1
Atmosphere Processes	Photochemistry & oxidation	8
	Precipitation	2
	Mixing	2
	Climatological seasonality	1
Not identified	Not identified	1

Combustion processes and mobile sources were identified eight times and twice, respectively (Table 2.3). This is not unexpected because the combustion process is one of the largest anthropogenic sources of mercury (UNEP, 2013). The chemical species like PBM, GOM, NO_x, SO_2, and CO were used to identify combustion processes (Lynam & Keeler, 2006; Huang et al., 2010; Ren et al., 2014). Sulfur dioxide was the indicator of fossil fuel (coal) combustion because almost all SO_2 sources are related to fossil fuel combustion (Liu et al., 2007). The mobile sources including the combustion of oil and gasoline are the main anthropogenic sources of nitrogen oxides (NO_x) (Liu et al., 2007). Carbon monoxide (CO) is used to identify all kinds of combustion process because it is formed by incomplete combustion. Combustion emission could release all three mercury forms but GOM and PBM are better markers of combustion processes compared to GEM. This is because the background concentration of GEM is 100-1000 times higher than GOM and PBM. Therefore, GEM released by combustion sources may not have a significant impact on ambient concentrations of GEM comparing to GOM and PBM (Huang et al., 2010).

Four studies (Li et al., 2008; Cheng et al., 2009; Cheng et al., 2012; Cheng et al., 2013) have identified industrial emissions and two of them are long range transport of industrial processes (Table 2.3). The identified industrial processes in these four studies had similar markers (SO_2, NO_x, SO_4^{2-}, HNO_3 and NO_3^-) as combustion process. The markers used in the studies were not enough to separate the industrial sources from the combustion emissions (Li et al., 2008; Cheng et al., 2009; Cheng et al., 2012; Cheng et al., 2013). Further analysis such as back trajectory, ratio analysis, and pollution roses were used to distinguish the industrial process from the combustion process in these studies. The back trajectories model shows the place the air mass passed. If the air mass passed industrial areas before reaching the receptor site, the factor is likely related to the industrial process. The ratio analysis could also be used to verify the industrial emission because the ratios of NO_2 to Hg, $PM_{2.5}$ to Hg, and SO_2 to Hg differs among different sources (Cheng et al., 2009). When the ratio is in the range of industrial process, then the factor was named as industrial process (Manolopoulos & Snyder, 2007). The pollution rose shows where the mercury and the combustion indicators come from. When mercury and the combustion indicators come from the direction of the industrialized area, the factor is very likely to be related to industrial sources.

Four factors (road salt, crustal/soil source, snow melting and mercury evasion from the ocean, Table 2.3) were found related to the natural emissions and the re-emission of mercury (Huang et al., 2010; Cheng et al., 2012; Cheng et al., 2013). The factor named as road salt (Cheng et al., 2012) had high positive loadings of Na^+ and Cl^- and a negative loading of temperature. Although these species are indicators of sea salt, the study site is less likely to be affected by marine aerosols because it is located in the middle of the continent (Cheng et al., 2012). The negative loading on temperature indicates that the factor is more likely to be related to road salts used to melt the snow in winter (Cheng et al., 2012).

The factor with high positive loadings of K^+, Mg^{2+} and Ca^{2+} (Cheng et al., 2012) was named as crustal/soil sources because these chemical species are the indicators of soil sources (Zhang et al., 2008). This factor also had a minor loading on PBM. This is probably related with the re-emission of the previously deposited mercury from crustal/soil sources.

The factor snow melting (Huang et al., 2010) had positive loadings of GEM, temperature and a variable called “melting”. This finding indicates that the GEM in this factor was probably emitted directly by the snow or formed by the reduction of GOM related to snow (by ultraviolet light) (Lalonde et al., 2002).

The factor mercury evasion from the ocean (Cheng et al., 2013) had positive loadings of GEM, precipitation, wind speed and relative humidity. Back trajectories analysis showed that GEM concentrations, precipitations, air temperature, wind speed and relative humidity are higher for oceanic trajectories. Sigler et al. (2009) suggested that the Atlantic storms might enhance the GEM evasion from the ocean and the precipitation scavenged the sea salt aerosols. This is consistent with the high precipitation loading and the lack of sea salt tracing ions.

Four processes affecting ambient mercury concentrations were identified. They were one photochemical process and three meteorological processes (Table 2.3). Eight out of ten papers identified the photochemical process or oxidation process. These factors usually had high positive loadings of O_3 and GOM (Lynam & Keeler, 2005; Gao, 2007; Huang et al., 2010; Cheng et al., 2012). The high loading on O_3 indicates ozone-rich environment and GEM could be oxidized to GOM under the ozone-rich environment (Pal & Ariya, 2004). Therefore, the presence of O_3 and GOM indicates photochemical processes.

Three meteorological processes identified are precipitation, mixing, and seasonality (Table 2.3). The factors named precipitation (Huang et al., 2010; Cheng et al., 2013) had negative loadings on PBM and/or GOM and positive loadings on relative humidity and precipitation. The positive loadings on precipitation and relative humidity indicate precipitation process. The negative loadings on PBM and/or GOM is consistent with precipitation process because precipitation removes PBM and GOM from the atmosphere (UNEP, 2008) This is consistent with the negative loading of PBM and GOM.

There are two mixing factors including diurnal mixing (Liu et al., 2007) and mixing of free-tropospheric (Swartzendruber et al., 2006). Factor diurnal mixing (Liu et al., 2007) had positive loadings on O_3, mixing height and wind speed and a negative loading on NO_3^-. The presence of mixing height and wind speed indicates mixing processes and the mixing processes dilute the concentrations of other pollutants (Liu et al., 2007). The negative loading of NO_3^- was caused by strong vertical mixing which led to low NO_3^- concentrations. The factor mixing of free-tropospheric air (Swartzendruber et al., 2006) was a unique factor identified at the mountain site. This factor had a negative loading on GEM and positive loadings on GOM, O_3, and water vapor. Daytime anabatic winds brought the marine air at daytime and nighttime katabatic winds brought free-tropospheric air at night to the site. The marine air containing higher water vapor compared to continental air mass brought to the site by wind is consistent with the positive loading of water vapor. Free tropospheric air brings O_3 resulted in the positive loading of O_3 (Swartzendruber et al., 2006). High O_3 concentrations may oxidize the GEM to GOM. This is consistent with the negative loading of GEM and the positive loading of GOM.

Factor named as seasonality (Liu et al., 2007) has a high positive loading for water vapor mixing ratio (WVMR), moderate positive loadings for GEM and O_3 and moderate negative loading for PBM and wind speed (WS). The values of WVMR, O_3 and GEM were highest in warm seasons and lowest in cold seasons while the values of PBM and WS were highest in the cold season and lowest in the warm season (Liu et al., 2007). Therefore, this factor was assigned to climatological seasonality.

One factor (Table 2.3, Cheng et al., 2009) was not determined because it only had high loadings of all three mercury forms and $PM_{2.5}$. Mercury and $PM_{2.5}$ have too many common sources such as combustion processes, industrial sources and condensation processes. Therefore, this factor was left undetermined.

2.3.3 Uncertainty of PCA and PMF results

There are uncertainties in PCA and PMF results which could be caused by measurements and interpretation process. The measurement of concentrations near the detection limit has larger uncertainties compared to the concentrations far beyond the detection limit (Croghan & Egeghy, 2003). The detection limits of GEM, GOM and PBM are 0.1 ng/m^3, 2 pg/m^3 and 2 pg/m^3 according to Tekran Inc. (Tekran Inc., 2010), respectively. In the ten reviewed studies, all of the mean concentrations or median concentrations of GEM were beyond a factor of 10 times MDL (Table 2.4) which led to low uncertainties of GEM concentrations measurement. The mean or median concentrations of GOM are within a factor of 5 times MDL in 10 out of 13 datasets (Table 2.4) while the mean or median concentrations of PBM are within 5 times of MDL in 8 out of 13 datasets (Table 2.4). Therefore, GOM and PBM concentrations in most of the studies were near the MDL which led to high uncertainties in the input file (Croghan & Egeghy, 2003).

Table 2.4: The mean concentrations of speciated mercury with the standard deviation or relative standard deviation (%) in bracket and their ratios to MDL

Paper	GEM		GOM		PBM	
	Concentration (ng/m^3)	Ratio to MDL	Concentration (pg/m^3)	Ratio to MDL	Concentration (pg/m^3)	Ratio to MDL
Cheng et al., 2013 (HF,2009)[a]	1.24(0.27)	12.4	0.96(1.86)	0.5	2.38(3.63)	1.2
Cheng et al., 2013 (KEJ, 2010)[a]	1.34(0.20)	13.4	0.44(0.96)	0.2	3.41(4.53)	1.7
Cheng et al., 2013 (HF,2010)[a]	1.37(0.37)	13.7	1.66(2.51)	0.8	4.06(2.95)	2
Cheng et al., 2013 (KEJ, 2009)[a]	1.39(0.31)	13.9	1.82(4.17)	0.9	2.87(2.93)	1.4
Ren et al., 2014	1.42(0.12)	14.2	5.4(10.2)	2.7	3.1(1.9)	1.6
Huang et al., 2010	1.49 (18%)	14.9	4.08 (192%)	2	6.57 (93%)	3.3
Liu et al., 2007	1.5(median)	15	2.9(median)	1.5	Not provided	-
Swartzendruber et al., 2006	1.54(0.176)	15.4	43(82)	**21.5**	5.2(4.4)	2.6
Cheng et al., 2012	1.57(0.22)	15.7	0.99(1.89)	0.5	4.42(3.67)	2.2
Gao, 2007	1.60(0.240)	16	5.03(5.44)	2.5	10.78(22.02)	**5.4**
Li et al., 2008	1.96(0.38)	19.6	2.53(4.09)	1.3	12.50(5.88)	**6.3**
Liu et al., 2007	2.2(1.3)	22	17.7(28.9)	**8.9**	20.8(30.0)	**10.4**
Cheng et al. (2009)	4.5(3.1)	45	21.5(16.4)	**10.8**	14.2(13.2)	**7.1**

a: KEJ, HF indicated site name.
GOM and PBM concentrations beyond a factor of 5 times MDL are bolded.

Other than the detection limits, uncertainties of the results could also be caused by the measuring method. Lynam and Keeler (2005) suggested that GOM could be collected on the particle filter of the Tekran system. This leads to an overestimate of PBM concentrations and an underestimate of GOM concentrations. Combining GOM and PBM concentrations to reactive mercury (RM) was thought to be able to improve the data quality (Cheng et al., 2016). In Cheng et al. (2016)'s study, GOM and PBM were combined to RM to improve the input data quality. An additional what-test factor was identified after the combination of GOM and PBM compared to the results using original dataset.

Uncertainties could also be caused in the factor interpretation process due to a lack of markers to identify the specific sources (Cheng et al., 2015). For example, in Cheng et al. (2009)'s study, one component was not assigned to any sources because the factor was dominated by the high factor loadings of all three Hg forms and $PM_{2.5}$. Because mercury and $PM_{2.5}$ have too many common sources, additional markers were needed to identify this factor. The rest three factors in the same study (Cheng et al., 2009) were not specific enough to keep them distinct (Cheng et al., 2009). All these three factors were related to combustion process because most of the makers used in this study (e.g. SO_2, NO, NO_2, NO_x, O_3, and $PM_{2.5}$) are from combustion processes (Cheng et al., 2009). More markers need to be included to identify the specific types of the combustion process. For example, black carbon and levoglucosan could be included to identify biomass combustion or wildfires (Puxbaum et al., 2007).

Uncertainties could also be caused in the factor interpretation process due to a lack of knowledge about physical and chemical mechanisms of mercury. For example, 8 out of 10 papers (Table 2.3) had assigned a factor with positive loadings of O_3 and

GOM to photochemistry. However, positive loadings of O_3 and GOM only account for the correlation between those two species and cannot reach a causal relationship between them. Theoretical studies also found that mercury oxidization by O_3 is inefficient (Calvert & Lindberg, 2005). More studies on the chemical and physical process of mercury need to be conducted to improve our understanding of the mechanism of mercury oxidation. These kinds of studies will help us interpreting the Hg source apportionment results (Cheng et al., 2015).

2.4 Inter-comparison of the PCA and PMF results

Three studies (Viana et al., 2008a; Belis et al., 2015; Cesari et al., 2016) about the methods to inter-compare the PCA and PMF results were reviewed. As suggested by Viana et al. (2008a), the number and the characteristic of the factors/components identified by PMF model and PCA should be compared. When a factor identified by receptor model "a" could be spilt into several more homogenous and distinct factors by receptor model "b", receptor model "b" has a better performance than receptor model "a" (Belis et al., 2015). For example, a factor secondary inorganic aerosol identified in PCA was separated into secondary sulfate and secondary nitrate by PMF in Cesari et al. (2016)'s study. This indicates that PMF model is better at resolving collinear sources than PCA.

Viana et al. (2008a) also suggested to compare model performances on reproducing the measured concentrations. In Belis et al. (2015)'s study, root mean squared difference (RMSD) was used to evaluate the agreement between the modelled and observed concentrations. When the RMSD is no larger than 1, the receptor model has a good performance on reproducing the observed concentrations (Belis et al., 2015). Z-score indicating whether the differences between the predicted and observed concentrations are within the target uncertainty was calculated (Belis et al., 2015). When the absolute z-score value is lower than 2, the reproduction of the observed concentrations is acceptable. When the absolute z-score value is between 2 and 3, the reproduction of the observed concentrations is questionable (Thompson et al., 2006). More details about RMSD and Z-score could be found in Belis et al. (2015).

The sensitivity of source apportionment to the variables included was compared in another recently published paper (Cesari et al., 2016). The PMF results derived from complete dataset and incomplete dataset were comparable, indicating that the PMF results were stable and robust when different chemical species were included. However, the PCA results were sensitive to the chemical species presented in the input file (Cesari et al., 2016).

CHAPTER 3

METHOD

3.1 Site

Kejimkujik (KEJ) site is a coastal site located in Kejimkujik National Park, Nova Scotia (44.32°N; 65.2°W). The site is 165 m above the sea level and the sample inlet was 5 m above the ground (Cheng et al., 2013). The site is approximately 50 km away from the nearest coast and approximately 140 km northwest of Halifax. The site is surrounded by forests and the terrain around the site is a flat plain (Figure 3.1). This site was originally designed to monitor acid rain in the area. The site is now a part of the Canadian Air and Precipitation Monitoring Network (CAPMoN) (EC, 2015), the National Atmospheric Deposition Program (NADP) (NADP, 2016), and the National Air Pollutants Surveillance (NAPS) networks (EC, 2011b). The variables monitored in these programs were used to conduct the mercury source apportionment at this site.

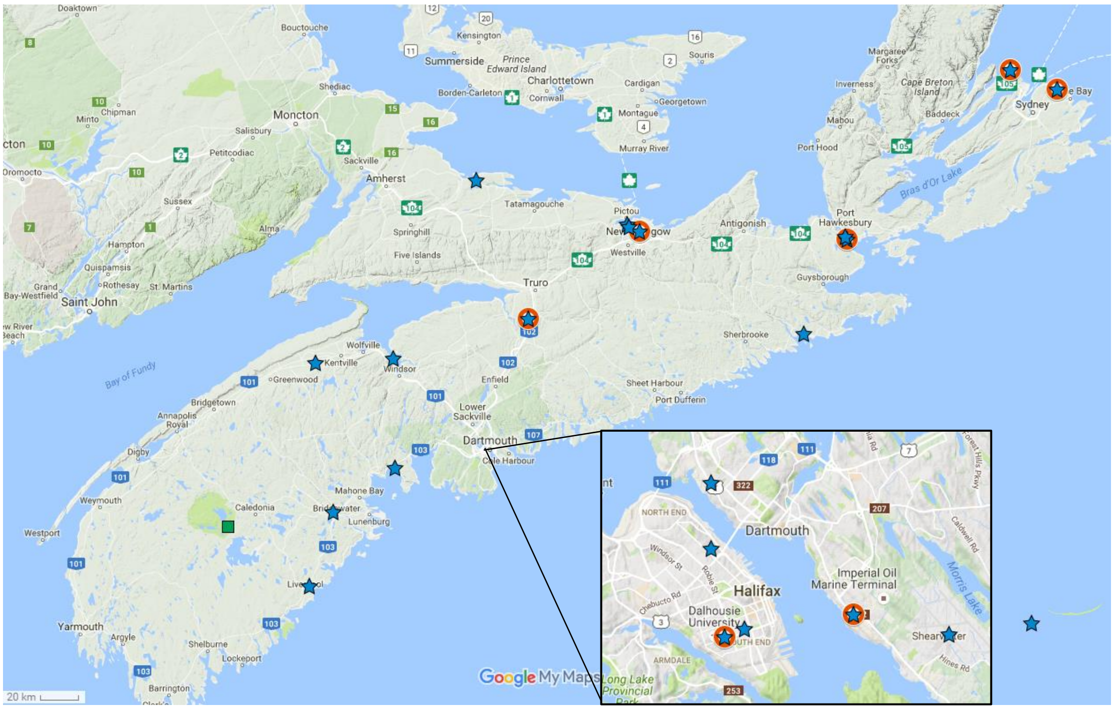

Figure 3.1: Map showing the locations of sampling site (🟩), the top 19 SO_2 or NO_x point sources (average of 2009 and 2010) (★), and all mercury point sources in 2009 and 2010 (○), in Nova Scotia.

The sampling site was selected due to the concern of the elevated Hg blood concentrations in the common loons in Kejimkujik National Park. It was found that the common loons in Kejimukujik National Park had the highest blood Hg concentrations (5.5±1.4 μg/g) in Southeastern Canada in 2007 (Evers et al., 2007). In another study conducted in the mid-1990s, yellow perch, the common loon's preferred pray, in Kejimukujik National Park and National Historic Site (Nova Scotia) was found to have the highest blood Hg concentrations across the North America (Wyn et al., 2010). In 2006 and 2007, the blood Hg concentrations in yellow perch increased an average of 29% in 10 out of 16 lakes even though the anthropogenic emission of mercury decreased form mid-1990s to 2006/07 (Wyn et al., 2010). Knowledge about the deposition rate of Hg at this site could help the interpretation of decreasing Hg emission and increasing blood Hg concentrations in common loons and yellow perch in the study area. As pointed out by Evers et al. (2007), the increasing blood Hg concentrations in common loons and yellow perch at this site posed a great threat the common loons and the ecosystem health. Therefore, the factors affecting the mercury deposition rate including Hg sources affecting the ambient mercury concentrations at this site need further investigationAccording to the National Pollutant Release Inventory (NPRI) of 2009 and 2010 (EC, 2016), seven sources emitting mercury in Nova Scotia were reported by Environment Canada (2016) in 2009 and 2010 (Figure 3.1). They were four electric power plants, one refinery, one cement plant, and one university. The nearest anthropogenic mercury sources (Dalhousie University, Halifax: 0.17 kg/yr, Imperial oil, Dartmouth Refinery: 2.8 kg/yr, Table A.1) were approximately 140 km northeast of the sampling site. The largest two mercury sources were Lingan Power Generating Station (450 km northeast) and Trenton Power Generating Station (250 km northeast), respectively.

Their annual mercury emissions during 2009 and 2010 were 71 kg/yr and 26 kg/yr on average, respectively. The provincial total mercury emission was 147.5 kg and 90.3 kg in 2009 and 2010, respectively. The top 19 SO_2 and NO_x point sources (average in 2009 and 2010) in Nova Scotia were also shown in Figure 3.1. Among the top 19 SO_2 and NO_x sources, 16 sources were common in the top 19 SO_2 and NO_x sources while 3 sources had large emissions only for SO_2 and NO_x, respectively. The top 19 SO_2 or NO_x sources contributed 99% and 98% to the total SO_2 emission and the total NO_x emission (average of 2009 and 2010), respectively. All seven mercury sources emitted both SO_2 and NO_x. The nearby combustion sources/industrial sources such as biomass-fueled power station and tire production factory were located approximately 50 km away from the KEJ according to the NPRI (Table A.1). The largest two common sources of SO_2 and NO_x are a power generation station and a refinery in Dartmouth approximately 140 km northeast of the sampling site (Table A.1). In addition to the pollutants emissions within the province, KEJ site was possibly affected by the mercury emissions from U.S. because there are high-density mercury sources in eastern U.S.

3.2 Data collection

GEM, GOM, and PBM concentrations were collected in 2009 and 2010 using Tekran® instruments (Models 1130/1135/2537). The speciated mercury concentrations were measured at a 3-hour interval. In the first two hours, the equipment measured GEM concentrations every 5 minutes and collected GOM and PBM. In the third hour, the GOM and PBM concentrations were quantified by the equipment (Cheng et al., 2013). The method detection limits of speciated mercury used in this study were 0.1 ng/m^3, 2 pg/m^3, and 2 pg/m^3 for GEM, GOM, and PBM, respectively (Tekran Inc., 2010).

The collection efficiencies of GOM and PBM were uncertain because there were no calibration standards for GOM and PBM (Huang et al., 2013). The collection efficiency for GEM could be quantified because the manual injection of GEM was performed to adjust the equipment (Gustin et al., 2015). The quality of the speciated mercury data was checked by Environment Canada using Research Data Management Quality (RDMQ) software (Steffen et al., 2012).

Other air pollutants beyond mercury species were included to aid the source apportionment of mercury species. The hourly concentrations of ground-level O_3 and meteorological parameters (relative humidity, wind speed, temperature, and precipitation amount, wind direction) were monitored at this site (EC, 2011a; EC, 2011b). The concentrations of $PM_{2.5}$, SO_2, HNO_3, and SO_4^{2-}, NO_3^-, Mg^{2+}, Cl^-, K^+, Ca^{2+}, NH_4^+, Na^+ on $PM_{2.5}$ at KEJ site were available from CAPMoN and NAPS (EC, 2011b; EC, 2015). The dataset with air pollutants only were used in both PCA and PMF model while the meteorological data were only used in PCA. All variables were converted to daily average because the time interval of the variables need to be the same. All of the data are provided by Environment Canada. (EC, 2011a; EC, 2011a; EC, 2015). It should be noted that the same datasets were also used in another PCA study conducted at this site (Cheng et al., 2013).

According to previous studies (Lynam & Keeler, 2006; Viana et al., 2008b; Watson et al., 2008; Zhang et al., 2008; Cheng et al., 2012), the variables included in this study were used as the indicators of specific sources. Variables including SO_2, HNO_3, SO_4^{2-} NO_3^- and $PM_{2.5}$ were the indicators of combustion and industrial source (Watson et al., 2008; Zhang et al., 2008; Cheng et al., 2012). Ozone indicated photochemical process or gas phase oxidation of mercury (Lynam & Keeler, 2006). Potassium ions and calcium ions were the indicators of biomass combustion and soil/crustal emission, respectively (Andersen et al., 2007; Zhang et al., 2008). Marine aerosols and road salts (i.e. salt used to melt ice in winter) were indicated by Mg^{2+}, Cl^- and Na^+ (Brennan et al., 2004). Ammonia indicated agriculture activities (Zhang et al., 2008). Meteorological parameters were used to identify the processes affecting ambient mercury concentrations in PCA. The positive loadings of relative humidity and precipitation indicated the wet deposition (Huang et al., 2010). Wind speed indicated air mass transport. Temperature affected the natural mercury sources or re-emission (Laurier et al., 2003).

The general statistics of the input data are listed in Table 3.1 and Table 3.2 for 2009 and 2010, respectively. Three speciated mercury forms were not missing at the same time in 2009. PBM had the largest amount of missing values among all chemical species in 2009. The ratios of the geometric mean value to the median value for GEM, GOM, and PBM in 2009 were all close to 1 (0.97, 1.36, and 0.83, respectively). The total ions were calculated by summing up all ions. It contributed 80% of the total PM mass. It should be noted that GOM has a larger geometric mean value than the median values. Three mercury forms and the non-mercury species were missing at the same time. The non-mercury species had more missing values in 2010. The ratios of the geometric mean to the median for all three mercury forms in 2010 were also close to 1 (0.97, 1.29, and 0.95, for GEM, GOM, and PBM, respectively). The number of missing values observed in 2010 were less than the missing values observed in 2009.

Table 3.1: General statistics of air pollutant concentrations ($\mu g/m^3$, for non-mercury species) in the input files in 2009.

Species	Number of the missing values	Percent of missing values	MDL	Percent of values below MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m^3)	112	31%	0.1	0%	1.37	1.41	1.39	0.26
GOM (pg/m^3)	115	32%	2	79%	0.57	0.42	1.77	3.70
PBM (pg/m^3)	151	41%	2	48%	1.78	2.15	2.81	2.72
PM	73	20%	1	9%	2.71	2.91	3.44	2.49
O_3	0	0%	4.3	0%	59.4	62.1	62.4	19.1
SO_2	12	3%	0.002	0%	0.20	0.22	0.40	0.51
HNO_3	11	3%	0.05	12%	0.13	0.12	0.19	0.22
Ca^{2+}	2	1%	0.002	0%	0.05	0.05	0.06	0.04
K^+	2	1%	0.029	17%	0.04	0.03	0.04	0.03
Na^+	2	1%	0.05	9%	0.25	0.30	0.43	0.47
Mg^{2+}	2	1%	0.0004	2%	0.04	0.04	0.06	0.06
Cl^-	2	1%	0.046	23%	0.19	0.23	0.46	0.64
NO_3^-	2	1%	0.06	9%	0.18	0.17	0.28	0.39
NH_4^+	2	1%	0.001	0%	0.19	0.17	0.28	0.32
SO_4^{2-}	2	1%	0.05	0%	0.78	0.76	1.14	1.27
Total ions	2	1%	-	-	2.13	2.05	2.76	2.23

Table 3.2: General statistics of air pollutant concentrations ($\mu g/m^3$ for non-mercury species) in the input files in 2010, MDL same as in Table 3.1.

Species	Number of the missing values	Percent of missing values	Percent of values below MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m^3)	16	4%	0%	1.34	1.38	1.35	0.17
GOM (pg/m^3)	16	4%	96%	0.27	0.21	0.44	0.64
PBM (pg/m^3)	16	4%	46%	2.08	2.20	3.40	4.13
O_3	3	1%	0%	62.2	63.4	64.5	16.6
SO_2	68	19%	1%	0.10	0.13	0.23	0.31
HNO_3	68	19%	25%	0.10	0.10	0.18	0.22
Ca^{2+}	68	19%	0%	0.04	0.04	0.07	0.13
K^+	68	19%	46%	0.04	0.03	0.06	0.07
Na^+	68	19%	16%	0.20	0.24	0.40	0.53
Mg^{2+}	68	19%	0 %	0.03	0.04	0.05	0.06
Cl^-	68	19%	27%	0.14	0.15	0.46	0.83
NO_3^-	68	19%	21%	0.14	0.13	0.25	0.36
NH_4^+	68	19%	0%	0.16	0.15	0.30	0.57
SO_4^{2-}	68	19%	0%	0.69	0.64	1.11	1.65
Total ions	68	19%	-	1.89	1.80	2.71	2.95

The large amount of GOM and PBM concentrations below the MDL (78% and 96% for GOM in 2009 and 2010, respectively; 48% and 46% for PBM in 2009 and

2010, respectively, Table 3.1 and Table 3.2) indicated the high corresponding uncertainties because the concentrations below the MDL usually have higher uncertainties compared to the concentrations above the MDL (Croghan & Egeghy, 2003). The high uncertainties observed in the GOM and PBM concentrations measurement was caused by the collection of GOM on the PBM filter (Lynam & Keeler, 2005). d in GOM concentrations biased low and PBM concentrations biased high. Combining GOM and PBM to reactive mercury (RM) (Gustin et al., 2013; Weiss-Penzias et al., 2015) and excluding GOM and PBM from the dataset were conducted to improve data quality. The MDL for RM was considered to be 4 pg/m^3 and was considered as missing when one of GOM and PBM was missing.

3.3 Selection of receptor models

PMF model, PCA, and CMB model are three commonly used receptor models. The inputs and outputs of these three receptor models are listed in Table 3.3 (US EPA, 2004; US EPA, 2014; Cheng et al., 2015). As seen in Table 3.3, source profiles are only required by the CMB model as input. Because the source profiles were not available in this study, CMB model was not used.

Table 3.3: Input files and output files for the three receptor models.

	CMB	PMF	PCA
Input files	1) Source profiles 2) Ambient concentration data 3) Uncertainty data	1) Ambient concentration data 2) Uncertainty data	1) Ambient concentration data and/or meteorological data
Output files	1) Source contributions 2) Model performances index	1) Source profiles 2) Source contributions 3) Model performances index	1) Source factors

PCA could also include meteorological parameters which could identify the meteorological processes affecting the ambient concentrations of pollutants. However, the interpretation of the PCA components could be subjective when the makers included are insufficient (Viana et al., 2008a). PCA results only provide qualitative components affecting ambient concentrations. The source profiles provided by PCA may have no physical meaning (Viana et al., 2008a).

The inclusion of the data uncertainty in PMF model makes the results more accurate. PMF model are also able to reduce the influence of the species with high uncertainties (US EPA, 2014a). Quantitative factor profiles and factor contributions provided by PMF output could aid future studies in factor interpretation. However, the numbers of the factors should be determined before running PMF model. The determination of the number of factors is time consuming (US EPA, 2014a). The interpretation of the PMF results could also be subjective like the interpretation of the PCA results (Viana et al., 2008a).

The source apportionment using the same dataset was conducted by PCA in another study (Cheng et al., 2013). Therefore, the source apportionment of the same dataset using PMF model were conducted in this study in order to compare the PMF results and PCA results. However, the data treatment and the input settings of PCA in Cheng et al. (2013)'s study (see details in section 3.6.1) were different from the data treatment and the input settings in PMF model (see details in section 3.5.2) in this study. This makes the PCA results in Cheng et al. (2013)'s study not comparable with the PMF results in this study. Therefore, PCA was also conducted in this study with the same data treatment and input settings as the PMF model.

3.4 Treatments of missing data

Listwise deletion was used in PMF model because it was the only deletion method accepted by PMF model. Cases using geometric mean imputations and median imputations were conducted to assess the model sensitivity to different treatments of missing values. The annual geometric mean and annual median were used because they are less affected by extreme values in the dataset than mean imputation (Pekey et al., 2004). In order to be compared with the PMF results, the listwise deletion was also used in PCA.

3.5 PMF model

3.5.1 Study design

The PMF model (version 5.0) developed by US EPA (2014a) was used in this study. The following description is based on the 2009 dataset. The run excluding missing values listwise was defined as Case 1. Case 1 was referred as the reference case in 2009. The runs replacing missing values with the annual geometric mean and the annual median were defined as Case 2 and Case 3, respectively. Case 2 and Case 3 were used to assess the sensitivity of PMF model to different treatments of missing values. The run excluding GOM and PBM and the run combining GOM and PBM into RM were defined as Case 4 and Case 5, respectively. The run using a scaling factor (Eqn 5) to increase the GOM and PBM concentrations was defined as Case 6.

$$Scaling\ factor = \sqrt{\frac{\max(x)}{x_i}} \qquad (5)$$

where

x_i: concentration of the species in the ith sample;

max(x): the maximum value of the species.

The scaling factor is large when the concentration is low, and vice versa. Case 4, Case 5 and Case 6 were used to evaluate the impact of the improvement of data quality on the model performances and factor interpretations.

The same case design was also used in 2010. The detailed information about the cases in 2009 and 2010 was listed in Table 3.4. It should be noted that $PM_{2.5}$ was not available in 2010. The details about the data process step for each case could be found in Appendix B.

Table 3.4: PMF case designs with different treatments of speciated mercury.

Case number		Variables used (n)	Treatment of missing value	Sample size	
2009	2010			2009	2010
1	7	All chemical species (15)	Excluding listwise	161	290
2	8	All chemical species (15)	Geometric mean imputation	365	365
3	9	All chemical species (15)	Median imputation	365	365
4	10	All chemical species except GOM and PBM (13)	Excluding listwise	201	290
5	11	All chemical species, but combining GOM and PBM to RM (14)	Excluding listwise	161	290
6	12	All chemical species and increasing GOM and PBM concentrations using a scaling factor (15)	Excluding listwise	161	290

The factors identified of the cases excluding or combining GOM and PBM were compared to the factors identified in the reference case. When all the factors identified in the cases excluding or combining GOM and PBM were the same as the reference case, excluding or combining GOM and PBM was acceptable. However, it should be noted that the factor contributions to GOM and PBM could be concealed even though the same factor were identified. When the factors in the reference case were not identified in the cases excluding or combining GOM and PBM, the missing factors were checked. When the missing factors contributed few to mercury concentrations, excluding or combining GOM and PBM was also acceptable. Otherwise, excluding or combining GOM and PBM was not acceptable.

3.5.2 Model setup

Equation (3) was used to calculate the equation-based uncertainties (U.S. EPA, 2014). The error fractions needed by equation (3) were assumed to be 15% of concentrations for mercury and 10% of concentrations for non-mercury compounds because most of the measured GOM and PBM concentrations had large uncertainties as concluded in Section 3.2. 100%, 200% and 1000% of geometric mean/median for GEM, GOM, and PBM, respectively, were used as the uncertainties for the samples with imputations (Polissar et al., 1998; Hopke, 2000). The ratios were chosen according to the distribution of the uncertainties of the below MDL values in Case 1 and Case 7.

After loading the input files, the time series of each variable were checked. In 2009, there were three spikes in GEM and GOM species indicating extreme events (US EPA 2014). The spikes were caused by real pollution events. In order to identify all possible Hg sources, no sample was excluded in 2009. No extreme event was observed in 2010. No step changes were observed in both years. This is expected because the collection equipment did not change in 2009 and 2010.

In this study, the S/N ratio was not useful in categorizing the variables because the uncertainties of all variables were set to a fixed fraction of the concentrations (U.S. EPA, 2014). The categorizations of the variables were set based on the scaled residuals after the first run (U.S. EPA, 2014). All variables were categorized as “strong” for the first run. All species except GOM and PBM had a good performance in scaled residual plot and observed/predicted statistics in the initial run. The second run with GOM and PBM categorized as “weak” had similar results. Because there was no significant difference between the results using different GOM and PBM categories, all species were set as “strong” in this study.

No species was selected as the total variable because dataset used in this study contained both particulate matter and gaseous pollutants and the study is focused on the speciated mercury. The default value (20) was used as the number of runs to evaluate the stability of results. A fixed seed number (25) were used so that the results could be reproduced later. The extra modeling uncertainty needs to be used when the result is not stable. It was not used because the initial run was quite stable (no large changes of the Q values among different runs).

The IM and IS were calculated to determine the number of factors. The IM and IS dropped dramatically in 2009 when the number of factor increased to 3 (Figure 3.2). In the line plot of Q(Robust) and Q(true) vs number of factors (Figure 3.3), no significant decreases were found when the number of factors is larger than 5 in 2009. Therefore, the PMF was conducted using the number of factor from 3-5 in 2009. In 2010, the decrease of IS value was gentle while the IM value experienced a drastic drop when the number of factors increased to 3 (Figure 3.4). The trend of the Q(Robust) and Q(True) in 2010 what-test is similar to 2009 (Figure 3.5). Therefore, the PMF runs with the number of factors from 3 to 5 were also conducted in 2010. The number of the factors is a compromise of the trends of these indexes and the physical meanings of the factors obtained (Cesari et al., 2016). The detailed comparison of the physical meanings of different solutions was provided in the Appendix C.

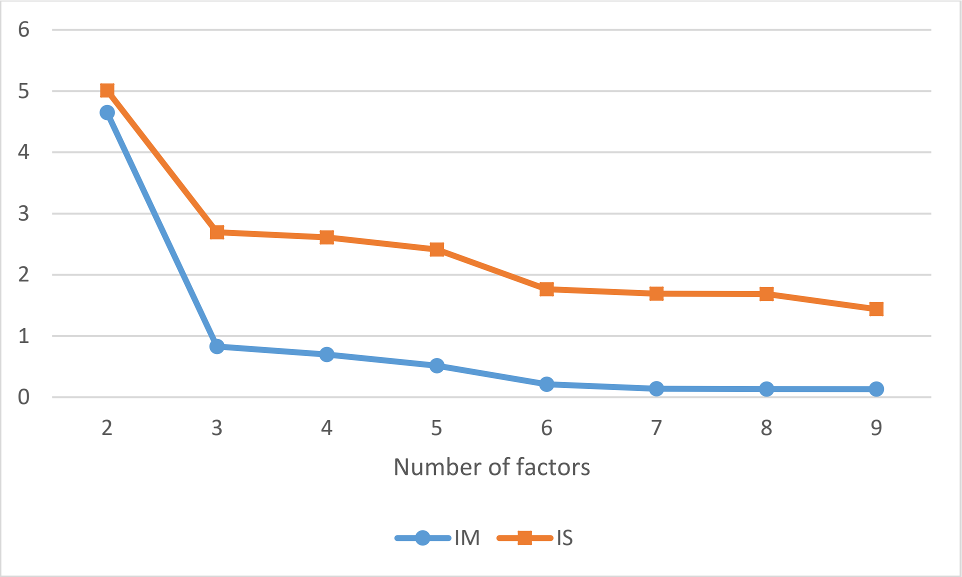

Figure 3.2: IM and IS plot vs number of factors in 2009.

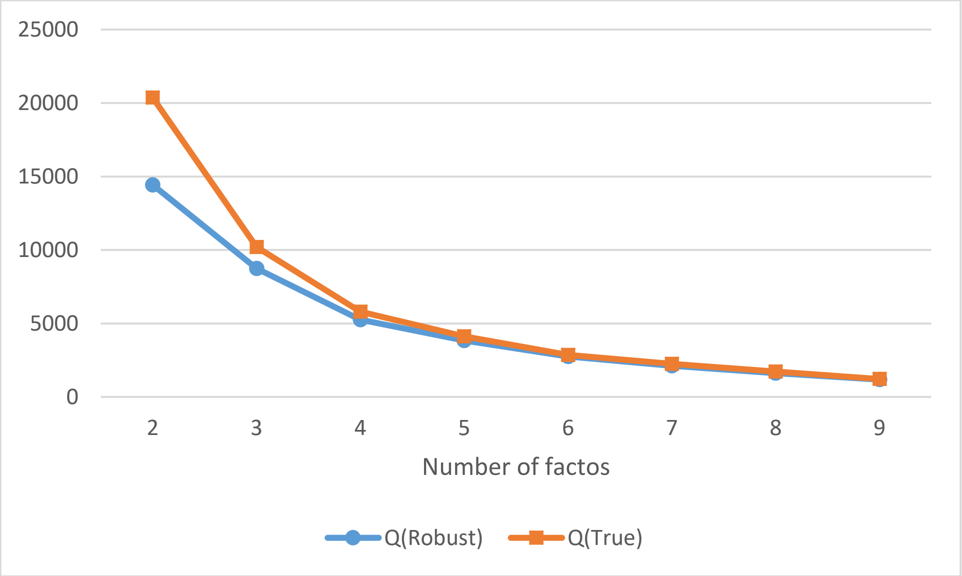

Figure 3.3: Q(Robust) and Q(true) plot vs number of factors in 2009.

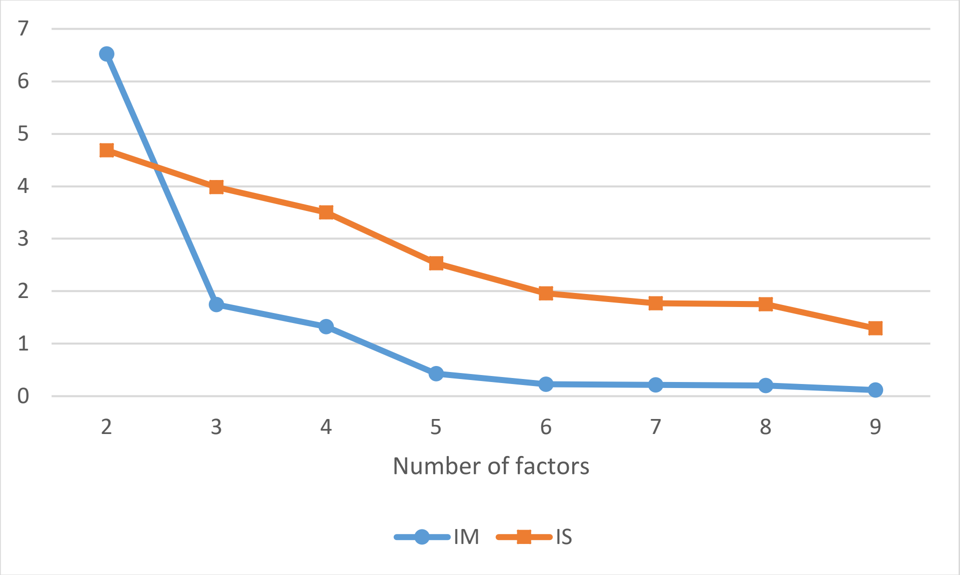

Figure 3.4: IM and IS plot vs number of factors in 2010.

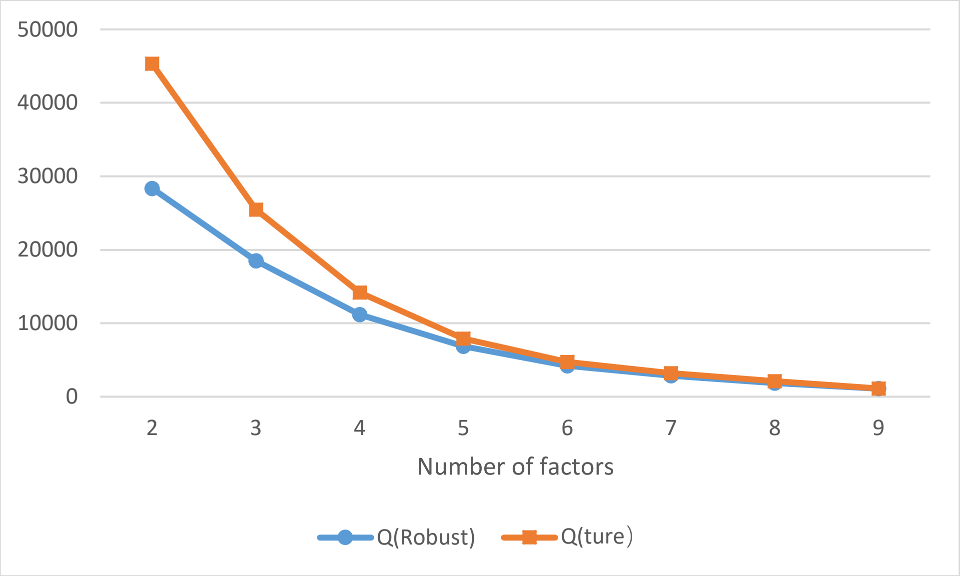

Figure 3.5: Q(Robust) and Q(true) plot vs number of factors in 2010.

3.5.3 PMF performance evaluation

The Q (robust), Q (true) and the convergence were checked. The converged run with the lowest Q (robust) value was highlighted and only the converged run was further analyzed (US EPA, 2014a). In this study, all runs in each case were convergent. Both Q (robust) and Q (true) values had small variations among 20 runs (smaller than 5, Appendix D) indicating that the results were quite stable (US EPA, 2014a). The run with the minimum Q (robust) value, which is also known as the best base run, was used for further analysis. Different in Q (true) and Q (robust) values (>8%, Appendix D) were observed indicating that uncertainties of input files were proper (US EPA, 2014a).

The model performance on reproducing the observed concentrations were evaluated by analyzing the scaled residual plot, Obs/Pred scatter plot and Obs/Pred time series. It should be noted that the predictions in graphic format such as Obs/Pred time series are provided by PMF directly but the numeric predictions of concentrations in this study was calculated using the factor contributions and factor profiles in PMF outputs. In the scaled residuals plot, when the distribution of the scaled residuals was normal and most of the residuals were below three standard deviations, the observed concentrations were reproduced well (US EPA, 2014a). The Obs/Pred scatter plot contained a regression line of observed and predicted concentrations and a one-to-one line. A small bias between the regression line and the one-to-one line indicated good agreement between the predicted concentrations and observed concentrations (US EPA, 2014a). The difference between the slopes of two lines was used to determine the bias between two lines. When the slope is larger than 1, the model overestimated the observed concentrations and vice versa. The R^2 was also provided in the Obs/Pred scatter plot screen. It represents the proportion of variance explained by the model. The closer to 1 the R^2 was, the better the agreement between the predicted and observed concentrations was. In the Obs/Pred time series, when the observed concentrations and the predicted concentrations tracked with each other, the observed concentrations were thought to be reproduced by the PMF model well (US EPA, 2014a).

Factor contributions to total predicted mercury concentrations, ratios of predicted mercury concentrations to observed mercury concentrations (Pred/Obs ratio) and ratios of predicted annual mean to observed annual mean (annual Predmean/Obsmean ratio) were calculated to verify the findings derived from Obs/Pred time series. Factors with average mercury contributions larger than 15% were considered as the major mercury sources which has a large impact on ambient mercury concentrations. Pred/Obs ratios were used to determine how well the observed concentrations were reproduced by PMF model. When the average Pred/Obs ratio was close to one, the reproduction of the observed concentrations was good. A large range of the Pred/Obs ratios indicated that the model predicted mercury concentrations did not track the observed mercury concentrations well. The annual Predmean/Obsmean ratio close to one indicated that the observed concentrations were well reproduced on an annual basis.

3.5.4 Factor interpretation

The factors were interpreted based on the comparison of the major variables of the factors and markers of the source profiles in the literature (Leuchner et al., 2015). The variables with the factor contribution larger than 25% were thought to be the major variables of the factor. The factors were assigned to the names of the sources with similar major variables.

3.6 PCA

3.6.1 Study design

In order to be compared with PMF model, the same datasets as in Case 1 and Case 7 for 2009 and 2010 were used in PCA, respectively. These two cases were defined as Case 13 and Case 15 for 2009 and 2010, respectively. Cases including both the chemical species and meteorological parameters were defined as Case 14 and Case 16 for 2009 and 2010, respectively. Case 14 and Case 16 were used to evaluate the effects of including the meteorological data on components identification. The dimension of the dataset (number of variables m × number of samples n) used in this study met the requirement to derive stable results (i.e. $n > 50+m$, Thurston & Spengler, 1985b). The detailed information about the cases in PCA is listed in Table 3.5.

Table 3.5: Input and Set-up used in PCA in this study.

Case number	Year	Variables (m)	Sample size (n)	Requirements (50+m)	Other settings
13	2009	all chemical species (15)	161	65	1) Missing value: Listwise deletion 2) Rotation: Varimax 3) Cut-off value for major variables: 0.25 4) Principal components: Kaiser criterion
14	2009	all chemical species and meteorological parameters (19)	159	69	
15	2010	all chemical species (15)	290	65	
16	2010	all chemical species and meteorological parameters (19)	285	69	

The source apportionment of speciated mercury by PCA using the same dataset was also conducted in another study (Cheng et al., 2013). The detailed information of the case design in that study was listed in Table 3.6. Pairwise deletion was used instead of listwise deletion to make the full use of the dataset (Cheng et al., 2013). Marine tracing species (Na^+, Mg^{2+}, and Cl^-) were excluded after the first run because the marine tracing species were not related to any mercury sources (Cheng et al., 2013). A fixed number of components (i.e. 4 and 3 for 2009 and 2010, respectively) was kept for further analysis (Cheng et al., 2013).

Table 3.6: Input and Set-up used in PCA in Cheng et al. (2013).

Case number	Dataset	Variables	Other settings
A	2009	All variables	1) Missing value: pairwise deletion 2) Rotation: Varimax 3) Cut-off value for major variables: 0.30 4) Principal components: criteria eigenvalue>1 used in case 1 and 2; fixed number of factors used in case 3 and case 4
B	2010	All variables	
C	2009	O_3, PM, GEM, GOM, PBM, SO_2, Ca^{2+}, HNO_3, K^+, NH_4^+, NO_3^-, SO_4^{2-}, temperature, relative humidity, wind speed, and precipitation.	
D	2010	O_3, GEM, GOM, PBM, temperature, relative humidity, wind speed, and precipitation.	

3.6.2 Model setup

The PCA analysis was conducted using SPSS 22.0 (IBM Corp., 2013). The suitability of the dataset for factor analysis was checked before running the PCA (Williams et al., 2012). Kaiser-Meyer-Olkin (KMO) index and Bartlett's Test of Sphericity were checked. When the KMO index of the dataset is larger than 0.50 and the Bartlett's Test of Sphericity is significant ($p<0.05$), the dataset is suitable for factor analysis (Tabachnick & Fidell, 2001; Hair et al., 2006). In this study, the data set for each case met both criteria (Table 3.7).

Table 3.7: KMO index with the Bartlett's Test significant at 0.001

	KMO index
Case 13	0.75
Case 14	0.70
Case 15	0.54
Case 16	0.58

After the dataset was automatically standardized, the principal components were extracted by PCA. The number of the components kept was determined according to the Kaiser Criterion (Kaiser, 1960) which kept the components with eigenvalues larger than 1 for further analysis. The varimax rotation was used because the sources were assumed to be independent.

3.6.3 Factor interpretation

The principal components after rotation were interpreted by comparing the major variables of the principal components and markers in the outcomes of other studies. Different cut-off values were tried and 0.25 was chosen as the cut-off value for the major variables in this study according to the interpretability of the results. The component was assigned to the sources and processes whose markers are the same as the major variables of this component.

3.7 Comparison of the results

3.7.1 Comparison of the results between 2009 and 2010

In PMF model, the factor profiles, factor contributions to mercury, and the model performances on reproducing the observed mercury concentrations in 2009 and 2010 were compared. When any differences between 2009 and 2010 were observed, the possible causes of the changes were investigated.

In PCA, only the sources identified were compared. The similarities and differences of the results between two years were summarized. The possible causes of the differences were investigated.

3.7.2 Comparison between PCA and PMF results

The PCA loadings and the factor profiles and contributions in PMF results have very different meanings. The variables with high loadings in the principal components in PCA indicates the strong correlations between the variable and the principal component (Beaumont, 2012). The PMF factor profiles indicate the percentage of total mass within the factor for each variable in each factor. Therefore, a comparison between the PCA and PMF results cannot be done. However, a comparison of the numbers and the characteristics of the factors/processes or components identified and the major variables could be conducted to find the weakness and strength of each method on source identification (Callén et al., 2009). The number of the factors identified by PCA and PMF model was compared. When the number of the factors and characteristic of the factors were similar in both models, the models were thought to have similar model performances on factor identification. When a model split one factor/component into more homogenous and distinct factors/components than other methods, the model was thought to be more powerful on resolving the collinear sources (Belist et al., 2015). When a factor was identified by one model but not identified by another model, the reason why it was not identified by another model was investigated.

Viana et al., (2008a) also recommended to compare the model performance on reproducing the observed concentrations. However, the reconstruction of the measured concentrations by PCA was not conducted due to the limited study period. Therefore, the comparison of the model ability to reproduce the measured concentrations was not conducted in this study.

CHAPTER 4

RESULT AND DISCUSSION

4.1 PMF results

4.1.1 Year 2009

Case 1: Reference case

The source profiles of Case 1 are listed in Table 4.1. Factor 1 is named as Combustion Emission because this factor accounts for high % of SO_4^{2-} (64%) and HNO_3 (54%) and moderate % of GOM (31%). The precursors of SO_4^{2-} and HNO_3 (SO_2 and NO_x) are from the coal combustion process and they were probably oxidized during the transport from sources to receptor sites (Liu et al., 2007). The presence of GOM is consistent with the combustion emission because the combustion process is one of the GOM sources (Carpi, 1997). The presence of NH_4^+ in this factor (71%, Table 4.1) is related to the long-range transport of the NH_3 from agriculture activities (Pitchford et al., 2009). NH_3 reacted with H_2SO_4 during the transport to form $(NH_4)_2SO_4$. The molar ratio of NH_4^+ to SO_4^{2-} is 1.7. This means the NH_3 was not enough to neutralize H_2SO_4 and the rest of the H_2SO_4 reacted with other compounds to form sulfate (Pavlovic et al., 2006). The moderate % of PM in the factor is consistent with the presence of secondary pollutants existing as particles. Because there were no combustion emissions near the sampling site according to the NPRI (EC, 2016), this factor is likely related to the transport process. Back trajectory model could be used to determine the origin of the combustion emission.

Table 4.1. Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 1 in 2009.

% of total species	F1	F2	F3	F4
GEM			76	
GOM	31		69	
PBM		29	63	
PM	42		34	(17)
O_3			72	
SO_2		82		
HNO_3	54	(21)	(25)	
Ca^{2+}	(19)		45	31
K^+	(22)		37	39
Na^+				86
Mg^{2+}				83
Cl^-				100
NO_3	(25)	(23)		40
NH_4^+	71			
SO_4^{2-}	64			(18)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Factor 2 is assigned to Industrial Source. The major variables PBM (29%) and SO_2 (82%) are the indicators of coal combustion process (Huang et al., 2010). However, no coal combustion source near the KEJ site in 2009 was reported according to the NPRI (EC, 2016). Additional variables were needed to confirm the identification of this factor. NO_3^- (23%) and HNO_3 (21%) were the third and the fourth largest variables the factor accounting for. They were included as the major variables. The presence of NO_3^- and HNO_3 indicate that the factor was related to industrial sources because NO_x, the precursor of NO_3^- and HNO_3, is mainly released by industrial sources (Liu et al., 2007). More analyses such as back trajectory or pollution wind rose are needed to verify this factor.

Factor 3 is named as Photochemistry and Re-emission of Hg because the factor accounts for high % of ozone (72%), GEM (76%), GOM (69%), and PBM (63%) and moderate % of Ca^{2+} (45%) and K^+ (37%). The high % of ozone indicates an ozone rich environment. GEM could be oxidized to GOM under this situation (Pal & Ariya, 2004). Although the results of recent studies showed that the oxidation of mercury by ozone had uncertainties, the oxidation rates of mercury by bromine were reported very fast (Goodsite et al., 2004). The location of the KEJ site is near the Atlantic which makes the oxidation of mercury by bromine possible. The PBM in this factor was likely formed by the condensation of mercury on particles. The presence of K^+ and Ca^{2+} is related to soil emission or biomass combustion (Andersen et al., 2007; Zhang et al., 2008). The site is located in Kejimkujik National Park, making the site possible to be affected by soil emission or biomass combustion. The mercury re-emission from biomass burning and land surface was estimated to contribute approximately 13% and 34% to the total global natural re-emission of mercury, respectively (Pirrone et al., 2010). Therefore, the high % of GEM the factor accounting for was likely related with the re-emission of GEM from biomass burning and soil. A factor related to soil and biomass combustion was also identified in another PCA study at this site (Cheng et al., 2013).

Factor 4 accounts for large % of Cl^- (100%), Mg^{2+} (82%) and Na^+ (86%) and moderate % of Ca^{2+} (31%), K^+ (49%) and NO_3^- (40%). The presence of Na^+, Mg^{2+}, Cl^- in the major variables indicates marine aerosols because they are rich in the composition of sea water (Brennan et al., 2004). The ions Ca^{2+} and K^+ are the fourth and fifth largest metal ions in the composition of sea water, respectively (Brennan et al., 2004). The presence of marine aerosols was reasonable because the sampling site is located in a national park near the Atlantic. The presence of NO_3^- was related to the reaction of HNO_3 and sea salt (Pakkanen, 1996). Therefore, this factor is named as marine aerosols.

The factor contributions and the ratios of predicted concentrations to observed concentrations are listed in Table 4.2. As seen in Table 4.2, the factor Photochemistry and Re-emission of Hg had the largest contributions to GEM (average 77%), GOM (average 74%), and PBM (average 69%) among all four factor indicating that the factor Photochemistry and Re-emission of Hg dominated the ambient mercury concentrations at this site. The factor Industrial Source contributes 22% on average to PBM while the factor Combustion Emission contributes 26% on average to GOM. The factor sea salt only contributes 14% on average to GEM. This is not unexpected because GEM is likely to be oxidized to GOM by the *in situ* photochemical process under the bromine-rich environment and most of the GOM in the marine boundary layer is absorbed by sea salt aerosols and deposited into the ocean again (Holmes et al., 2009). Case 1 was used as the reference case in 2009.

Table 4.2: Factor contributions to speciated mercury in Case 1 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	20	56	97	97
	Average	4	6	77	14
	Median	2	3	83	9
GOM (%)	Min	0	0	0	0
	Max	100	0	100	0
	Average	26	0	73	0
	Median	21	0	79	0
PBM (%)	Min	0	0	0	0
	Max	0	93	100	92
	Average	0	22	69	9
	Median	0	14	74	5

Case 2: Geometric mean imputation

The factor profiles of Case 2 are listed in Table 4.3. Factor 2 has the same major variables as the factor Industrial Source in Case 1. Therefore, the factor is assigned to the same name. Factor 1, factor 3 and factor 4 have similar major variables as the factor Combustion Emission, factor Photochemistry and Re-emission of Hg, and factor Sea Salt in Case 1, respectively. Factor 1 lacks GOM in the major variables but the presence of HNO_3 (62%, Table 4.3) and SO_4^{2-} (79%, Table 4.3) in the major variable is enough to identify combustion process. Factor 3 has an additional major variable of NO_3^- (34%, Table 4.3) compared to the factor Photochemistry and Re-emission of Hg in Case 1. The presence of NO_3^- is consistent with the soil emissions identified in this factor (Parmar et al., 2001). Factor 4 lacks Ca^{2+} in the major variables compared to the factor Sea Salt in Case 1. However, the major variables of Cl^-, Mg^{2+}, and Na^+ were enough to indicate that this factor was related to sea salt (Brennan et al., 2004). Therefore, factor 1, factor 3 and factor 4 were assigned to the same names as in Case 1.

Table 4.3: Factor profiles (% of species >25%, between 15% and 25% in bracket) in Case 2 in 2009.

% of total species	F1	F2	F3	F4
GEM			78	
GOM	(19)		81	
PBM		29	62	
PM	38		41	
O_3			67	
SO_2		96		
HNO_3	62	(18)	(19)	
Ca^{2+}			61	(21)
K^+	(24)		41	34
Na^+				85
Mg^{2+}				81
Cl^-				100
NO_3^-		(19)	34	37
NH_4^+	84			
SO_4^{2+}	79			(16)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

The factor contributions to speciated mercury of Case 2 are listed in Table 4.4. As seen in Table 4.4, the factor Photochemistry and Re-emission of Hg also had the largest contributions to speciated mercury (81% of GEM, 96% of GOM, and 73% of

PBM on average, Table 4.4) as in Case 1. The factor Industrial Source contributed 26% on average to PBM concentrations. The other two factors (Combustion Emission and Sea Salt) only had small contributions (<20%, Table 4.4) to mercury.

Table 4.4: Factor contributions to speciated mercury in Case 2 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	79	34	98	99
	Average	5	4	78	13
	Median	3	2	83	7
GOM (%)	Min	0	0	0	0
	Max	97	0	100	0
	Average	17	0	83	0
	Median	12	0	88	0
PBM (%)	Min	0	0	0	0
	Max	0	87	100	94
	Average	0	23	67	10
	Median	0	19	71	5

The differences in the source profiles and source contributions of Case 1 and Case 2 may relate to the differences of the correlation between mercury and chemicals markers in these two cases. The Pearson correlation coefficients between speciated mercury and other chemical species in Case 1, Case 2 and Case 3 are listed in Table 4.5. The Pearson correlation coefficients between GEM and other chemical species, especially for the species with large correlation coefficients (GOM, PBM, and O_3), did not change a lot between Case 1 and Case 2 (Table 4.5). The Pearson correlation coefficients between GOM and NH_4^+ (from 0.29 to 0.09) and SO_4^{2-} (from 0.27 to 0.06) become insignificant and the correlations between GOM and PM (from 0.31 to 0.19) and HNO_3 (from 0.45 to 0.27) were also reduced to a low level after imputation in Case 2 compared to Case 1. This may be the reason that the factor Combustion Emission contained no GOM in major variables and contributed less to GOM concentrations in Case 2 (Table 4.4) because SO_4^{2-}, HNO_3, and NH_4^+ are the markers in the factor Combustion Emission. For PBM, the correlation coefficients between PBM and SO_2 (from 0.63 to 0.50), HNO_3 (from 0.42 to 0.25), and NO_3^- (from 0.49 to 0.41) decreased. However, the correlation coefficient values were still strong enough to make these variables (PBM, SO_2, HNO_3, and NO_3^-) cluster in one factor (Industrial Source, Table 4.3). The changes of the correlation coefficients between the mercury and other chemicals were due to a large number (up to 41%) of imputation.

Table 4.5: Pearson correlation coefficients between speciated mercury and other chemical species in Case 1, Case 2 and Case 3 in 2009. (asterisk indicates insignificant at p=0.05)

	GEM			GOM			PBM		
	Case1	Case2	Case3	Case1	Case2	Case3	Case1	Case2	Case3
GEM	1.00	1.00	1.00	0.37	0.37	0.35	0.28	0.35	0.35
GOM	0.37	0.37	0.35	1.00	1.00	1.00	0.11*	0.20	0.19
PBM	0.28	0.35	0.35	0.11*	0.20	0.19	1.00	1.00	1.00
PM	0.15	0.11	0.11	0.31	0.19	0.18	0.48	0.30	0.31
O_3	0.48	0.52	0.50	0.28	0.24	0.25	0.56	0.44	0.45
SO_2	0.11*	0.16	0.16	0.21	0.19	0.19	0.63	0.50	0.52
HNO_3	0.18	0.14	0.16	0.45	0.27	0.27	0.42	0.27	0.29
Ca^{2+}	0.13*	0.10 *	0.09 *	0.39	0.36	0.36	0.28	0.26	0.26
K^+	0.01*	0.03 *	0.02 *	0.17	0.17	0.17	0.11*	0.14	0.13
Na^+	0.06*	0.12	0.10 *	-0.09*	-0.02 *	-0.02 *	0.20	0.22	0.21
Mg^{2+}	0.07*	0.12	0.10	-0.01*	0.03*	0.04 *	0.23	0.24	0.23
Cl^-	-0.01*	0.06 *	0.04 *	-0.18	-0.09*	-0.08 *	0.06*	0.11	0.10*
NO_3^-	0.14	0.15	0.13	0.17	0.15	0.15	0.49	0.41	0.41
NH_4^+	0.18	0.12	0.14	0.28	0.09*	0.09*	0.53	0.22	0.23
SO_4^{2-}	0.24	0.13	0.15	0.27	0.06*	0.05 *	0.53	0.17	0.19

Case 3: Median imputation

The factor profiles and factor contributions of Case 3 are listed in Table 4.6 and Table 4.7, respectively. The major variables of all four factors are similar to the factors in Case 2. Therefore, the factors in Case 3 are assigned to the same names as the factors in Case 2. The contributions of each factor to speciated mercury are also similar to those in Case 2.

Table 4.6: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 3 in 2009.

% of total species	F1	F2	F3	F4
GEM			79	
GOM	(15)		84	
PBM		26	65	
PM	38		42	
O_3			69	
SO_2		84		
HNO_3	63	(15)	(21)	
Ca^{2+}			62	(21)
K^+	25		41	35
Na^+				85
Mg^{2+}				81
Cl^-				100
NO_3^-		(16)	36	38
NH_4^+	85			
SO_4^{2-}	79			(16)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Table 4.7: Factor contributions to speciated mercury in Case 3 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	89	28	98	99
	Average	6	3	79	13
	Median	3	1	83	7
GOM (%)	Min	0	0	0	0
	Max	95	100	100	0
	Average	14	1	85	0
	Median	10	0	90	0
PBM (%)	Min	0	0	0	0
	Max	0	86	100	96
	Average	0	20	70	10
	Median	0	15	75	6

As seen in Table 4.5, the differences of the Pearson correlation coefficients between Case 3 and Case 2 were less than 0.03. Therefore, similar source profiles and source contributions were expected in Case 3 which is consistent with the actual results.

Case 4: Excluding GOM and PBM

The source profiles of Case 4 are listed in Table 4.8. The major variables in factor 1, factor 2, factor 3 and factor 4 are similar to the factors Combustion Emission, Industrial Source, Photochemistry and Re-emission of Hg, and Sea Salt in Case 1 (Table 4.1), respectively. Factor 1 and factor 2 lack the presence of GOM and/or PBM in major variables compared to the factors Combustion Emission and Industrial Source due to the exclusion of GOM and PBM. However, they are assigned to the same names because the presence of major species is enough to identify these factors.

Table 4.8: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 4 in 2009.

% of total species	F1	F2	F3	F4
GEM			78	
PM	41		38	(16)
O_3			75	
SO_2		82		
HNO_3	55	(19)	27	
Ca^{2+}	(16)		51	29
K^+	(25)		35	40
Na^+				86
Mg^{2+}				82
Cl^-				100
NO_3^-	(24)	(23)		39
NH_4^+	72			
SO_4^{2-}	64		(15)	(17)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

The factor contributions to GEM are listed in Table 4.9. The factor contributions to GEM are similar to that in Case 1 (Table 4.2). The factor Photochemistry and Re-emission of Hg has the largest contribution to GEM (79% on average) while each of the other three factors contributes less than 20% to the GEM concentrations.

Table 4.9: Factor contributions to GEM in Case 4 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	34	15	98	96
	Average	7	1	79	13
	Median	4	1	84	9

Excluding GOM and PBM had no impact on the source identification in this case because the identification of the factors in this study did not rely on these two species. However, the exclusion of GOM and PBM will have impacts on the identification of the factors relying on mercury such as mercury condensation process, theoretically. Excluding GOM and PBM did affect the factor contributions to mercury. For example, the factor Industrial Source had a contribution (22% on average, Table 4.2) to PBM but the factor was no longer contributing to PBM after it was excluded from the input file. Therefore, excluding GOM and PBM may affect the source identification and the source contributions.

Case 5: Combining GOM and PBM

The factor profiles of are listed in Table 4.10. The major variables of factor 1, factor 2, factor 3 and factor 4 are similar to factors Combustion Emission, Industrial Source, Photochemistry and Re-emission of Hg, and Sea Salt in Case 1 (Table 4.1), respectively. However, the loadings of GOM and PBM in related factors were replaced by RM but this did not affect the identification of the factors. Therefore, they are assigned to the same names.

Table 4.10: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 5 in 2009.

% of total species	F1	F2	F3	F4
GEM			78	
RM		(20)	68	
PM	44		35	(17)
O_3			73	
SO_2		84		

Table 4.10 – Continued

% of total species	F1	F2	F3	F4
HNO_3	56	(18)	26	
Ca^{2+}	(19)		46	31
K^+	(22)		38	39
Na^+				86
Mg^{2+}				82
Cl^-				100
NO_3^-	(24)	(23)		39
NH_4^+	76			
SO_4^{2-}	69			(19)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

The factor contributions to GEM and RM in Case 5 are listed in Table 4.11. The factor contributions to GEM and RM are similar to the factor contributions to GEM and PBM in Case 1 (Table 4.2), respectively. This is because the median concentration of PBM is approximately 5 times of the median concentration of GOM (Table 3.1). After these two species are combined to RM, the variance of the RM is dominated by PBM.

Table 4.11: Factor contributions to speciated mercury in Case 5 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	16	54	98	97
	Average	3	5	78	13
	Median	2	2	83	9
RM (%)	Min	0	0	0	0
	Max	37	83	100	18
	Average	10	16	73	1
	Median	7	11	78	0

Overall, combining GOM and PBM to RM did not affect the identification of the factor names, but it had small impact on factor contributions. Sources (i.e. Combustion Emission, Table 4.2) with large contributions to GOM in Case 1 are no longer contributing to GOM or RM (Table 4.11) when GOM and PBM were replaced by RM.

Case 6: Scaling factor

The factor profiles and factor contributions of Case 6 are listed in Table 4.12 and Table 4.13, respectively. Factor 1, factor 2 and factor 4 have the same major variables as the factor Combustion Emission, Industrial Source and Sea Salt in Case 1 (Table 4.1), respectively. Therefore, they are assigned to the same name. Factor 3 has an additional major variable HNO_3 (26%) compared to the factor Photochemistry and Re-emission of Hg in Case 1 (Table 4.1). The presence of HNO_3 is consistent with the presence of O_3 because they have the same precursor (NO_x). Therefore, this factor is also assigned to Photochemistry and Re-emission of Hg. The factor contributions to speciated mercury in this case are similar to those of Case 1 (Table 4.2). Using the scaling factor to increase the GOM and PBM concentrations did not affect the source identifications and source contributions much.

Table 4.12: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 6 in 2009.

% of total species	F1	F2	F3	F4
GEM			78	(15)
GOM(scale)	27		72	
PBM(scale)		(20)	71	
PM	40		36	(17)
O_3			73	
SO_2		85		
HNO_3	53	(21)	26	
Ca^{2+}	(17)		46	32
K^+	(21)		38	40
Na^+				86
Mg^{2+}				82
Cl^-				100
NO_3^-	(24)	(23)		40
NH_4^+	71			
SO_4^{2-}	65			(17)
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Table 4.13: Factor contributions to speciated mercury in Case 6 in 2009.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	0	65	100	99
	Average	0	7	78	15
	Median	0	4	83	10
GOM (%)	Min	0	0	0	0
	Max	97	0	100	67
	Average	23	0	75	2
	Median	18	0	81	1
PBM (%)	Min	0	0	0	0
	Max	0	88	100	96
	Average	0	16	74	10
	Median	0	10	80	6

Factors identified in each case and their ranks of mercury contribution were summarized in Table 4.14. The factor Photochemistry and Re-emission of Hg always had the largest contribution to mercury while the factor Sea Salt always had a small contribution to mercury in all cases in 2009. The factors Combustion Emission and Industrial Source had the second and third largest contributions to mercury in Case 1 and Case 6 in 2009, respectively. However, the factor Industrial Source had second largest contributions in Case 2 and Case 3 to mercury. The factor Combustion Emission did not contribute to mercury in Case 2 and Case 3 because the correlation coefficients between mercury and the markers changed due to the large amount imputations (Table 4.5).

Table 4.14: Factors identified in each case and their ranks of mercury contribution ("s" indicates the factor contributions to mercury less than 15%).

Factor Name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
2009	Case 1	2	3	1	4(s)
	Case 2	4(s)	2	1	3(s)
	Case 3	4(s)	2	1	3(s)
	Case 4	3(s)	4(s)	1	2(s)
	Case 5	4(s)	2	1	3(s)
	Case 6	2	3	1	4(s)

Table 4.14 – Continued

Factor Name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
2010	Case 7	3(s)	2	1	4(s)
	Case 8	4(s)	2	1	3(s)
	Case 9	4(s)	2	1	3(s)
	Case 10	2(s)	4(s)	1	3(s)
	Case 11	2(s)	4(s)	1	3(s)
	Case 12	3(s)	2	1	4(s)

Performances

The distribution of scaled residuals and the number of scaled residuals larger than 3 for speciated mercury are listed in Table 4.15. GEM had the best performances (i.e. normal distribution and no scaled residuals larger than 3) among all three mercury forms followed by PBM and GOM in 2009. Case 1, Case 4, Case 5 and Case 6 had similar performances on scaled residual plot. This indicates that combining or excluding GOM and PBM or increasing GOM and PBM using a scaling factor did not affect the performances on scaled residual plot much in 2009. The model performances of Case 2 and Case 3 were worse than other cases on scaled residual plot because the scaled residuals in these two cases were concentrating near zero (Table 4.15). This was resulted from the high uncertainties related to mercury due to a large number (up to 41%, Table 3.1) of imputations (i.e. the uncertainties of missing values were set to 100%, 200% and 1000% of the concentrations for GEM, GOM and PBM, respectively) (Polissar et al., 1998).

Table 4.15: PMF model performances for speciated mercury in scaled residual plot in 2009.

Species	Case number	Criteria	
		Normal distribution	Number of scaled residuals larger than 3
GEM	1	Normal	0
	2	Concentrated near zero,	5
	3	Concentrated near zero	5
	4	Normal	1
	5	Normal	0
	6	Normal	0

Table 4.15 – Continued

Species	Case number	Criteria	
		Normal distribution	Number of scaled residuals larger than 3
GOM	1	Right skewed	17
	2	Concentrated near zero, right skewed	17
	3	Concentrated near zero, right skewed	17
	4	-	-
	5	-	-
	6	Right skewed	26
PBM	1	Yes	5
	2	Right skewed	6
	3	Right skewed	6
	4	-	-
	5	Right skewed (RM)	8
	6	Left skewed	2

The coefficient of determination (R^2) and the slope of the regression line for speciated mercury in Obs/Pred scatter plot in 2009 (the Obs/Pred scatter plot could be found in Appendix D) are listed in Table 4.16. The R^2 and slope close to 1 indicate good model performances. GEM and PBM have the best performances on the slope (0.59 for GEM, Table 4.16) and the R^2 (0.59 for PBM, Table 4.16), respectively. GOM was thought to have the worst performances in Obs/Pred scatter plot among three mercury forms in 2009 because it has the lowest slope of the regression line (0.09, Table 4.16) and the R^2 value (0.23, Table 4.16). The model performance of Case 2 and Case 3 are worse than the other cases in 2009. A large number of imputations reduced the variance of the dataset and led to the worse performances on the R^2 and slope. Case 1, Case 4, and Case 5 had similar performances on GEM indicating that excluding or combining GOM and PBM did not affect the reproduction of GEM. The performances of RM in Case 5 were similar to that of PBM because RM concentrations are dominated by PBM due to its higher concentrations compared to GOM (i.e. the median concentrations of PBM is 5 times of GOM, Table 3.1). The case using scaling factor to increase the GOM and PBM concentrations had better performances on GOM and PBM (Table 4.16) than other cases because increasing the low GOM and PBM concentrations reduced the number of the concentrations below MDL (Table 4.17) and the corresponding uncertainties were also reduced.

Table 4.16: PMF model performances for speciated mercury in Obs/Pred scatter plot in 2009 (all significant at $p<0.001$).

Species	Case number	Criteria	
		Coefficient of determination (R^2)	Slope of regression line
GEM	1	0.28	0.59
	2	0.17	0.57
	3	0.15	0.54
	4	0.25	0.59
	5	0.29	0.59
	6	0.28	0.58
GOM	1	0.23	0.09
	2	0.08	0.05
	3	0.09	0.05
	4	-	-
	5	-	-
	6	0.33	0.18
PBM	1	0.57	0.39
	2	0.33	0.32
	3	0.34	0.34
	4	-	-
	5	0.48	0.31
	6	0.59	0.48

Table 4.17: General statistics of speciated Hg with different data treatment options in 2009.

Species	Case	Percent of missing values	MDL	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m^3)	1	31%	0.1	0%	1.37	1.41	1.39	0.26
	2	0%		0%	1.37	1.37	1.38	0.22
	3	0%		0%	1.38	1.41	1.39	0.22
GOM (pg/m^3)	1	32%	2	79%	0.27	0.42	1.77	3.98
	2	0%		86%	0.57	0.57	1.39	3.11
	3	0%		86%	0.51	0.42	1.34	3.12
	5	-		-	-	-	-	-
	6	32%		20%	3.91	3.35	5.02	4.76
PBM (pg/m^3)	1	41%	2	48%	1.79	2.15	2.81	2.71
	2	0%	2	70%	1.79	1.79	2.39	2.14
	3	0%	2	28%	1.93	2.15	2.53	2.11
	5	42%	4 (RM)	61%	2.73	3.02	4.69	5.19
	6	41%	2	4%	5.52	6.05	6.19	3.09

In Obs/Pred time series of Case 1 (Figure 4.1, the Obs/Pred time series of the other cases are provided in Appendix D), the graph of each mercury forms could be split into three time periods by the missing values. The three periods are January to February (period 1), March to July (period 2), and October to December (period 3). GEM had better performances than GOM and PBM because the peak values of GEM were reproduced well by the model in all three periods. However, the valley values for GEM in period 3 were too low compared to observed concentrations. The performance of PBM was better than GOM because the predicted PBM concentrations tracked the observed concentrations well in Period 2. However, PBM concentrations were underestimated in period 1 and overestimated in period 3. The observed GOM concentrations were poorly reproduced because the time series of predicted GOM concentrations is rather flat comparing to the observed GOM concentrations. Similar trends were observed in the other cases in 2009 except Case 6. In Case 6 (Figure D.3f), the predicted concentrations tracked the observed concentrations well for GEM and PBM in all three periods while the predicted concentrations tracked the observed concentrations well only in period 1 and period 3 for GOM. Increasing the low concentrations using a scaling factor improved the reproduction of the observed scaled concentrations by PMF model.

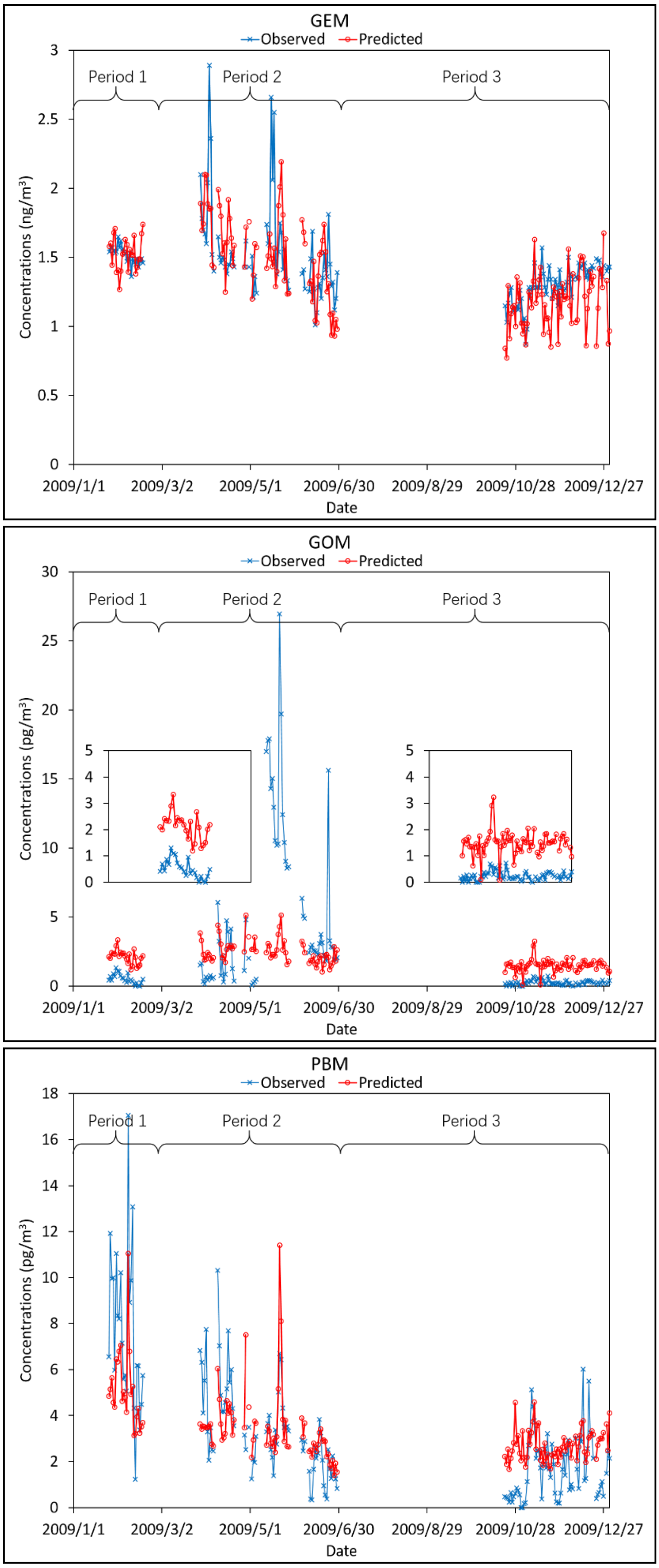

Figure 4.1: Obs/Pred time series for speciated Hg in 2009.

The ratio of predicted concentrations to observed concentrations (Pred/Obs ratio) is used to verify the results derived from Obs/Pred time series. In Case 1, the predicted GEM concentrations track the observed GEM concentrations well due to the narrow range of Pred/Obs ratios (Table 4.18, 0.57-1.32). The observed concentrations were also reproduced well on the annual basis because the ratio of annual mean concentration predicted to annual mean concentration observed (annual Predmean/Obsmean, 0.97, Table 4.18) is close to 1. The narrower range of the Pred/Obs ratio of PBM and the annual Predmean/Obsmean ratio for PBM closer to 1 compared to GOM (0.40-13.4 vs 0.13-53, 1.03 vs 0.86, respectively, Table 4.18) indicated that the PBM concentrations were better reproduced than GOM by PMF. Similar ratios were observed in other cases with an exception of Case 6 in 2009. In Case 6, the range of the Pred/Obs ratios of GOM and PBM were narrower (0-4.79 and 0-2.75, respectively, Table 4.18) compared to other cases indicating that the GOM and PBM were reproduced well by PMF. The performances derived from the Pred/Obs ratios are consistent with the performances derived from Obs/Pred time series.

Table 4.18: Ratios of PMF predicted to observed Hg concentrations of 2009

Case		1	2	3	4	5	6
GEM	Min	0.56	0.44	0.42	0.53	0.58	0.56
	Max	1.32	1.62	1.58	1.31	1.34	1.35
	Average	0.98	0.94	0.94	0.97	0.98	0.98
	Median	0.98	0.93	0.92	0.99	0.98	0.98
	Ratio of annual mean	0.97	0.93	0.93	0.97	0.97	0.97
GOM	Min	0.13	0.00	0.00	-	-	0.00
	Max	53.0	54.0	55.0	-	-	4.79
	Average	5.89	4.83	5.06	-	-	1.07
	Median	3.82	3.03	3.64	-	-	1.04
	Ratio of annual mean	0.86	1.19	1.20	-	-	0.75
PBM	Min	0.40	0.00	0.00	-	0.00 (RM)	0.00
	Max	13.4	17.3	17.6	-	17.0(RM)	2.75
	Average	2.09	1.93	1.86	-	2.29 (RM)	1.06
	Median	1.14	1.36	1.22	-	1.36 (RM)	0.90
	Ratio of annual mean	1.03	1.19	1.14	-	1.04 (RM)	0.94

Overall, too many imputations biased the source profiles and source contributions and made the model performance worse in 2009. Combining or excluding GOM and PBM from the input file had no impact on PMF model performances. The case using scaling factor to increase the GOM and PBM concentrations has better performances on GOM and PBM. However, increasing the low GOM and PBM concentrations using a scaling factor may contort the original variance of the dataset.

4.1.2 Year 2010

Case 7: Reference case

The source profiles of Case 7 are listed in Table 4.19. Case 7 is used as reference case in 2010. Factor 1 is similar to the factor named Combustion Emission in Case 1 in 2009 (Table 4.1). Factor 1 account for large % of NH_4^+ (87%), SO_4^{2-} (79%), HNO_3 (64%), and K^+ (51%). It is assigned to the same name because the SO_2 and NO_x, the precursors of the variables SO_4^{2-} and HNO_3, are both released by combustion process (Liu et al., 2007). The major variable K^+ may relate to the biomass combustion emission or the agriculture activities (Andersen et al., 2007). This is consistent with the presence of NH_4^+ in the major variables because NH_4^+ indicates the long-range transport of NH_3 from agriculture activities (Pitchford et al., 2009). NH_4^+ is formed by the reaction of NH_3 and HNO_3 or H_2SO_4. Therefore, this factor is assigned to the Combustion Emission.

Table 4.19: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 7 in 2010.

% of total species	F1	F2	F3	F4
GEM			79	
GOM		37	59	
PBM			81	
O_3			80	
SO_2		93		

Table 4.19 – Continued

% of total species	F1	F2	F3	F4
HNO_3	64	26		
Ca^{2+}		29	36	(21)
K^+	51		27	(23)
Mg^{2+}				83
Na^+				75
Cl^-				100
NO_3^-	(18)	41	(23)	(18)
NH_4^+	87			
SO_4^{2-}	79			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

The major variables of factor 2 are also similar to the major variables of the factor Industrial Source in Case 1 in 2009 (Table 4.1). The major variables of SO_2 (93%), HNO_3 (26%) and NO_3^- (41%) are the indicators of industrial sources and combustion sources (Huang et al., 2010; Liu et al., 2007). The factor accounted for moderate % of GOM (37%) instead of PBM compared to Case 1. It is not unexpected to observe the presence of GOM because the combustion process/industrial sources are sources of GOM (Carpi, 1997). This factor has an additional major variable of Ca^{2+} (29%) compared to Case 1. The presence of Ca^{2+} is not unexpected because $Ca(NO_3)_2$ could be formed through the reaction of soil aerosols and HNO_3 (Pakkanen, 1996; Zhang et al., 2008). This is consistent with the presence of NO_3^- in this factor. Because there were no combustion sources reported near the sampling site in 2010 (Table A.1), this factor is likely related to Industrial Source.

Factor 3 has the similar major variables as the factor named as Photochemistry and Re-emission of Hg in Case 1 (Table 4.2). This factor lacks the major variable of PM because PM data were not available in 2010. The lack of PM in the major variables did not affect the name of this factor. Factor 4 only has Cl^-, $Na^{+,}$ and Mg^{2+} as major variables. This factor was named as sea salt as well because the sea water contains a large amount of these three ions (Brennan et al., 2004).

The factor contributions of Case 7 are listed in Table 4.20. The factor Photochemistry and Re-emission of Hg has the largest contributions to mercury (79%, 67% and 80% for GEM, GOM and PBM, respectively). The factor Industrial Source has minor contributions to GOM (average, 29%) instead of PBM compared to Case 1 (Table 4.2). Other two factors Combustion Emission and Sea Salt only have minor contributions (<20%) to all three mercury forms.

Table 4.20: Factor contributions to speciated mercury in of Case 7 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	100	9	99	100
	Average	11	1	79	9
	Median	7	1	85	4
GOM (%)	Min	0	0	0	0
	Max	100	100	100	0
	Average	5	29	67	0
	Median	2	28	68	0
PBM (%)	Min	0	0	0	0
	Max	100	28	98	99
	Average	11	4	80	5
	Median	6	3	86	2

Case 8: Geometric mean imputation

The factor profiles of Case 8 are listed in Table 4.21. The major variables of the factors in Case 8 are similar to those in Case 7 (Table 4.19). Factor 1 has an additional major variable of NO_3^- (31%) compared to the factor Combustion Emission in Case 7. The presence of NO_3^- is reasonable because it has the same precursor as HNO_3 (NO_x) which could be release by combustion emission (Liu et al., 2007).

Table 4.21: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 8 in 2010.

% of total species	F1	F2	F3	F4
GEM			83	
GOM		34	63	
PBM			95	
O_3			84	
SO_2		95		
HNO_3	70	23		
Ca^{2+}			66	(17)
K^+	51		28	(21)
Mg^{2+}				81
Na^+				74
Cl^-				100
NO_3^-	31	(16)	35	(18)
NH_4^+	92			
SO_4^{2-}	83			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Factor 2 only has GOM (34%) and SO_2 (95%) as major variables which may not be enough to identify the specific factor. Therefore, the largest four variables (i.e. SO_2, GOM, HNO_3, and NO_3^-) were considered as the major variables of factor 2. Factor 2 lacks the presence of Ca^{2+} in the major variables compared to the factor Industrial Source in Case 7. However, this did not affect the name of the factor because the presence of HNO_3, SO_2, and NO_3^- is enough to identify the industrial source (see Case 1).

Factor 3 has an additional major variable of NO_3^- (35%) compared to the factor Photochemistry and Re-emission of Hg in Case 7 (Table 4.19). The factor is assigned to the same name because NO_3^- could also present in soil emission (K^+ and Ca^{2+} in factor 3, Table 4.21) (Parmar et al., 2001). Factor 4 has the same major variables as the factor Sea Salt in Case 7. Therefore, it is assigned to the same source.

The factor contributions to speciated mercury are listed in Table 4.22. The contributions of each factor to mercury are similar to those in Case 7 (Table 4.20).

This is because the missing values of mercury are only 4% and the mercury and other chemical species were missing at the same time. Therefore, the imputation did not affect the correlation coefficients between mercury and other chemical species a lot (changes <0.05 for significant correlations, Table 4.23).

Table 4.22: Factor contributions to speciated mercury in Case 8 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	29	70	99	100
	Average	3	6	83	8
	Median	2	3	89	4
GOM (%)	Min	0	0	0	0
	Max	32	99	100	0
	Average	2	27	70	0
	Median	2	21	76	0
PBM (%)	Min	0	0	0	0
	Max	0	84	100	98
	Average	0	4	93	2
	Median	0	2	97	1

Table 4.23: Pearson correlation coefficients between speciated mercury species and other chemical species in Case 7, Case 8 and Case 9 in 2010 (asterisk indicates insignificant at p =0.05).

	GEM			GOM			PBM		
	Case 7	Case 8	Case 9	Case 7	Case 8	Case 9	Case 7	Case 8	Case 9
GEM	1.00	1.00	1.00	0.29	0.29	0.29	0.19	0.19	0.19
GOM	0.29	0.29	0.29	1.00	1.00	1.00	0.14	0.15	0.15
PBM	0.19	0.19	0.19	0.14	0.15	0.15	1.00	1.00	1.00
O_3	0.71	0.68	0.68	0.54	0.51	0.51	0.30	0.29	0.29
SO_2	0.01*	0.00*	0.00*	0.29	0.29	0.29	0.05*	-0.04*	-0.03*
HNO_3	-0.12	-0.11	-0.11	0.23	0.24	0.24	0.04*	-0.04*	-0.04*
Ca^{2+}	-0.11*	-0.02*	-0.02*	0.07*	0.08*	0.09*	-0.17*	-0.05*	-0.05*
K^+	-0.13	-0.12	-0.13	0.07*	0.09*	0.10*	0.07*	-0.07*	-0.08*
Na^+	0.08*	0.07*	0.06*	-0.03*	0.00*	-0.01*	-0.07*	-0.10	-0.09*
Mg^{2+}	0.07*	0.06*	0.06*	-0.01*	0.02*	0.02*	-0.07*	-0.10*	-0.09*
Cl^-	0.09*	0.07*	0.07*	-0.10*	-0.07*	-0.07*	-0.07*	-0.11	-0.11
NO_3^-	-0.18*	-0.02*	-0.03*	0.18	0.19	0.20	0.03*	-0.03*	-0.04*
NH_4^+	-0.11*	-0.10*	-0.10	0.06*	0.08*	0.08*	0.01*	-0.04*	-0.04*
SO_4^{2-}	-0.11*	-0.10*	-0.10	0.06*	0.08*	0.08*	0.00*	-0.05*	-0.05*

Case 9: Median imputation

The factor profiles and factor contributions to speciated mercury of are listed in Table 4.24 and Table 4.25, respectively. The major variables of all four factors are the same as the major variables of the factors in Case 7 (Table 4.19), respectively. Therefore, all factors are assigned to the same sources. The factor contributions to speciated mercury are also similar to those in Case 7 (Table 4.20). The similar results between this case and Case 7 are because only a small amount of Hg concentrations are missing and the Hg species and the non-Hg species were missing at the same time (Table 4.23). Therefore, the correlation coefficients between mercury and the other chemical species changed little between Case 9 and Case 7 (Table 4.23).

Table 4.24. Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 9 in 2010.

% of total species	F1	F2	F3	F4
GEM			88	
GOM		44	56	
PBM			98	
O_3			89	
SO_2		97		
HNO_3	61	28		
Ca^{2+}		30	47	(17)
K^+	48		30	(22)
Mg^{2+}				80
Na^+				72
Cl^-				100
NO_3^-		46	(23)	(17)
NH_4^+	87			
SO_4^{2-}	79			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Table 4.25: Factor contributions to speciated mercury in Case 9 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	39	1	100	100
	Average	3	0	88	8
	Median	2	0	93	4
GOM (%)	Min	0	0	0	0
	Max	5	100	100	0
	Average	0	36	64	0
	Median	0	36	63	0

Table 4.25 – Continued

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
PBM (%)	Min	0	0	0	0
	Max	0	0	100	100
	Average	0	0	97	3
	Median	0	0	99	1

The number of missing values in mercury was relatively small (4%, Table 3.2) and the mercury and non-mercury species were missing at the same time. This led to the similar correlation coefficients between the reference case and the imputation cases. Therefore, it is not unexpected that the imputation cases had similar results as the reference case.

Case 10: Excluding GOM and PBM

The source profiles and source contributions to speciated mercury of Case 10 are listed in Table 4.26 and Table 4.27, respectively. The major variables of all four factors are similar to those in Case 7 (Table 4.19). Although GOM and PBM are removed from the input files, all the factors are assigned to the same sources because the rest of the variables did not change from Case 7 and they were enough to identify the factors (Table 4.19 and Table 4.26). The factor contributions to GEM are also similar to those of Case 7 (Table 4.20). Only the factor Photochemistry and Re-emission of Hg has a large contribution (78%) to GEM. The factor Industrial Source is no longer contributing to GOM due to the exclusion of GOM. Overall, the exclusion of GOM and PBM have few impacts on the source profiles but it concealed the factor contributions to GOM or PBM.

Table 4.26: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 10 in 2010.

% of total species	F1	F2	F3	F4
GEM			79	
O_3			79	
SO_2		94		
HNO_3	64	26		
Ca^{2+}		30	36	(21)
K^+	51		27	(22)
Mg^{2+}				83
Na^+				75
Cl^-				100
NO_3^-	(18)	41	(23)	(18)
NH_4^+	87			
SO_4^{2-}	79			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea salt

Table 4.27: Factor contributions to GEM in Case 10 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	100	10	99	100
	Average	11	1	78	9
	Median	7	1	85	4

Case 11: Combining GOM and PBM

The factor profiles of Case 11 are listed in Table 4.28. In this case, the major variables of all four factors were similar to the factors in Case 7 (Table 4.19). However, the major variable GOM was removed while the major variable PBM was replaced by RM. This is because the variance of RM is dominated by the variance of PBM due to the higher PBM concentrations than GOM concentrations (i.e. the PBM median concentration is approximately 10 time of the GOM median concentrations in 2010, Table 3.2). The substitution of GOM and PBM with RM did not affect the identification of the factors because other major variables in the factors remained the same. Therefore, all four factors are assigned to the same names as the factors in Case 7.

Table 4.28: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 11 in 2010.

% of total species	F1	F2	F3	F4
GEM			79	
RM			81	
O_3			80	
SO_2		94		
HNO_3	64	26		
Ca^{2+}		30	36	(21)
K^+	51		27	(23)
Mg^{2+}				83
Na^+				75
Cl^-				100
NO_3^-	(18)	41	(23)	(18)
NH_4^+	87			
SO_4^{2-}	79			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

The factor contributions to speciated mercury of Case 11 are listed in Table 4.29. The factor contributions are similar to the factor contributions to GEM and PBM in Case 7, respectively, because the variance of RM is dominated by the variance of PBM. The factor Photochemistry and Re-emission of Hg has the largest contributions to GEM and RM while the other three factors only have small factor contributions (<20%) to GEM and RM. Therefore, combining GOM and PBM to RM does not affect the factor identifications but it concealed the factors' contribution to GOM species.

Table 4.29: Factor contributions to speciated mercury in Case 11 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	100	9	99	100
	Average	11	1	79	9
	Median	7	1	85	4
RM (%)	Min	0	0	0	0
	Max	100	47	98	96
	Average	9	8	80	3
	Median	5	6	86	1

Case 12: Scaling factor

The factor profiles and factor contributions to speciated mercury of Case 12 are listed in Table 4.30 and Table 4.31, respectively. Factor 1, factor 3 and factor 4 have the same major variables as the factors Combustion Emission, Photochemistry and Re-emission of Hg, and Sea Salt in Case 7 (Table 4.19), respectively. Therefore, they are assigned to the same names. Factor 2 lacks GOM in the major variables compared to the factor Industrial Source in Case 7 (Table 4.19). However, this factor has the same major variables as the factor Industrial Source in Case 6 in 2009 (Table 4.12). Therefore, this factor is also assigned to Industrial Source. The factor contributions to speciated mercury in this case are similar to those of Case 7 (Table 4.20). The results of this case indicate that increasing the GOM and PBM concentrations by a scaling factor did not affect the source identification. However, the source contributions to mercury had a minor change.

Table 4.30: Factor profiles (% of species >25%, between 15% and 25% in bracket) of Case 12 in 2010.

% of total species	F1	F2	F3	F4
GEM			78	
GOM (scaled)		(16)	75	
PBM (scaled)			76	
O_3			80	
SO_2		91		
HNO_3	64	(25)		
Ca^{2+}		30	34	(22)
K^+	50		27	(22)
Mg^{2+}				83
Na^+				75
Cl^-				100
NO_3^-	(18)	40	(23)	(18)
NH_4^+	87			
SO_4^{2-}	79			
Factor name	Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt

Table 4.31: Factor contributions to speciated mercury in Case 12 in 2010.

Factor name		Combustion Emission	Industrial Source	Photochemistry & Re-emission of Hg	Sea Salt
GEM (%)	Min	0	0	0	0
	Max	100	10	99	100
	Average	11	1	78	9
	Median	7	1	85	4
GOM (scale) (%)	Min	0	0	0	0
	Max	100	69	99	75
	Average	8	14	77	1
	Median	4	11	80	0
PBM (scale) (%)	Min	0	0	0	0
	Max	100	40	97	98
	Average	12	6	76	6
	Median	7	4	82	2

As seen in Table 4.14, factors Photochemistry and Re-emission and of Hg and Sea Salt have the largest and smallest contributions to speciated mercury in 2010, respectively. The factor Industrial Source has the second largest contributions to GOM in Case 7, Case 8, Case 9, and Case 12. The combination of GOM and PBM into RM and exclusion of GOM and PBM makes the factor Industrial Source only has small contributions to mercury in Case 10 and Case 11 because the variance of RM is dominated by PBM.

Performances

The distribution of scaled residuals and the number of the scaled residuals larger than 3 for speciated mercury in 2010 are listed in Table 4.32. As seen in Table 4.32, the model has the best performance on reproducing the observed GEM concentrations in 2010 because the distributions of scaled residuals are normal in all cases and nearly no scaled residuals are larger than 3. The GOM concentrations are better reproduced than the PBM concentrations in 2010 according to the scaled residual plot because no GOM scaled residuals are larger than 3. The narrow distribution of GOM was caused by the large uncertainties due to the large amount

(96%, Table 3.2) of below MDL observations (Polissar et al., 1998). Similar distributions and number of scaled residuals were observed in Case 7, Case 8, Case 9, Case 10 and Case 11. This indicates that small amount of imputation, combining or excluding GOM and PBM do not affect the model performances on reproducing the GOM and PBM concentrations very much. In Case 12, the PBM performance was improved (normal distribution). This indicates that using a scaling factor may improve the model performances on reproducing the GOM and PBM concentrations.

Table 4.32: PMF model performances for speciated mercury in scaled residual plot in 2010.

Species	Case number	Criteria	
		Normal distribution	Number of scaled residuals larger than 3
GEM	7	Normal	2
	8	Normal	19
	9	Normal	2
	10	Normal	2
	11	Normal	2
	12	Normal	1
GOM	7	Narrower	0
	8	Narrower	0
	9	Narrower	0
	10	-	-
	11	-	-
	12	Narrower	0
PBM	7	Right skewed	14
	8	Right skewed	28
	9	Right skewed	29
	10	-	-
	11	Right skewed (RM)	5
	12	Normal	18

The R^2 value and the slope of the regression line for speciated mercury in Obs/Pred scatter plot in 2010 are listed in Table 4.33. GEM has the best performance in both R^2 and the slope of the regression line among all three mercury forms. Case 8 and Case 9 have similar performances on R^2 and the slope of the regression line with the reference case in 2010. The similar performances on GEM observed in Case 10 and Case 11 indicates that excluding or combining GOM and PBM did not cause many differences in the model performances on reproducing the GEM concentrations in 2010. The performances of RM are similar to that of PBM because the variance of RM is dominated by the variance of PBM due to its higher concentrations compared to GOM (i.e. the median concentration of PBM is 10 times of GOM in 2010, Table 3.2). Case 12 has the best performances on reproducing GOM, and PBM observations among all the cases in 2010 but similar performances on reproducing GEM compared to Case 1. This indicates that increasing GOM and PBM concentrations by a scaling factor improved the model performances because the uncertainty of the concentrations was decreased as the number of below MDL value decreased (Table 4.34).

Table 4.33: PMF model performances for speciated mercury in Obs/Pred scatter plot in 2010 (all significant at $p < 0.001$).

Species	Case number	Criteria	
		Coefficient of determination (R^2)	Slope of regression line
GEM	7	0.46	1.29
	8	0.32	1.26
	9	0.41	1.26
	10	0.47	1.31
	11	0.46	1.31
	12	0.44	1.19
GOM	7	0.31	0.29
	8	0.23	0.22
	9	0.28	0.28
	10	-	-
	11	-	-
	12	0.42	0.33
PBM	7	0.13	0.09
	8	0.15	0.09
	9	0.16	0.08
	10	-	-
	11	0.19	0.15
	12	0.25	0.24

Table 4.34: General statistics of speciated Hg with different data treatment options in 2010.

Species	Case	Percent of missing values	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m^3)	7	4%	0%	1.33	1.37	1.34	0.17
	8	0%	0%	1.34	1.37	1.35	0.16
	9	0%	0%	1.34	1.38	1.35	0.16
	11	4%	0%	1.33	1.37	1.34	0.17
	12	4%	0%	1.33	1.38	1.34	0.17

Table 4.34 – Continued

Species	Case	Percent of missing values	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation
GOM (pg/m^3)	7	4%	96%	0.27	0.21	0.44	0.64
	8	0%	96%	0.27	0.24	0.43	0.63
	9	0%	96%	0.27	0.21	0.43	0.63
	11	-	-	-	-	-	-
	12	4%	85%	1.12	0.99	1.12	0.90
PBM (pg/m^3)	7	4%	51%	2.08	2.2	3.4	4.13
	8	0%	44%	2.08	2.12	3.35	4.04
	9	0%	44%	2.08	2.20	3.35	4.04
	11	4%	70%	2.45	2.62	3.85	4.27
	12	4%	1%	7.55	7.77	8.55	4.52

In Case 7 (Figure 4.2, the time series of all cases in 2010are provided in Appendix D), the time series could be split into two periods: January-May (period 1) and June-December (period 2). In period 1, GEM concentrations were overestimated while the GEM concentrations are underestimated in period 2. The fluctuation of the predicted time series is stronger than the observed time series for GEM. The predicted GOM concentrations track the observed GOM concentrations well in period 1 while the predicted GOM concentrations are overestimated in period 2. The predicted PBM concentrations and observed PBM concentrations did not track with each other in both periods. The peak of the observed PBM time series was not reproduced by the model either. In 2010, the time series of Case 8, Case 9, and Case 12 (Figure D.8, Figure D.9, and Figure D.12) have different trrends from Case 7 . In Case 8 and Case 9, the PBM concentrations were not reproduced well at the beginning of 2010 because the imputation brought back some peak PBM concentrations (Table 4.33). Similar to 2009, the observed time series and the predicted time series in the case increasing GOM and PBM with a scaling factor tracked with each other. This indicated a better reproduction of the observed concentrations for GOM and PBM species due to the reduced number of concentrations below MDL and the reduction in corresponding uncertainties.

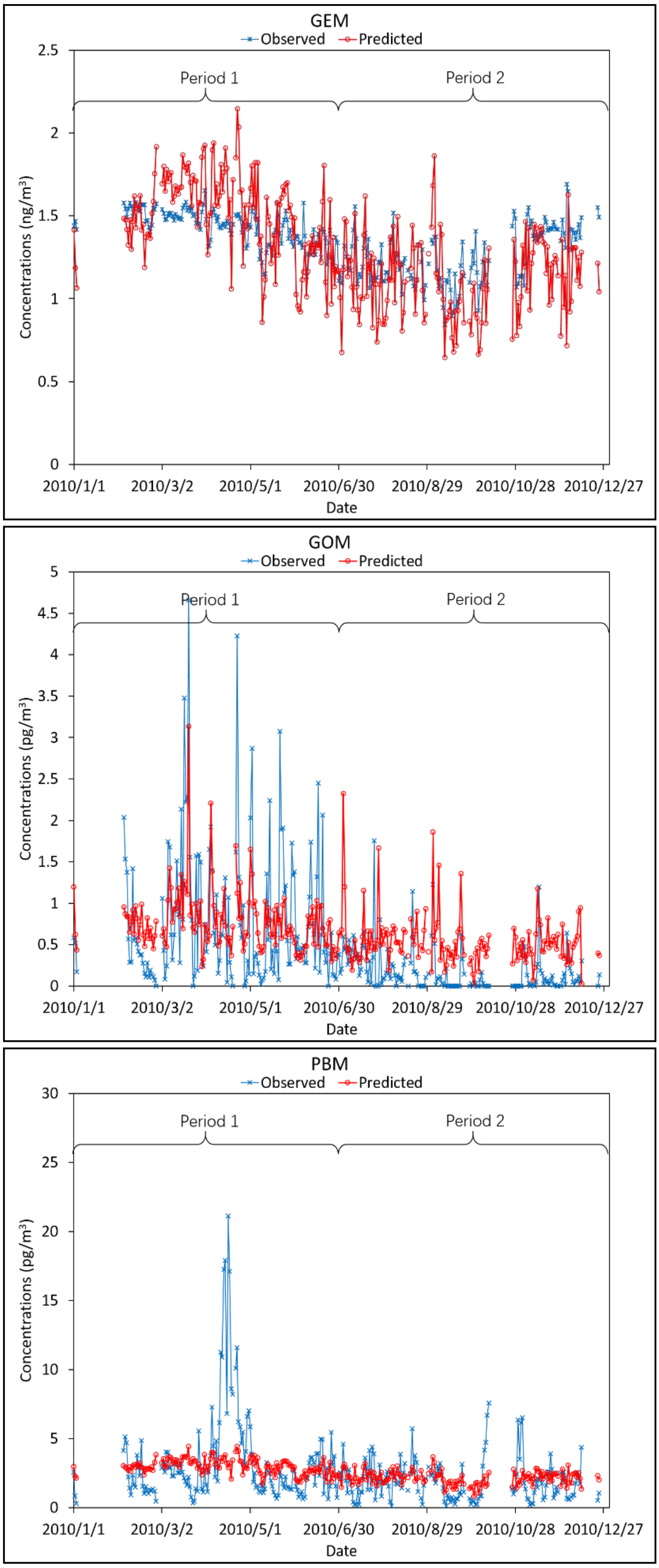

Figure 4.2: Obs/Pred time series for speciated Hg in 2010.

The Pred/Obs ratios of speciated mercury in Case 7 were also analyzed. The predicted concentrations of GEM track the observed concentrations well because the range of GEM Pred/Obs ratios (Table 4.35, 0.42-1.43) is short and the annual Predmean/Obsmean is close to 1 (Table 4.35, 0.98). The reproduction of GOM is worse than that of PBM due to a larger range of Pred/Obs ratios (GOM: 0.19-193; PBM: 0.14-18.33, Table 4.35) and the closer to 1 annual Predmean/Obsmean ratio (1 vs 1.34, Table 4.35). Similar ratios for GEM were observed in other cases in 2010. In Case 8 and Case 9, the ranges of Obs/Pred ratios for PBM were slightly larger than the reference case while the annual Predmean/Obsmean ratios were further from 1 (Table 4.35) indicating that the model reproduction of PBM was poor after imputation. This is because the imputation of PBM brought the peak PBM concentration back. In Case 12, the range of Obs/Pred ratio for GOM and PBM are smaller compared to the base case indicating increasing the GOM and PBM concentrations improved the model performance on reproducing the observed concentrations. Overall, the results derived from the ratios are consistent with the results from the time series (Appendix D).

Table 4.35: Ratios of PMF predicted to observed Hg concentrations in 2010

Case		7	8	9	10	11	12
GEM	Min	0.42	0.20	0.22	0.42	0.42	0.51
	Max	1.43	1.49	1.48	1.39	1.41	1.44
	Average	0.97	0.95	0.97	0.97	0.97	0.97
	Median	0.97	0.99	0.98	0.97	0.97	0.97
	Ratio of annual mean	0.98	0.96	0.97	0.98	0.98	0.98
GOM	Min	0.00	0.00	0	-	-	0.00
	Max	193	141	196	-	-	10.6
	Average	4.44	3.62	3.66	-	-	1.27
	Median	1.48	1.47	1.85	-	-	1.08
	Ratio of annual mean	1.34	1.35	1.32	-	-	1.23

Table 4.35 – Continued

Case		7	8	9	10	11	12
PBM	Min	0.14	0.12	0.14	-	0.18 (RM)	0.32
	Max	18.3	22.8	21.5	-	16.2 (RM)	3.09
	Average	1.98	2.03	2.09	-	2.19 (RM)	1.01
	Median	1.37	1.34	1.37	-	1.51 (RM)	0.96
	Ratio of annual mean	1.00	0.87	0.88	-	1.16 (RM)	0.88

4.1.3 Comparison between 2009 and 2010

In terms of factor contributions, factors Photochemistry and Re-emission of Hg and Sea Salt have the largest and smallest mercury contributions in both years, respectively. The factors Combustion Emission and Industrial Source are the second and third largest contributors to mercury in 2009 while only the factor Industrial Source has a large mercury contribution in 2010. The lack of the contribution to mercury in factor Combustion Emission in 2010 is probably due to the large reduction of SO_2 (2,425,000 tons, 32%) and NO_2 (894,000 tons, 32%) by reducing the coal combustion in the United States from 2008 to 2010 (U.S. EPA, 2011) and the large reduction of SO_2 emission (38265 tonnes, 35%) and Hg emission (57.2 kg, 39%) between 2009 and 2010 in Nova Scotia Table A.1). One study also shows that the shutdown of the coal-fired power plant reduces the correlations between GOM and SO_2 (Huang et al., 2010). After replacing the missing values with geometric mean or median, factor contributions to mercury changed in 2009 while the factor contributions to mercury in 2010 were similar to the reference case. This is likely due to the larger number of missing values in Hg in 2009 (up to 41%, Table 3.1) than in 2010 (4%, Table 3.2). Excluding or combining the GOM and PBM to RM and scaling GOM and PBM did not affect the factor profiles or factor contributions in both years.

The differences of the factor profiles in 2009 and 2010 were consistent with the differences between factor contributions.

The interpretability of the factor Industrial Source in 2009 was slightly worse than in 2010. The major variables of the factor Industrial Source in 2009 only contain SO_2 and PBM. However, PBM and SO_2 have many common sources such as coal combustion, industrial sulfur and wildfires (Wang et al., 2010, Huang et al., 2011). The factor was confirmed by considering the third and fourth largest species (NO_3^- and HNO_3, respectively) the factor accounting for in 2009. However, the major variables of the factor Industrial Source in 2010 were SO_2, GOM, HNO_3, Ca^{2+}, and NO_3^-. The factor Industrial Source in 2010 was identified without including additional variables. Therefore, the interpretability of the factor Industrial Source was better in 2010 compared to 2009. The interpretability of other three factors was similar between 2009 and 2010.

The overall model performances in 2010 are better than that in 2009 based on the reference cases. The cases with imputations had worse performance on reproducing GOM and PBM concentrations in 2009 but similar performances on reproducing GOM and PBM concentrations in 2010. This is probably due to the smaller amount of imputations in 2010 (up to 41% in 2009 vs 4% in 2010, Table 3.1 and Table 3.2 respectively). The cases excluding or combining GOM and PBM to RM had similar performances on reproducing GEM concentrations to the reference case in both years. The cases increasing the low GOM and PBM concentrations use scaling factor improved the model performances on reproducing mercury concentration in both years.

4.1.4 Sensitivity of the PMF results to data treatments

Different treatments including geometric mean imputation, median imputation, excluding GOM and PBM, combining GOM and PBM to RM, and increasing the low GOM and PBM concentrations by the scaling factor was adopted to improve the data quality. Overall, a large number of imputations affected the factor profiles and the factor contributions to mercury in 2009. However, it did not affect the identification of the factors because major variables of the factors did not change a lot after the imputations in both years.

Similar to the imputation cases, excluding or combining GOM and PBM did not affect the source identification in PMF model in both years in this study (Table 4.36). However, the identification of the factors relying on GOM or PBM only (i.e. mercury condensation process) may be affected after combining or excluding GOM and PBM, theoretically, but this kind of factors did not exist in this study. Excluding or combining GOM and PBM did affect the source contributions in both years. For example, the factors contributing to GOM only (Combustion Emission, 2009; Industrial Source 2010, Table 4.36) did not contribute to mercury after combining or excluding GOM and PBM. The factor (Industrial source, 2009, Table 4.36) contributing to PBM only is contributing to RM after the combination of GOM and PBM because the variance of RM is dominated by the variance of PBM in 2009. Using speciated mercury led to more major mercury sources identified. Therefore, monitoring speciated mercury could help us understanding the mercury cycle better. The cases increasing the factors using scaling factors have similar factor profiles and factor contributions as the reference cases.

Table 4.36: Comparison of the major mercury factors (>15% contribution) between the reference case and the cases combining or excluding GOM and PBM.

Year	Case No.	Combustion Emission	Industrial Source	Photochemistry and Re-emission of Hg	Sea Salt
2009	Case 1	GOM	PBM	GEM, GOM, and PBM	
	Case 4			GEM	
	Case 5		RM	GEM and RM	
2010	Case 7		GOM	GEM, GOM, and PBM	
	Case 10			GEM	
	Case 11			GEM and RM	

As seen in Figure 4.3, the ratio of predicted concentrations to observed concentrations indicates the model performance on reproducing the observed concentrations. The model performance on reproducing GEM did not change a lot using the different treatments to improve the data quality in both years because GEM only had fewer missing values (31% in 2009, Table 3.1, and 4% in 2010, Table 3.2) and no values below MDL (0%, Table 3.1 and Table 3.2), compared to GOM and PBM. The reproduction of the GOM and PBM were improved in both years using different treatment of input data with an exception for PBM in Case 8 and Case 9 in 2010. The reproduction of the GOM and PBM concentrations in Case 8 and Case 9 were similar to Case 7 due to the small amount (4%, Table 3.2) of missing values in mercury. Increasing the low GOM and PBM concentrations improved the model performance on reproducing the observed GOM and PBM concentrations most in both years. This is likely due to the large reduction of the values below MDL by increasing the low concentrations (57% reduction for GOM and 33% reduction for PBM in 2009, Table 4.17; 29% reduction for GOM and 50% reduction for PBM in 2010, Table 4.34).

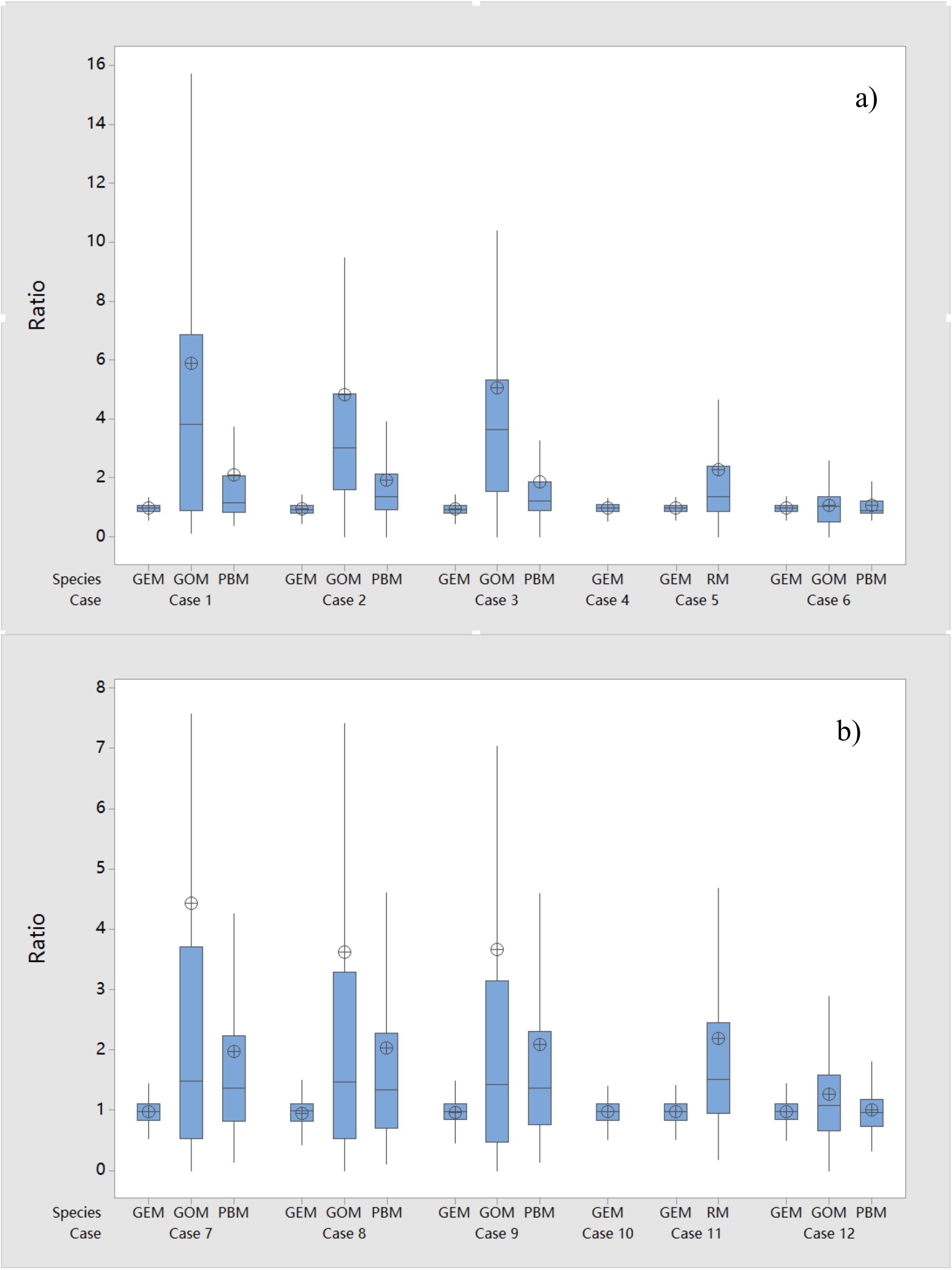

Figure 4.3. Box plot of the predicted to observed concentrations ratios (upper whisker- the upper 25% of the distribution excluding outliers; interquartile range box - middle 50% of the data; the horizontal line in the box: the median of the data; lower whisker- the lower 25% of the distribution excluding outliers; ⊕ - the average of the data) a) 2009, b) 2010.

4.2 PCA results

4.2.1 Year 2009

Case 13: Reference Case

The variables with loadings larger than 0.25 in PCA results were considered as major variables of the component. PCA results with all loadings could be found in Appendix E. The component loadings of the major variables in Case 13 are listed in Table 4.37. PC1 is named as Combustion/industrial Emission due to the high positive loadings (>0.25) of O_3, SO_2, Ca^{2+}, HNO_3, K^+, NH_4^+, NO_3^-, and SO_4^{2-}. The positive loadings of these species indicate that the concentrations of these species increase or decrease together. These chemical species were also found in the factor related to industrial emissions in another study using the same dataset (Cheng et al., 2013). The presence of HNO_3, NO_3^-, and SO_4^{2-} in the major variables indicates the transport of combustion/industrial emission. This is because their precursors ($NO_{x,}$ and SO_2) are mainly emitted by combustion/industrial process (Liu et al., 2007). The precursors may be oxidized during the transport process. The high positive loading of NH_4^+ is related to the NH_3 released by the local or regional agriculture activity and livestock and it could react with HNO_3 or H_2SO_4 to form NH_4^+, NO_3^-, and SO_4^{2-} (Pakkanen, 1996; Pitchford et al., 2009). The moderate loading of ozone is also related to the transport of combustion emission because the precursors of O_3 (NO_x and VOC) are mainly released by the combustion processes from mobile sources and stationary sources (New Jersey Department of Environmental Protection (NJDEP), 2011). The high loading of PBM indicates the coal combustion as well (Huang et al., 2010). The moderate loading of Ca^{2+} and minor loading of K^+ are related to the soil emission or the biomass burning (Zhang et al., 2008; Andersen et al, 2007). More markers such as levoglucosan are needed to verify the biomass combustion process (Puxbaum et al., 2007). It is possible for one factor to representing two sources because these two sources could affect the site at the same time (Cheng et al., 2013). Therefore, this factor is mainly related to Combustion/industrial Emission.

Table 4.37: Component loadings (>0.25) of Case 13 in 2009.

Factor Number	PC1	PC2	PC3	PC4
GEM			0.86	0.27
GOM			0.26	0.84
PBM	0.63		0.50	-0.33
PM	0.80			
O_3	0.50		0.70	
SO_2	0.88			
HNO_3	0.86			0.34
Ca^{2+}	0.59	0.39		0.45
K^+	0.29	0.70		0.33
Na^+		0.97		
Mg^{2+}		0.95		
Cl^-		0.97		
NO_3^-	0.73	0.48		
NH_4^+	0.92			
SO_4^{2-}	0.86			
Factor name	Combustion/industrial Emission	Sea Salt	Gas Phase Oxidation of Hg	Gas-particle Partition of Hg
Variance explained	37%	25%	11%	9%

The loadings of GEM and GOM were low in PC1 although all three speciated mercury could be released by coal combustion (UNEP, 2002). This is because the increase of GOM and PBM concentrations is more easily to be observed due to their significant lower background concentrations compared to GEM (Huang et al., 2010). However, GOM was removed rapidly after its formation in the atmosphere and led to the low loading on GOM (Stamenkovic et al., 2007). Therefore, it is reasonable that only PBM is present in this factor.

PC2 has high loadings of Cl^- (0.97), Mg^{2+} (0.95), K^+ (0.70), and Na^+ (0.97) and moderate loadings of Ca^{2+} (0.39) and NO_3^- (0.48) (Table 4.37). The presence of Cl^-, Mg^{2+}, K^+, Na^{2+}, and Ca^+ indicate the marine aerosols because all these ions are rich in sea water (Brennan et al., 2004). The loading of NO_3^- in this factor is probably due to the reaction of HNO_3 and sea salt (Pakkanen, 1996). The presence of sea salt at this site is not unexpected because the site is near the Atlantic. Therefore, this factor is named as Sea Salt.

PC3 is named as Gas Phase Oxidation of Hg. This factor has positive loadings of GEM (0.86), GOM (0.26), PBM (0.50) and O_3 (0.70) (Table 4.37). The positive loadings on O_3 and GOM indicate the photochemical production of GOM (Huang et al., 2010). However, the positive loading of GEM is not expected because the photochemical production of GOM consumes GEM and will lead to different signs of GEM and GOM like in Huang et al. (2010)'s study. It should be noted that the two-hour average concentrations were used in Huang et al. (2010)' study while the daily average concentrations were used in this study. The daily GEM and GOM concentrations in this study are positively correlated (r=0.37 in 2009, Table 4.5, r=0.29 in 2010, Table 4.23). In another PCA study (Cheng et al., 2013) using the same dataset, a further analysis on %GOM/TGM ratios (TGM=GEM+GOM) was conducted. The ratio is indicative of the degree of oxidation. In the analysis, the ratio increased with O_3 when O_3 concentration is over 40 ppb which indicates the oxidation of GEM by ozone existed at this site (Cheng et al., 2013).

PC4 is named as Gas-particle Partition of Hg. The negative loading of PBM (-0.33, Table 4.37) and the positive loadings of other two mercury forms (0.27 and 0.84 for GEM and GOM, respectively, Table 4.37) indicates the partition process (i.e. the PBM concentrations increase as the GOM concentrations decrease). The positive loadings of Ca^{2+} and K^+ represent soil aerosols (Andersen et al., 2007; Zhang et al., 2008). The soil aerosols are abundant at the site because it is located in a national park. The soil aerosols in this factor were the particles for partitioning. Therefore, this factor is named as Gas-particle Partitioning of Hg.

Three out of four principal components have an impact on the ambient mercury concentrations and only the principal component Sea Salt has little impact on ambient mercury concentrations. The principal component Gas-particle Partition of Hg was the additional factor identified by PCA compared to PMF results using the same dataset. This is because the identification of partitioning process depends on the negative correlation between GOM and PBM. However, this kind of relationship cannot be identified by PMF model due to its non-negative property.

Case 14: Including meteorological parameters

Five principal components are extracted in Case 14. The component loadings of Case 14 are listed in Table 4.38. The loadings of chemical species of PC1, PC2, PC3, and PC4 are similar to the principal components Combustion/industrial Emission, Sea Salt, Gas-particle Partitioning of Hg, and Gas Phase Oxidation of Hg in Case 13, respectively. Each of these of factors has an additional loading of meteorological parameters but the meteorological parameters did not affect the identification of the component. PC1 had an additional negative loading on relative humidity (-0.26) while PC2 had an additional positive loading on wind speed (0.32) compared to Case 13. The loadings on meteorological parameters in these two principal components are relatively low compared to other major variables. Therefore, they had little impact on the component identification.

Table 4.38: Component loadings (>0.25) of Case 14 in 2009.

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM				0.80	
GOM			0.64	0.41	-0.29
PBM	0.59		-0.47	0.34	
PM	0.81				
O_3	0.47			0.72	-0.27
SO_2	0.86				
HNO_3	0.88				

Table 4.38 – Continued

Factor Number	PC1	PC2	PC3	PC4	PC5
Ca^{2+}	0.60	0.38	0.33		
K^+	0.36	0.66	0.39		
Na^+		0.96			
Mg^{2+}	0.28	0.95			
Cl^-		0.98			
NO_3^-	0.76	0.45			
NH_4^+	0.94				
SO_4^{2-}	0.88				
Temperature			0.94		
Relative humidity	-0.26				0.79
Wind speed		0.32		0.52	0.49
Precipitation					0.79
Factor name	Combustion/ industrial Source	Sea Salt	Gas-particle Partitioning of Hg	Gas Phase Oxidation of Hg	Hg Wet Deposition
Variance explained	30%	20%	10%	10%	10%

PC3 lacks the presence of GEM and has an additional loading of temperature (0.94, Table 4.38). The lack of GEM did not affect the identification of this factor because the partitioning of GEM from the gas phase to particles is much weaker than GOM (Liu et al., 2007). The negative relation between temperature and PBM is consistent with the condensation process because the low temperature is in favor of the formation of PBM (Rutter & Schauer, 2007). PC4 has an additional loading of wind speed indicating that the air mass containing mercury and/or O_3 is transported from the urban or industrial area (Cheng et al., 2013). This is reasonable because mercury emissions and the sources of other pollutants in 2009 were located in the north of the site (Figure 3.1) where the temperature should be lower.

PC5 is named as Hg Wet Deposition due to the negative loading of GOM (-0.29, Table 4.38) and positive loadings of relative humidity (0.79, Table 4.38) and precipitation (0.79, Table 4.38). The positive loadings of precipitation and relative humidity indicates the precipitation process (Huang et al., 2010). The negative loading on GOM is consistent with precipitation because GOM is easily removed by precipitation process compared to GEM due to its higher water solubility (Gaffney & Marley, 2014). Therefore, this factor is named as Mercury Wet Deposition.

Similar to Case 13, all factors except the factor Sea Salt had an impact on the ambient mercury concentrations according to loadings on mercury. After including the meteorological parameters, an additional component Mercury Wet Deposition was identified. Each of the factors contains, at least, one meteorological parameter in major variables but the meteorological parameter did not play a critical role in factor identification in all factors except Mercury Wet Deposition. The loadings of chemical species in each of the components did not change a lot compared to Case 13. Therefore, similar components were identified in this Case.

4.2.2 Year 2010

Case 15: Reference Case

The component loadings are listed in Table 4.39. PC1 is named as Combustion Emission due to the positive loadings of HNO_3 (0.34), NO_3^- (0.79) and SO_4^{2-} (0.90). This is because their precursors (NO_2 and SO_2) are good indicators of combustion emissions (Liu et al., 2007). They might be oxidized during the transport to the sampling site because there were no combustion/industrial emissions near the site (Table A.1). The high positive loading of NH_4^+ (0.94) may relate to emission of NH_3 resulted from the excess use of fertilizer in agriculture activities and the reaction with HNO_3 or H_2SO_4 during the transport (Pakkanen, 1996; Pitchford et al., 2009). The positive loadings of Ca^{2+} (0.89) and K^+ (0.77) indicate soil or biomass combustion (Andersen et al., 2007; Zhang et al., 2008). Therefore, this component is named as Combustion Emissions.

Table 4.39: Component loadings (>0.25) of Case 15 in 2010.

Factor Number	PC1	PC2	PC3	PC4
GEM			0.79	
GOM			0.71	0.33
PBM			0.48	
O_3			0.91	
SO_2				0.89
HNO_3	0.34			0.83
Ca^{2+}	0.89			
K^+	0.77			
Na^+		0.99		
Mg^{2+}	0.34	0.93		
Cl^-		0.98		
NO_3^-	0.79			
NH_4^+	0.94			
SO_4^{2-}	0.90			0.26
Factor name	Combustion emission	Sea salt	Gas phase oxidation of mercury	Industrial source
Variance explained	28%	21%	16%	13%

PC2 is named as Sea Salt due to high positive loadings of Na^+ (0.99), Mg^{2+} (0.93), and Cl^- (0.98). These three chemical species are rich in the sea water (Brennan et al., 2004). PC3 has the same major variables as the component Gas Phase Oxidation of Hg in Case 13. Therefore, PC3 is also named as Gas Phase Oxidation of Hg.

PC4 is assigned to Industrial Source. The positive loadings of GOM (0.33) and SO_2 (0.89) indicate coal combustion process (Lynam & Keeler, 2006). The positive loading of SO_4^{2-} (0.26) is consistent with the combustion process because SO_2, the precursor of SO_4^{2-}, is released by combustion process (Liu et al., 2007). However, no combustion process was reported near the KEJ site in 2010 according to the NPRI (Table A.1). The positive loading of HNO_3 indicates industrial sources because NO_2, the precursor of HNO_3 is mainly released by industrial sources (Liu et al., 2007). Therefore, this factor is more likely to relate to industrial sources. However, this factor should be verified by other analyses such as back trajectory analysis.

Only two factors (i.e. Gas Phase Oxidation of Hg and Industrial Source) have an impact on ambient mercury. The factor Combustion Emission is no longer contributing to mercury in 2010 compared to Case 13 (Table 4.37). This was probably caused by the reduction of coal combustion in the USA (US EPA, 2011) and the reduction of SO_2 and Hg emission in Canada (Table A.1). However, a long term study should be conducted to evaluated the contributions of the reduction of SO_2, NO_x and Hg emission in the USA and Canada to the ambient Hg concentrations at this site.

Case 16: Including meteorological parameters

The component loadings of Case 16 are listed in Table 4.40. The loadings of the chemical species in PC1, PC2, PC3 and PC5 are similar to the factors Combustion Emission, Sea Salt, Gas Phase Oxidation of Hg and Industrial Source in Case 15, respectively. Each of the factors has additional loadings on meteorological parameters. The additional loadings on temperature (0.27) and wind speed (0.26) in PC1 and PC2 are relatively low. They have little impact on the factor identification. Therefore, PC1 and PC2 were named as Combustion Emission and Sea Salt, respectively.

Table 4.40: Component loadings (>0.25) of Case 16 in 2010.

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM			0.87		
GOM			0.51	-0.51	0.38
PBM			0.29	-0.62	
O_3			0.87		
SO_2					0.84
HNO_3	0.33				0.82
Ca^{2+}	0.89				
K^+	0.77				
Na^+		0.99			
Mg^{2+}	0.34	0.92			
Cl^-		0.97			
NO_3^-	0.80				
NH_4^+	0.94				
SO_4^{2-}	0.89				0.26

Table 4.40 – Continued

Factor Number	PC1	PC2	PC3	PC4	PC5
Temperature	0.27		-0.52		0.27
Relative humidity				0.74	-0.33
Wind speed		0.26	0.52	0.57	
Precipitation				0.76	
Factor name	Combustion Emission	Sea Salt	Gas Phase Oxidation of Mercury	Mercury Wet Deposition	Industrial Source
Variance explained	22%	17%	14%	12%	10%

PC3 is also named as Gas Phase Oxidation of Hg due to the high loadings of O_3 (0.87), GEM (0.87), GOM (0.51) and PBM (0.29). The additional negative loading of temperature (-0.52) and positive loading of wind speed (0.52) in major variables may indicate the air flows containing more O_3 and mercury from the cold area (Cheng et al., 2013). This is reasonable because the mercury emissions in Nova Scotia were mainly located in the north of the site according to the NPRI where the temperature is usually lower than the wind flows from other directions (Figure 3.1). PC5 has an additional negative loading of relative humidity (-0.33,) and an additional positive loading of temperature (0.27) in major variables. These two loadings did not affect the name of this factor because the loadings of these two factors are relatively low and their presence did not reveal new possible sources. Therefore, this factor is also named as Industrial Source. PC4 is the additional factor extracted and is named as Mercury Wet Deposition due to negative association between mercury (GOM and PBM) and precipitation which is similar to the component Mercury Wet Deposition in Case 14.

Similar to Case 15, the components Gas Phase Oxidation of Hg and Industrial Source had the impact on ambient mercury concentrations. After including the meteorological parameters, an additional component Wet Deposition of Hg was identified to affect the ambient mercury concentrations. The rest of the factors contained at least one meteorological parameter in the major variables. However, they were assigned to the same names before including the meteorological parameters because the meteorological parameters did not play an important role in component interpretations.

4.2.4 Comparison between 2009 and 2010

In both years, four components were extracted using the chemical species only. Both components Gas Phase Oxidation of Hg and Sea Salt were extracted in 2009 and 2010. The component Gas Phase Oxidation of Hg had a large impact on the ambient concentrations of speciated mercury while the component Sea Salt had little impact on the mercury in both years. The factor Gas-particle Partitioning of Hg were only identified in 2009. This is consistent with the strong correlations between temperatures and GOM and PBM (r=0.46 and -0.43) in 2009 but weak correlations (r=0.04 and -0.16) in 2010.

Although the components Combustion Emission and Industrial Sources were identified in both years, they were existing in one factor (Combustion/industrial Emission) in 2009 but in two separate factors (Combustion Emission and Industrial Source) in 2010. The factor Combustion/industrial Emission had impacts on ambient mercury concentrations in 2009 while only the Industrial Source affected the ambient GOM concentrations in 2010. Same as the PMF results, the lack of impact on mercury for the Combustion Emission in 2010 was caused by the reduction of coal combustion in the USA (US EPA, 2011) and the reduction of SO_2 and Hg emission in Canada (Table A.1). The contribution of the component Industrial Source to mercury was shifted from PBM in 2009 to GOM in 2010. The shift is likely related to the change of the correlations between mercury and SO_2. The correlation between PBM and SO_2

was strong in 2009 (r=0.63, Table 4.5) but low correlation coefficient in 2010 (r=0.1, Table 4.22) while the correlations between GOM and SO_2 were moderate in both years (r=0.21 & 0.29, respectively, Table 4.5 and Table 4.22). The result was consistent with the change of the contributions of the factor Industrial Source to mercury from PBM in 2009 to GOM in 2010 in PMF results. After including the meteorological parameters in the input, Hg Wet Deposition, an additional component related to meteorological conditions, was identified in both years. The loadings of the chemical species in other factors and the variances they explained were similar to the cases using only chemical species in both years.

4.2.3 Comparison of the PCA results to Cheng et al. (2013)'s study

The results of Cheng et al. (2013)'s study are listed in Table 4.41 and Table 4.42 for 2009 and 2010, respectively. Four components were extracted in Cheng et al. (2013)'s study in 2009. Three out of four components were similar as the components Combustion/Industrial Source, Gas-particle partitioning of Hg, and Gas-phase Oxidation of Hg in Case 13. The component loadings of the components Combustion/industrial Source and Gas-phase Oxidation of Hg in 2009 were similar in this study and in Cheng et al. (2013)'s study. The component loadings of the components Condensation on Particles in Winter (Cheng et al., 2013) and Gas-particle Partitioning of Hg (this study) in 2009 were very different. Only the negative association between temperature and PBM were the same between Cheng et al. (2013)'s study and this study. Three components were extracted by Cheng et al. (2013) in 2010. However, none of the major variables of these three components is similar to the five components identified in this study in Case 16.

Table 4.41: Component loadings (>0.3) of 2009 in Cheng et al. (2013)

Factor Number	PC1	PC2	PC3	PC4
GEM			0.66	0.39
GOM			0.77	
PBM	0.37	0.77		
PM	0.85			
O_3		0.66	0.56	
SO_2	0.43	0.69		
HNO_3	0.82			
Ca^{2+}	0.56			
K^+	0.69			
Na^+				
Mg^{2+}				
Cl^-				
NO_3^-	0.54	0.51		
NH_4^+	0.88			
SO_4^{2-}	0.84			
Temperature	0.45	-0.83		
Relative humidity			-0.55	0.58
Wind speed				0.78
Precipitation				0.76
Factor name	Combustion/industrial/wildfires	Condensation on particles in winter	Photochemical production of GOM	GEM evasion from ocean
Variance explained	28.6%	17.1%	12.4%	11.7%

Table 4.42: Component loadings (>0.3) of 2010 in Cheng et al. (2013)

Factor Number	PC1	PC2	PC3
GEM	0.34	0.55	0.60
GOM	0.89		
PBM			0.73
O_3	0.69	0.41	0.46
SO_2			
HNO_3			
Ca^{2+}			
K^+			
Na^+			
Mg^{2+}			
Cl^-			
NO_3^-			
NH_4^+			
SO_4^{2-}			
Temperature			-0.89
Relative humidity	-0.82	0.38	
Wind speed		0.86	
Precipitation		0.77	
Factor name	Transport of free troposphere air	GEM evasion from ocean/regional background	Condensation on particles in winter
Variance explained	26.9%	25.4%	24.0%

The differences are probably caused by the following three items: exclusion method, variables included, and method used to retain the number of principal components. In Cheng et al. (2013)'s study, pairwise exclusion was used to make the full use of the dataset. The marine tracing species were excluded in 2009 while SO_2, HNO_3, and all ions were excluded in 2010 because they were not related with mercury. However, listwise exclusion and all species were used in order to be compared with the PMF results in this study. The method used to retain the number of components for further analysis was different. Fixed number (4 and 3 for 2009 and 2010, respectively) of components was retained in Cheng et al. (2013)'s study but the Kaiser criterion (eigenvalue>1) was used to retain the number of components in this study. All these differences could result in the differences in the results. Therefore, the comparison of the results suggests that the PCA results are sensitive to the input parameters.

4.3 Comparison between the PCA results and PMF results

The comparison between PMF model result and PCA result is based on the cases including only chemical species because the input of PMF model cannot include meteorological parameters. Both PMF model and PCA identified four factors using the same dataset. Among them, three and four factors are the same in 2009 and 2010, respectively. The major variables used to identify these factors or components were similar in both methods. The process photochemistry was identified to affect the ambient mercury concentrations most while Sea Salt was found to have little impact on ambient mercury concentrations using both methods in both years. The factors Combustion Emission and Industrial Source were identified as separate factors in PMF model in both 2009 and 2010. However, in PCA results, they were separate factors in 2010 but in one factor in 2009. PCA is appeared to be more sensitive to the correlations between different variables. The lack of the impact of the factors Combustion Emission on speciated mercury in 2009 and the shift of the factor impact on the mercury (from PBM in 2009 to GOM in 2010) were observed using different methods. An additional factor Gas-particle Partition of Hg was identified only by PCA in 2009. This is because the identification of the factor depended on the negative association between GOM and PBM. However, this kind of association cannot be revealed by the PMF model because all the variables in the factor are non-negative.

After including the meteorological parameters in the input of PCA, a new component related to meteorological process Mercury Wet Deposition was identified in both years. This is the advantage of PCA over the PMF model. Other analysis such as back trajectory, pollutant rose and ratio analysis could be adapted to verify the factors.

Overall, similar factors affecting mercury concentrations were identified using PCA and PMF model. The good agreement between the PCA results and PMF results is consistent with other inter-comparison exercises of receptor models in PM source apportionment (Viana et al., 2008a; Cesari et al., 2016). PCA identified more processes affecting ambient mercury concentrations because it could identify the processes according to negative correlations between the variables. However, PMF model could provide factor contributions and factor profiles. The quantitative evaluation and the considerations of the uncertainties makes PMF model a better receptor model in speciated mercury source apportionment. Both methods are suggested to be conducted to verify the factors.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Source apportionment was conducted with PMF model using the concentrations of speciated Hg and other air pollutants collected at KEJ site in 2009 and 2010. The same four factors were identified by PMF model in each year. They are Combustion Emission, Industrial Source, Photochemistry and Re-emission, and Sea Salt. In both years, the factor Photochemistry and Re-emission had the largest contributions to atmospheric mercury while the factor Sea Salt was not a significant source of mercury. The Combustion Emission and the Industrial Source had moderate contributions to GOM and PBM in 2009, respectively, while only the Industrial Source contributes to GOM in 2010. The differences of the factor contributions to mercury between two years were caused by the reduction of the coal combustion in the USA (US EPA, 2011) and the reduction of the SO_2 and mercury emission in Nova Scotia (Table A.1).

The PMF model performance on reproducing speciated mercury concentrations in each of the 12 cases was evaluated based on the performance indexes. The performance indexes include scaled residual plot, Obs/Pred scatter plot, Obs/Pred time series, the Pred/Obs ratios and the annual Predmean/Obsmean ratios. The observed GEM concentrations were best reproduced by PMF model among speciated mercury. The model performances on PBM and GOM are poor and their concentrations were underestimated in both years. The PMF model performance in 2010 is better than that of 2009, including better interpretability of the factors, and better reproduction of Hg concentrations.

The sensitivity of the PMF results to different treatment of input data was tested. The cases using geometric mean imputation or median imputation identified the same factors as the reference cases in both years. Similar model performances and factor contributions to the reference case were observed in 2009 and 2010. Similar to the imputation cases, excluding or combining GOM and PBM did not affect the source identification and model performances in PMF model. However, excluding or combining GOM and PBM did affect the source contributions as expected. Excluding or combining GOM and PBM concealed the factor contributions to GOM or PBM in both years. For example, the factors contributing to GOM only (Combustion emission, 2009; industrial source 2010) did not contribute to mercury after combining GOM and PBM. The use of scaling factor to increase the GOM and PBM concentrations improved the model performances in both years without affecting the factor identification and factor contribution. Therefore, PMF model seems to have difficulties in reproducing the species with low concentrations.

Source apportionment was also conducted by PCA in this study. The components identified by PCA were largely consistent with the factors in PMF results. Three and four components had the same names as the factors identified in PMF in 2009 and 2010, respectively. The shift of the factor impacts of Industrial Source on mercury (i.e. from PBM in 2009 to GOM in 2010) was also observed in the PCA results. The lack of the factor impacts of Combustion Emission on Hg species in 2010 is consistent with the PMF results. An additional component Gas-particle partitioning was identified by PCA in 2009 according to the negative association between GOM and PBM. The ability to identify the negative associations between two species is the advantage of PCA over PMF model because the results of the PMF model are required to be positive. After including the meteorological parameters in the input of PCA, an additional meteorological process Wet deposition of mercury was identified.

5.2 Recommendation

Based on the results and the limitations of this study, several recommendations were made for future studies:

- The comparison of the mercury source apportionment using PCA and PMF model was not conducted before. Therefore, more studies about the comparison between PCA and PMF results should be conducted to ensure the result in this study is not a one-time event.

- The factor combustion emission contributing to Hg species in 2009 only was attributed to the reduction of power plant. However, there is limitations in this interpretation because only a two-year study was conducted. A long term study is needed to verify the impact of the shutdown of power plants and reduction of Hg and SO_2 emissions in Nova Scotia on the reduction of ambient concentrations.

- The spikes in Figure 4.1 and Figure 4.2 are not well reproduced. Most air quality models including receptor models cannot reproduce the spikes. Right now, the causes of the spikes are unknown. The time series of other variables should be checked and more makers should be included to help interpreting the spikes in future studies.
- The comparison of the factor characterization and identification in this study suggested a good agreement between PCA and PMF result but the model performance on reproducing the observations were not compared. due to the limited length of the time. It is recommended to conduct the reconstruction of

observed concentrations using PCA-APCS model and compare the model ability on reproducing the observed concentrations to PMF results.

- There were uncertainties in the interpretation of some factors such as the factor Industrial Source because the variables included in this study was limited. More markers should be monitored to identify more specific sources. For example, Cu^{2+} could be included to identify Cu smelters.
- Although wind direction is an important meteorological parameter in source apportionment, wind direction cannot be used in PMF and PCA. Other analysis utilizing wind direction such as back trajectory and wind rose are recommended in future mercury source apportionment studies to verify the factors identified in PCA and PMF.
- The accuracy of the sources identified from the models are not evaluated in this studies due to the time limitation. Future studies are recommended to investigate the accuracy of the results.
- The identification of the factor Photochemistry is uncertain because the mechanism of photochemical oxidation of Hg is uncertain. Therefore, more studies about the mechanisms of the atmospheric mercury cycle should be conducted to improve the accuracy of factor interpretation.
- Although increasing the low GOM and PBM concentrations with a scaling factor improved the PMF model performance, the scaling factor changed the original variability of the data. More studies are recommended to investigate whether the results derived from the scaled data could more accurately reflect the source affecting mercury concentrations in the real word.

- All data treatment methods used in this study including the scaling factor are recommended for Hg source apportionment using PMF with data from a different site to see whether similar improvements on PMF performances on reproducing PBM and GOM concentrations could be observed.

REFERENCES

Acuña, E., & Rodriguez, C. (2004). The treatment of missing values and its effect on classifier accuracy. In *Classification, clustering, and data mining applications* (pp. 639-647). Berlin Heidelberg: Springer-Verlag.

Andersen, Z. J., Wahlin, P., Raaschou-Nielsen, O., Scheike, T., & Loft, S. (2007). Ambient particle source apportionment and daily hospital admissions among children and elderly in copenhagen. *Journal of Exposure Science and Environmental Epidemiology, 17*(7), 625-636. doi: 10.1038/sj.jes.7500546

Beaumont, R. (2012). An introduction to principal component analysis & factor analysis using spss 19 and r (psych package). Retrieved May 30, 2016, from http://www.floppybunny.org/robin/web/virtualclassroom/stats/statistics2/pca1.pdf

Belis, C. A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B. R., Gerboles, M., & Hopke, P. K. (2015). A new methodology to assess the performance and uncertainty of source apportionment models in intercomparison exercises. *Atmospheric Environment, 119*, 35-44. doi: 10.1016/j.atmosenv.2015.08.002

Blackwood, L. G. (1992). The lognormal distribution, environmental data, and radiological monitoring. *Environmental Monitoring and Assessment, 21*(3), 193-210. doi: 10.1007/BF00399687

Brennan, S. T., Lowenstein, T. K., & Horita, J. (2004). Seawater chemistry and the advent of biocalcification. *Geology, 32*(6), 473-476. doi: 10.1130/G20251.1

Brown, J. D. (2009). Choosing the right type of rotation in pca and efa. *Shiken: JALT Testing and Evaluation SIG Newsletter, 13*(3), 20-25.

Brown, S. G., Eberly, S., Paatero, P., & Norris, G. A. (2015). Methods for estimating uncertainty in pmf solutions: Examples with ambient air and water quality data and guidance on reporting pmf results. *Science of the Total Environment, 518-519*, 626-635. doi: dx.doi.org/10.1016/j.scitotenv.2015.01.022

Callén, M. S., de La Cruz, M. T., López, J. M., Navarro, M. V., & Mastral, A. M. (2009). Comparison of receptor models for source apportionment of the pm10 in zaragoza (spain). *Chemosphere, 76*(8), 1120-1129. doi: 10.1016/j.chemosphere.2009.04.015

Calvert, J. G., & Lindberg, S. E. (2005). Mechanisms of mercury removal by o 3 and oh in the atmosphere. *Atmospheric Environment, 39*(18), 3355-3367. doi: 10.1016/j.atmosenv.2005.01.055

Cangelosi, R., & Goriely, A. (2007). Component retention in principal component analysis with application to cdna microarray data. *Biology Direct, 2*(2), 1-21. doi: 10.1186/1745-6150-2-2

Carpi, A. (1997). Mercury from combustion sources: A review of the chemical species emitted and their transport in the atmosphere. *Water, Air, and Soil Pollution, 98*(3-4), 241-254. doi: 10.1023/A:1026429911010

Cesari, D., Amato, F., Pandolfi, M., Alastuey, A., Querol, X., & Contini, D. (2016). An inter-comparison of pm10 source apportionment using pca and pmf receptor models in three european sites. *Envionment Science and Pollution Research, 23*(15), 15133-15148. doi: 10.1007/s11356-016-6599-z

Cheng, I., Lu, J., & Song, X. (2009). Studies of potential sources that contributed to atmospheric mercury in toronto, canada. *Atmospheric Environment, 43*(39), 6145-6158. doi: 10.1016/j.atmosenv.2009.09.008

Cheng, I., Zhang, L., Blanchard, P., Graydon, J. A., & Louis, V. L. S. (2012). Source-receptor relationships for speciated atmospheric mercury at the remote experimental lakes area, northwestern ontario, canada. *Atmospheric Chemistry and Physics: Atmospheres, 12*(4), 1903-1922. doi: 10.5194/acp-12-1903-2012

Cheng, I., Zhang, L., Blanchard, P., Dalziel, J., Tordon, R., Huang, J., & Holsen, T. M. (2013). Comparisons of mercury sources and atmospheric mercury processes between a coastal and inland site. *Journal of Geophysical Research: Atmospheres, 118*(5), 2434-2443. doi: 10.1002/jgrd.50169

Cheng, I., Xu, X., & Zhang, L. (2015). Overview of receptor-based source apportionment studies for speciated atmospheric mercury. *Atmospheric Chemistry and Physics, 15*(14), 7877-7895. doi: 10.5194/acp-15-7877-2015

Cheng, I., Zhang, L., & Xu, X. (2016). Impact of measurement uncertainties on receptor modeling of speciated atmospheric mercury. *Scientific Reports, 6*, 20676. doi: 10.1038/srep20676

Chow, J. C., Watson, J. G., Lowenthal, D. H., Chen, L.-W. A., & Motallebi, N. (2011). Pm2.5 source profiles for black and organic carbon emission inventories. *Atmospheric Envrionment, 45*(31), 5407-5414. doi: 10.1016/j.atmosenv.2011.07.011

Clever, H. L., Johnson, S. A., & Derrick, M. E. (1985). The solubility of mercury and some sparingly soluble mercury salts in water and aqueous electrolyte solutions. *Journal of Physical and Chemical Reference Data, 14*(3), 631-680. doi: 10.1063/1.555732

Compeau, G. C., & Bartha, R. (1985). Sulfate-reducing bacteria: Principal methylators of mercury in anoxic estuarine sediment. *Applied and Environmental Microbiology, 50*(2), 498-502.

Croghan, C. W., & Egeghy, P. P. (2003). Methods of dealing with values below the limit of detection using sas. Retrieved May 30, 2016, from http://analytics.ncsu.edu/sesug/2003/SD08-Croghan.pdf

Environment Canada. (2011a). National climate data and information archive. Retrieved May 29, 2016, from http://climate.weatheroffice.gc.ca/climateData/canada_e.html

Environment Canada. (2011b). National air pollution surveillance data for 2009-2010 (naps). Retrieved May 29, 2016, from http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx

Environment Canada . (2013a). Chemical properties about mercury. Retrieved May 29, 2016, from http://www.ec.gc.ca/mercure-mercury/default.asp?lang=En&n=10C3AF2D-1

Environment Canada . (2013b). Mercury in the food chain. Retrieved May 26, 2016, from https://www.ec.gc.ca/mercure-mercury/default.asp?lang=En&n=D721AC1F-1

Environment Canada . (2015). The canadian air and precipitation monitoring network (capmon). Retrieved Aug 20, 2015, from https://www.ec.gc.ca/rs-mn/default.asp?lang=En&n=752CE271-1

Environment Canada . (2016). National pollutants release inventory database (npri). Retrieved May 29, 2016, from https://www.ec.gc.ca/inrp-npri/default.asp?lang=en&n=0EC58C98-1

Evers, D. C., Han, Y.-J., Driscoll, C. T., Kamman, N. C., Goodale, M. W., Lambert, K. F., Holsen, T. M., Chen, C. Y., Clair, T. A., & Butler, T. (2007). Biological mercury hotspots in the northeastern united states and southeastern canada. *BioScience, 57*(1), 29-43. doi: 10.1641/B570107

Gaffney, J. S., & Marley, N. A. (2014). In-depth review of atmospheric mercury: Sources, transformations, and potential sinks. *Energy and Emission Control Technologies, 2*, 1-21. doi: 10.2147/EECT.S37038

Gao, F. (2007). *A comprehensive investigation of ambient hg in the ohio river valley: Source-receptor relationship and meteorological impact* (Master's thesis), Ohio University. Retrieved from https://etd.ohiolink.edu/!etd.send_file?accession=ohiou1194624797&disposition=inline

Goodsite, M. E., Plane, J. M. C., & Skov, H. (2004). A theoretical study of the oxidation of hg0 to hgbr2 in the troposphere. *Environmental Science & Technology, 38*(6), 1772-1776. doi: 10.1021/es034680s

Gustin, M. S., Huang, J., Miller, M. B., Peterson, C., Jaffe, D. A., Ambrose, J., Finley, B. D., Lyman, S. N., Call, K., Talbot, R., Feddersen, D., Mao, H., & Lindberg, S. E. (2013). Do we understand what the mercury speciation instruments are actually measuring? Results of ramix. *Environmental Science & Technology, 47*(13), 7295-7306. doi: 10.1021/es3039104

Gustin, M. S., Amos, H. M., Huang, J., Miller, M. B., & Heidecorn, K. (2015). Measuring and modeling mercury in the atmosphere: A critical review. *Atmospheric Chemistry and Physics, 15*(10), 5697-5713. doi: 10.5194/acp-15-5697-2015

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hastings, A., & Gross, L. J. (2012). *Encyclopedia of theoretical ecology*. London, England: University of California Press.

Henry, R. C., Lewis, C. W., Hopke, P. K., & Williamson, H. J. (1984). Review of receptor model fundamentals. *Atmospheric Environment, 18*(8), 1507-1515. doi: 10.1016/0004-6981(84)90375-5

Hopke, P. K. (2000). A guide to positive matrix factorization. Retrieved May 30, 2016, from https://www3.epa.gov/ttnamti1/files/ambient/pm25/workshop/laymen.pdf

Huang, J., Choi, H.-D., Hopke, P. K., & Holsen, T. M. (2010). Ambient hg sources in rochester, ny: Results from principle components analysis (pca) of hg monitoring network data. *Environmental Science & Technology, 44*(22), 8441-8445. doi: 10.1021/es102744j

Huang, J., Miller, M. B., Weiss-Penzias, P., & Gustin, M. S. (2013). Comparison of gaseous oxidized hg measured by kcl-coated denuders, and nylon and cation exchange membranes. *Environmental Science & Technology, 47*(13), 7307-7316. doi: 10.1021/es4012349

Huang, S., Rahn, K. A., & Arimoto, R. (1999). Testing and optimizing two factor-analysis techniques on aerosol at narragansett, rhode island. *Atmospheric Environment, 33*(14), 2169-2185. doi: 10.1016/S1352-2310(98)00324-0

IBM Corp. (2013). Ibm spss statistics for windows (version 22.0). Retrieved Septemeber 9, 2015, from http://www-01.ibm.com/support/docview.wss?uid=swg21646821

IBM Corp. (2014). Pairwise vs. Listwise deletion: What are they and when should i use them? Retrieved June 30, 2016, from http://www-01.ibm.com/support/docview.wss?uid=swg21475199

Jackson, J. E. (1991). *A uses's guide to principal components*. New York: John Wiley & Sons.

Jolliffe, I. T. (2002). *Principal component analysis*. New York: Springer-Verlag.

Karagulian, F., & Belis, C. A. (2012). Enhancing source apportionment with receptor models to foster the air quality directive implementation. *International Journal of Environment and Pollution, 50*(14), 1-4. doi: 10.1504/IJEP.2012.051192

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151. doi: 10.1177/001316446002000116

Lalonde, J. D., Poulain, A. J., & Amyot, M. (2002). The role of mercury redox reactions in snow on snow-to-air mercury transfer. *Envrionmental Science & Technology, 36*(2), 174-178. doi: 10.1021/es010786g

Laurier, F. J. G., Mason, R. P., Whalin, L., & Kato, S. (2003). Reactive gaseous mercury formation in the north pacific ocean's marine boundary layer: A potential role of halogen chemistry. *Journal of Geophysical Research: Atmospheres, 108*(D17), 4529. doi: 10.1029/2003JD003625

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*(19), 3201-3212. doi: 10.1016/S1352-2310(99)00113-2

Leuchner, M., Gubo, S., Schunk, C., Wastl, C., Kirchner, M., Menzel, A., & Plass-Dülmer, C. (2015). Can positive matrix factorization help to understand patterns of organic trace gases at the continental global atmosphere watch site hohenpeissenberg? *Atmospheric Chemistry and Physics, 15*(3), 1221-1236. doi: 10.5194/acp-15-1221-2015

Li, J., Sommar, J., Wängberg, I., Lindqvist, O., & Wei, S. Q. (2008). Short-time variation of mercury speciation in the urban of göteborg during g te-2005. *Atmospheric Environment, 42*(36), 8382-8388. doi: 10.1016/j.atmosenv.2008.08.007

Lindberg, S. E., & Stratton, W. J. (1998). Atmospheric mercury speciation: Concentrations and behavior of reactive gaseous mercury in ambient air. *Environmental Science & Technology, 32*(1), 49-57. doi: 10.1021/es970546u

Lindqvist, O., & Rodhe, H. (1985). Atmospheri mercury - a review. *Tellus, 37B*(3), 136-159. doi: 10.1111/j.1600-0889.1985.tb00062.x

Liu, B., Keeler, G. J., Dvonch, J. T., Barres, J. A., Lynam, M. M., Marsik, F. J., & Morgan, J. T. (2007). Temporal variability of mercury speciation in urban air. *Atmospheric Environment, 41*(9), 1911-1923. doi: 10.1016/j.atmosenv.2006.10.063

Liu, W., Hopke, P. K., Han, Y.-J., Yi, S.-M., Holsen, T. M., Cybart, S., Kozlowski, K., & Milligan, M. (2003). Application of receptor modeling to atmospheric constituents at potsdam and stockton, ny. *Atmospheric Environment, 37*(36), 4997-5007. doi: 10.1016/j.atmosenv.2003.08.036

Lynam, M. M., & Keeler, G. J. (2005). Artifacts associated with the measurement of particulate mercury in an urban environment: The influence of elevated ozone concentrations. *Atmospheric Environment, 39*(17), 3081-3088. doi: 10.1016/j.atmosenv.2005.01.036

Lynam, M. M., & Keeler, G. J. (2006). Source-receptor relationships for atmospheric mercury in urban detroit, michigan. *Atmospheric Environment, 40*(17), 3144-3155. doi: 10.1016/j.atmosenv.2006.01.026

Manolopoulos, H., & Snyder, D. C. (2007). Sources of speciated atmospheric mercury at a residential neighborhood impacted by industrial sources. *Environmental Science & Technology, 41*(16), 5626-5633. doi: 10.1021/es0700348

Morel, F. M. M., Kraepiel, A. M. L., & Amyot, M. (1998). The chemical cycle and bioaccumulation of mercury. *Ecology, Evolution and Systematics, 29*, 543-

566. doi: 10.1146/annurev.ecolsys.29.1.543

National Atmospheric Deposition Program (NADP). (2016). Amnet data access. Retrieved May 30, 2016, from http://nadp.sws.uiuc.edu/amn/data.aspx

New Jersey Department of Environmental Protection (NJDEP). (2011). Nox and voc emission trends: Ozone precursors Retrieved June 4, 2016, from http://www.nj.gov/dep/dsr/trends/pdfs/nox-voc.pdf

Nierenberg, D. W., Nordgren, R. E., Chang, M. B., Siegler, R. W., Blayney, M. B., Hochberg, F., Toribara, T. Y., Cernichiari, E., & Clarkson, T. (1998). Delayed cerebellar disease and death after accidental exposure to dimethylmercury. *The New England Journal of Medicine, 338*, 1672-1676. doi: 10.1056/NEJM199806043382305

Paatero, P., Eberly, S., Brown, S. G., & Norris, G. A. (2014). Methods for estimating uncertainty in factor analytic solutions. *Atmospheric Measurement Techniques, 7*, 781-797. doi: 10.5194/amt-7-781-2014

Pack, E. C., Kim, C. H., Lee, S. H., Lim, C. H., Sung, D. G., Kim, M. H., Park, K. H., Hong, S.-S., Lim, K. M., Choi, D. W., & Kim, S. W. (2014). Effects of environmental temperature change on mercury absoption in aquatic organisms with respect to climate warming. *Journal of Toxicology and Environmental Health, 77*, 1477-1490. doi: 10.1080/15287394.2014.955892

Pakkanen, T. A. (1996). Study of formation of coarse particle nitrate aerosol. *Atmospheric Environment, 30*(14), 2475-2482. doi: 10.1016/1352-2310(95)00492-0

Pal, B., & Ariya, P. A. (2004). Studies of ozone initiated reactions of gaseous mercury: Kinetics, product studies, and atmospheric implications. *Physical Chemistry Chemical Physics, 6*(3), 572-579. doi: 10.1039/B311150D

Parkin, T. B., & Robinson, J. A. (1992). Analysis of lognormal data. In *Advances in soil science* (pp. 193-235). New York: Springer

Pavlovic, R. T., Nopmongcol, U., Kimura, Y., & Allen, D. T. (2006). Ammonia emissions, concentrations and implications for particulate matter formation in houston, tx. *Atmospheric Environment, 40*(Supplement 2), 538-551. doi: 10.1016/j.atmosenv.2006.04.071

Pekey, H., Karakaş, D., & Bakogˇlu, M. (2004). Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. *Marine Pollution Bulletin, 49*(9-10), 809-818. doi: 10.1016/j.marpolbul.2004.06.029

Pennsylvania State University (PennState). (2016). Applied multivariate statistical analysis. Retrieved September 1, 2016, from https://onlinecourses.science.psu.edu/stat505/node/54

Pirrone, N., Cinnirella, S., Feng, X., Finkelman, R. B., Friedli, H. R., Leaner, J., Mason, R., Mukherjee, A. B., Stracher, G. B., Streets, D. G., & Telmer, K. (2010). Global mercury emissions to the atmosphere from anthropogenic and natural sources. *Atmospheric Chemistry and Physics, 10*(13), 5951-5964. doi: 10.5194/acp-10-5951-2010

Pitchford, M. L., Poirot, R. L., Schichtel, B. A., & Malm, W. C. (2009). Characterization of the winter midwestern particulate nitrate bulge. *Journal of the Air & Waste Management Association, 59*(9), 1061-1069. doi: 10.3155/1047-3289.59.9.1061

Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C., & Sisler, J. F. (1998). Atmospheric aerosol over alaska: 2. Elemental composition and sources. *Journal of Geophysical Research: Atmospheres, 103*(D15), 19045-19057. doi: 10.1029/98JD01212

Puxbaum, H., Caseiro, A., Sánchez-Ochoa, A., Kasper-Giebl, A., Claeys, M., Gelencsér, A., Legrand, M., Preunkert, S., & Pio, C. (2007). Levoglucosan levels at background sites in europe for assessing the impact of biomass combustion on the european aerosol background. *Journal of Geophysical Research: Atmospheres, 112*(D23), 1-11. doi: 10.1029/2006JD008114

Reff, A., Eberly, S. I., & Bhave, P. V. (2007). Receptor modeling of ambient particulate matter data using positive matrix factorization: Review of existing methods. *Journal of the Air & Waste Management Association, 57*(2), 146-154. doi: 10.1080/10473289.2007.10465319

Reff, A., Bhave, P. V., Simon, H., Pace, T. G., Pouliot, G. A., Mobley, J. D., & Houyoux, M. (2009). Emissions inventory of pm2.5 trace elements across the united states. *Environmental Science & Technology, 43*(15), 5790-5796. doi: 10.1021/es802930x

Ren, X., Luke, W. T., Kelley, P., Cohen, M., Ngan, F., Artz, R., Walker, J., Brooks, S., Moore, C., Swartzendruber, P., Bauer, D., Remeika, J., Hynes, A., Dibb, J., Rolison, J., Krishnamurthy, N., Landing, W. M., Hecobian, A., Shook, J., & Huey, L. G. (2014). Mercury speciation at a coastal site in the northern gulf of mexico: Results from the grand bay intensive studies in summer 2010 and spring 2011. *Atmosphere, 5*(2), 230-251. doi: 10.3390/atmos5020230

Rutter, A. P., & Schauer, J. J. (2007). The effect of temperature on the gas-particle partitioning of reactive mercury in atmospheric aerosols. *Atmospheric Environment, 41*(38), 8647-8657. doi: 10.1016/j.atmosenv.2007.07.024

Sigler, J. M., Mao, H., Sive, B. C., & Talbot, R. (2009). Oceanic influence on atmospheric mercury at coastal and inland sites: A springtime noreaster in new england. *Atmospheric Chemistry and Physics, 9*(12), 4023-4030. doi: 10.5194/acp-9-4023-2009

Slemr, F., Schuster, G., & Seiler, W. (1985). Distribution, speciation, and budget of atmospheric mercury. *Journal of Atmospheric Chemistry, 3*(4), 407-434. doi: 10.1007/BF00053870

Stamenkovic, J., Lyman, S., & Gustin, M. S. (2007). Seasonal and diel variation of atmospheric mercury concentrations in the reno (nevada, USA) airshed. *Atmospheric Environment, 41*(31), 6662-6672. doi: 10.1016/j.atmosenv.2007.04.015

Steffen, A., Scherz, T., Olson, M., Gay, D., & Blanchard, P. (2012). A comparison of data quality control protocols for atmospheric mercury speciation measurements. *Journal of Environmental Monitoring, 14*(3), 752-765. doi: 10.1039/C2EM10735J

Swartzendruber, P. C., Jaffe, D. A., Prestbo, E. M., Weiss-Penzias, P., Selin, N. E., Park, R., Jacob, D. J., Strode, S., & Jaeglé, L. (2006). Observations of reactive gaseous mercury in the free troposphere at the mount bachelor observatory. *Journal of Geophysical Research: Atmospheres, 111*(D24), 1-12. doi: 10.1029/2006JD007415

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. MA: Allyn and Bacon.

Talmi, Y., & Mesmer, R. E. (1975). Studies on vaporization and halogen decomposition of methyl mercury compounds using gc with a microwave detector. *Water Research, 9*(5-6), 547-552. doi: 10.1016/0043-1354(75)90080-9

Taylor, S. L., Ruhaak, L. R., Kelly, K., Weiss, R. H., & Kim, K. (2016). Effects of imputation on correlation: Implications for analysis of mass spectrometry data from multiple biological matrices. *Briefings in Bioinformatics*. doi: 10.1093/bib/bbw010

Tekran Inc. (2010). Products-ambient air-overview. Retrieved May 29, 2016, from http://www.tekran.com/products/ambient-air/overview/

Thompson, M., Ellison, S. L. R., & Wood, R. (2006). The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. *Pure and Applied Chemistry, 78*(1), 145-196. doi: 10.1351/pac200678010145

Thurston, G. D., & Spengler, J. D. (1985a). A multivariate assessment of meteorological influences on inhalable particle source impacts. *Journal of Climate and Applied Meteorology, 24*(11), 1245-1256.

Thurston, G. D., & Spengler, J. D. (1985b). A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan boston.

Atmospheric Environment, 19(1), 9-25. doi: doi:10.1016/0004-6981(85)90132-5.

Ullrich, S. M., Tanton, T. W., & Abdrashitova, S. A. (2001). Mercury in the aquatic environment: A review of factors affecting methylation. *Critical Reviews in Environmental Science and Technology, 31*(3), 241-293. doi: 10.1080/20016491089226

United Nations Environmental Programme (UNEP). (2002). Inter-organization programme for the sound management of chemicals: In global mercury assessment. Retrieved May 30, 2016, from http://www.unep.org/gc/gc22/Document/UNEP-GC22-INF3.pdf

United Nations Environmental Programme (UNEP). (2008). The global atmospheric mercury assessment: Sources, emissions and transport. Retrieved June 10, 2016, from http://www.unep.org/chemicalsandwaste/Portals/9/Mercury/Documents/Publications/UNEP_GlobalAtmosphericMercuryAssessment_May2009.pdf

United Nations Environmental Programme (UNEP). (2013). Global mercury assessment 2013: Sources, emissions, releases, and transport. Retrieved May 30, 2016, from http://www.unep.org/PDF/PressReleases/GlobalMercuryAssessment2013.pdf

US Environmental Protection Agency (US EPA). (2004). Protocol for applying and validating the cmb model for pm2.5 and voc. Retrieved July 20, 2015, from https://www3.epa.gov/scram001/models/receptor/CMB_Protocol.pdf

US Environmental Protection Agency (US EPA). (2011). Clean air interstate rule, acid rain program and former nox budget trading program: 2010 progress report emission, compliance, and market analyses. Retrieved May 30, 2016, from https://www.epa.gov/sites/production/files/2015-08/documents/arpcair10_analyses.pdf

US Environmental Protection Agency (US EPA). (2014a). Epa positive matrix factorization (pmf) 5.0 fundamentals and user guide. Retrieved May 30, 2016, from https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf

US Environmental Protection Agency (US EPA). (2014b). Speciate version 4.4. Retrieved Sepember 11, 2016, from https://www3.epa.gov/ttn/chief/software/speciate/

Vanderzee, C. E., & Swanson, J. A. (1974). The enthalpy of precipitation of mercury(i) chloride, and the thermodynamic properties of aqueous mercury(i) and mercury(ii) ions *The Journal of Chemical Thermodynamics, 6*(9), 827-843. doi: 10.1016/0021-9614(74)90228-6

Viana, M., Pandolfi, M., Minguillón, M. C., Querol, X., Alastuey, A., Monfort, E., & Celades, I. (2008a). Inter-comparison of receptor models for pm source apportionment: Case study in an industrial area. *Atmospheric Environment, 42*(16), 3820-3832. doi: 10.1016/j.atmosenv.2007.12.056

Viana, M., Kuhlbusch, T. A. J., Querol, X., Alastuey, A., Harrison, R. M., Hopke, P. K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A. S. H., Hueglin, C., Bloemen, H., Wåhlin, P., Vecchi, R., Miranda, A. I., Kasper-Giebl, A., Maenhaut, W., & Hitzenberger, R. (2008b). Source apportionment of particulate matter in europe: A review of methods and results. *Journal of Aerosol Science, 39*(10), 827-849. doi: 10.1016/j.jaerosci.2008.05.007

Wang, Y., Huang, J., Hopke, P. K., Rattigan, O. V., Chalupa, D. C., Utell, M. J., & Holsen, T. M. (2013). Effect of the shutdown of a large coal-fired power plant on ambient mercury species. *Chemosphere, 92*(4), 360-367. doi: 10.1016/j.chemosphere.2013.01.024

Watson, J. G., L. -W. Antony Chen, Chow, J. C., Doraiswamy, P., & Lowenthal, D. H. (2008). Source apportionment: Findings from the u.S. Supersites program. *Journal of the Air & Waste Management Association, 58*(2), 265-288. doi: 10.3155/1047-3289.58.2.265

Weiss-Penzias, P., Amos, H. M., Selin, N. E., Gustin, M. S., Jaffe, D. A., Obrist, D., Sheu, G.-R., & Giang, A. (2015). Use of a global model to understand speciated atmospheric mercury observations at five high-elevation sites. *Atmospheric Chemistry and Physics, 15*(3), 1161-1173. doi: 10.5194/acp-15-1161-2015

Wiener, J. G., Krabbenhoft, D. P., Heinz, G. H., & Scheuhammer, A. M. (2002). Ecotoxicology of mercury. In *Handbook of ecotoxicology* (2nd ed., pp. 409-443). Boca Raton, FL: CRC Press

Williams, B., Brown, T., & Onsman, A. (2012). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care (JEPHC), 8*(3), 1-14.

Wyn, B., Kidd, K. A., Burgess, N. M., Curry, R. A., & Munkittrick, K. R. (2010). Increasing mercury in yellow perch at a hotspot in atlantic canada, kejimkujik national park. *Environmental Science & Technology, 44*(23), 9176–9181. doi: 10.1021/es1018114

Zhang, L., Vet, R., Wiebe, A., Mihele, C., Sukloff, B., Chan, E., Moran, M. D., & Iqbal, S. (2008). Characterization of the size-segregated water-soluble inorganic ions at eight canadian rural sites. *Atmospheric Chemistry and Physics, 8*(23), 7133-7151. doi: 10.5194/acp-8-7133-2008

APPENDICES

Appendix A:
Point Sources of Hg and Other Pollutants

Table A.1. Point emissions of Hg and other pollutants reported in NPRI within Nova Scotia (EC, 2016). Bold sites are marked on the Figure 3.1.

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO2 (Tonnes)		NO2 (Tonnes)		NH3 (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Brooklyn Power	Brooklyn (44N, 64W)	50 Km southeast	0	0	9.9	26	309	259	0	0
Michelin North America (Canada)-Bridgewater Plant	Bridgewater (44N, 64W)	53 Km east	0	0	195	184	68	63	0	0
High Liner Foods Inc.	Lunenburg (44N,64W)	72 Km east	0	0	27	27	0	0	0	0
Department of National Defence – 14 Wing Greenwood	Greenwood (45N, 66W)	75 Km north	0	0	55	68	19	18	0	0
Louisana Pacific Canada Ltd.	East River (44N, 64W)	88 Km northeast	0	0	122	102	100	99	0	0
Maple Leaf Foods – Larsen Packers Limited	Berwick (45N, 64W)	89 Km northeast	0	0	51	38	0	0	0	0
Michelin North America (Canada) - Waterville Plant	Waterville (45N, 64W)	92 Km northeast	0	0	162	182	57	62	0	0
Acadia University – Acadia Campus	Wolfville (45N, 64W)	108 Km northeast	0	0	77	73	27	26	0	0
CKF. Inc.	Hantsport (45N, 64W)	116 Km northeast	0	0	66	57	21	72	0	0
Minas Basin Pulp and Power	Hantsport (45N, 64W)	116 Km northeast	0	0	225	260	66	76	0	0
Mount Saint Vincent University	Halifax (44N, 63W)	129 Km northeast	0	0	27	13	7.2	3.9	0	0
Department of National Defence – Canadian Forces Ammunition Depot	Bedford (44N, 64W)	131 Km northeast	0	0	56	50	0	0	0	0

Table A.1 – Continued 1

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO2 (Tonnes)		NO2 (Tonnes)		NH3 (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Department of National Defence - Windsor Park	Halifax (44N, 63W)	132 Km northeast	0	0	59	44	36	30	0	0
Department of National Defence – Stadacona/Dockyard	Halifax (44N, 63W)	133 Km northeast	0	0	211	177	58	51	0	0
Capital Health – Camp Hill Site Heating Plant	Halifax (44N, 63W)	133 Km northeast	0	0	15	12	14	20	0	0
DalHousie University	Halifax (44N, 63W)	133 Km northeast	0.18	0.15	253	260	89	72	0	0
Saint Mary's University	Halifax (44N, 63W)	133 Km northeast	0	0	1.2	0	3	0	0	0
Oland Brewery	Halifax (44N, 63W)	133 Km Northeast	0	0	31	0	0	0	0	0
Nova Scotia Power – Tufts Cove Generating Station	Dartmouth (44N, 63W)	134 Km northeast	0	0	2,205	2,205	3,054	3,054	0	0
Capital Health-Victoria General Hospital Central Heating Plant	Halifax (44N, 63W)	134 Km northeast	0	0	215	7.6	60	19	0	0
Maritime Paper Products Ltd.	Dartmouth (44N, 63W)	134 Km northeast	0	0	7.2	0.868	3.1	2.1	0	0
Nova Scotia Power –Burnside Combustion Turbines	Dartmouth (44N, 63W)	134 Km northeast	0	0	0	0	60	40	0	0
Capital Health – Nova Scotia Hospital Central Heating Plant	Dartmouth (44N, 63W)	136 Km northeast	0	0	3.3	1.1	9.3	8.7	0	0
Imperial Oil – Dartmouth Refinery	Dartmouth (44N, 63W)	137 Km northeast	2.6	2.9	4,231	3,073	1,543	1,251	0.593	2.2
Department of National Defence – 12 Wing Shearwater	Shearwater (44N, 63W)	138 Km northeast	0	0	150	127	43	38	0	0
Martells Contracting	Elmsdale (45N, 64W)	154 Km northeast	0	0	28	17	4.5	2.8	0	0

Table A.1 – Continued 2

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO2 (Tonnes)		NO2 (Tonnes)		NH3 (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
The Shaw Group Ltd.	Hardwoodlands (45N, 64W)	160 Km northeast	0	0	0	0	27	19	0	0
Lafarge Canada Inc. – Brookfield Cement Plant	Brookfield (45N, 64W)	180 Km northeast	5	5.9	562	667	498	591	0	0
Polycello	Amherst (46N, 64W)	183 Km northeast	0	0	0.003	0.002	0.462	0.335	0	0
Enligna Canada Inc.	Middle Musquodoboit (45N, 63W)	188 Km northeast	0	0	2.8	2.9	25	26	0	0
Oxford Frozen Foods	Oxford (46N, 64W)	188 Km northeast	0	0	66	59	0	0	0.9	0
Municipality of the county of Colchester – Wastewater Treatment Facility	Truro (45N, 63W)	188 Km northeast	0	0	0	0	0	0	2	0.08
Crossley Carpet Mills Limited	Truro (45N, 63W)	189 Km northeast	0	0	40	32	12	11	0	0
Rothsay	Truro (45N, 63W)	189 Km northeast	0	0	77	60	0	0	0	0
Stanfield’s Ltd.	Truro (45N, 63W)	191 Km northeast	0	0	21	21	0	0	0	0
Stella-Jones Inc.	Truro (45N, 63W)	192 Km northeast	0	0	12	19	2.9	4.2	0	0
The Canadian Salt Company Limited – Pugwash Mine and Refinery	Pugwash (46N, 64W)	209 Km northeast	0	0	168	153	32	31	0	0
Michelin North America (Canada) – Pictou County Plant	New Glasgow (46N, 63W)	245 Km northeast	0	0	209	229	72	78	0	0
Maritime Steel and Foundries Limited	New Glasgow (46N, 63W)	245 Km northeast	0	0	0.25	0	0.875	0	0	0

Table A.1 – Continued 3

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO2 (Tonnes)		NO2 (Tonnes)		NH3 (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Nova Scotia Power – Trenton Generating Station	Trenton (46N, 63W)	248 Km northeast	33	19	30,429	19,257	5,126	5,577	0	0
Nova Forge Corporation	Trenton (46N, 63W)	248 Km northeast	0	0	3.1	0	0	0	0	0
Northern Pulp Nova Scotia Corporation	New Glasgow (46N, 63W)	266 Km northeast	0	0	246	89	688	676	42	46
St. Francis Xavier University	Antigonish (46N, 62W)	291 Km northwast	0	0	41	36	25	17	0	0
Exxonmobil Canada Properties – Goldboro Gas Plant	Goldboro (45N, 62W)	300 Km northeast	0	0	0	0	521	415	0	0
Nova Scotia Power – Point Tupper Generating Station	Port Hawkesbury (46N, 61W)	335 Km northeast	12	9.5	9,394	5,721	1,952	1,952	0	0
Newpage Port Hawkesbury Corp.	Port Hawkesbury (46N, 61W)	355 Km northeast	0	0	294	85	404	306	0.23	0.23
Exxonmobil Canada Properties – Point Tupper Fractionation Plant	Port Hawkesbury (46N, 61W)	335 Km northeast	0	0	0	0	48	23	0	0
Exxonmobil Canada Properties – Thebaud Platform	Offshore (43N, 60W)	402 Km east	0	0	0	0	135	126	0	0
Exxonmobil Canada Properties – North Triumph Platform	North Triumph Platform (43N, 60W)	433 Km east	0	0	0	0	26	29	0	0
Nova Scotia Power – Point Aconi Generating Station	Point Aconi (46N, 60W)	442 Km northeast	2.7	2.8	3,627	3,365	1,759	1,747	0	0
Exxonmobil Canada Properties – Venture Platform	Venture Platform (44N, 59W)	450 Km east	0	0	18	0	54	51	0	0

Table A.1 – Continued 4

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO2 (Tonnes)		NO2 (Tonnes)		NH3 (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Cape Breton University	Sydney (46N, 60W)	450 Km northeast	0	0	60	57	0	0	0	0
Nova Scotia Power – Lingan Generating Station	Lingan (46N, 60W)	457 Km northeast	92	50	55,208	33,479	5,106	5,219	0	0
Provincial total emission			147.5	90.3	108,961	70,336	22,165	22,166	45.7	48.5

Appendix B:
Input Data Processing Step

Case 1

1. Copy the original data to a new spreadsheet and exclude the meteorological data from the table.
2. Replace the blanks with -999 as the indicator of the missing value.
3. Insert the MDL of each species to the spreadsheet.
4. Rename the sheet as original data.
5. Create a new sheet named as uncertainty.
6. Copy the species name to the header and the copy the date in the original data sheet to the uncertainty sheet.
7. When the concentration of the species is not smaller than its MDL, set the uncertainty to 5/6*original data. When the concentration is less than its MDL, set the uncertainty to $\sqrt{(a * oringinal\ data)^2 + (0.5 * its\ MDL)^2}$, a= 0.15 for Hg, a=0.1 for other pollutants.
8. Create a new sheet named as concentration and copy the original data table to this sheet.
9. Save the file and change the name to Case 1 input.

Case 2

1. Copy the original data to a new spreadsheet and exclude the meteorological data from the table.

2. Replace the blanks with the geometric mean of the species. When there was 0 value in the species, the geometric mean was calculated using the data excluding the missing values and the 0 values.
3. Insert the MDL of each species to the spreadsheet.
4. Rename the sheet as original data.
5. Create a new sheet named as uncertainty.
6. Copy the species name to the header and the copy the date in the original data sheet to the uncertainty sheet.
7. When the concentration of the species is not smaller than its MDL, set the uncertainty to 5/6*original data. When the concentration is less than its MDL, set the uncertainty to $\sqrt{(a * oringinal\ data)^2 + (0.5 * its\ MDL)^2}$, a= 0.15 for Hg species, a=0.1 for other pollutants. For the Hg samples using geometric mean imputations, the corresponding uncertainties were set to 1*MDL for GEM, 10*MDL for GOM, and 2*MDL for PBM.
8. Create a new sheet named as concentration and copy the original data table to this sheet.
9. Save the file and change the name to Case 2 input.

Case 3

The treatment process is the same as Case 2 except using the median value of the species to impute the missing values.

Case 4

The treatment process is the same as Case 1 except step 1. In step 1, the GOM and PBM should also be excluded in this case.

Case 5

The treatment process is the same as Case 1 except step 1. In step 1, the GOM and PBM should be combined to RM after copied in the original data. When one of the GOM and PBM is missing, the RM concentration is considered as missing. The MDL for RM is considered to be 4 pg/m^3.

Case 6

The treatment process is the same as Case 1 except Step 1. In step 1, the GOM and PBM concentrations need to be multiplied by the scaling factor.

The input treatment in 2010 is the same as in 2009.

Appendix C:
Comparison of the 3-factor, 4-factor, 5-factor PMF Results

C.1 Comparison in 2009

The 3-factor and 5-factor results of 2009 derived from PMF are listed in Table C.1 and Table C.2, respectively. In 3-factor results (Table C.2), the factor Combustion Emission in 4-factor results were reapportioned to Factor 1 and Factor 2. This led to the combination of the three factors Combustion Emission, Industrial Source, and Photochemistry & Re-emission of Hg. Factor 1 in 3-factor results is the combination of the factors Photochemistry & Re-emission of Hg and Combustion Emission in 4-factor results. Factor 2 in 3-factor results is the combination Industrial Source and Combustion Emission in 4-factor results. Factor 3 is the same as the factor Sea Salt in Case 1. The interpretability of 3-factor results was poor because Factor 1 and Factor 2 contained the parts of the same sources. In the 5-factor results, an additional factor accounting for large percent of Ca^{2+} (53%) and GOM (71%) and moderate percent of HNO_3 (30%) and NO_3^- (36%). This factor is likely related to the reaction between soil aerosols or sea salt and the HNO_3 (Zhang et al., 2008). However, the lack of the other markers such as Na^+ and Cl^- (sea salt), and Mg^{2+} or K^+ (soil derived aerosols) made it hard to interpret this factor (Zhang et al., 2008). The high percent of GOM the factor accounting for was also hard to interpret because GOM and the other species have no common sources. Therefore, factor 5 cannot be assigned to one specific source/category and left undetermined. The interpretability of 5-factor result is worse than 4-factor result Overall, the interpretability of the four-factor results was the most reasonable results in 2009.

Table C.1: Factor profiles (% of species >25%, between 15% and 25% in the bracket) of 3-factor solution in 2009.

% of total species	F1	F2	F3
GEM	80		(20)
GOM	96		
PBM	65	(20)	
PM	57	26	(17)
O_3	74		(18)
SO_2		98	
HNO_3	49	51	
Ca^{2+}	53		34
K^+	49		41
Mg^{2+}			85
Na^+			80
Cl^-			100
NO_3^-		46	38
NH_4^+	43	51	
SO_4^{2-}	44	40	(16)

Table C.2: Factor profiles (% of species >25%, between 15% and 25% in the bracket) of 5-factor solution in 2009

% of total species	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
GEM	68		(18)		
GOM	29				71
PBM	57	(24)	(17)		
PM	(23)		32		(25)
O_3	65				
SO_2		81			
HNO_3		(23)	37		30
Ca^{2+}				26	53
K^+	29		(18)	36	(17)
Mg^{2+}				82	
Na^+				77	
Cl^-				94	
NO_3^-		(25)		36	36
NH_4^+			61		(17)
SO_4^{2-}			60	(17)	

C.2 Comparison in 2010

The 3-factor results and 5-factor results of 2010 are listed in Table C.3 and Table C.4, respectively. Similar to 2009, factor 3 in the 3-factor results was a combination of the factor Combustion Emission and Industrial Source in Case 7 indicating a worse interpretability compared to 4-factor results. Similar to the 5-factor results in 2009, a factor accounting for high percent of Ca^{2+} and moderate percent of NO_3^- were identified in the 5-factor results in 2010. The Ca^{2+} and NO_3^- could be derived from several sources such as the reaction of HNO_3 with soil aerosols or marine aerosols. The factor cannot be assigned to a specific source due a lack of markers in the factor. Therefore, the interpretability of the 5-factor results was worse than that of the 4-factor results. The 4-facotr results has the best interpretability in 2010.

Table C.3: Factor profiles (% of species >25%, between 15% and 25% in the bracket) of 3-factor solution in 2010. Suggest adding factor names to aid your discussion

% of total species	Factor 1	Factor 2	Factor 3
GEM	49	(19)	32
GOM	63		36
PBM	55		32
O_3	51	(16)	33
SO_2	100		
HNO_3			89
Ca^{2+}	(24)	31	44
K^+		29	68
Mg^{2+}		84	
Na^+		77	(20)
Cl^-		100	
NO_3^-	(22)	26	52
NH_4^+			93
SO_4^{2-}		(16)	83

Table C.4: Factor profiles (% of species >25%, between 15% and 25% in the bracket) of 5-factor solution in 2010

% of total species	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
GEM	82				
GOM	59		41		
PBM	83				
O_3	83				
SO_2			96		
HNO_3		66	26		
Ca^{2+}	(17)				72
K^+	(25)	38		(18)	(19)

Table C.4 – Continued 1

% of total species	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Mg^{2+}				82	
Na^{+}				68	
Cl^{-}				99	
NO_3^{-}	(20)		30		37
NH_4^{+}		77			(23)
SO_4^{2-}		71			(18)

The 4-factor results had the best interpretability among the 3-factor, 4-factor results and 5-factor results in both years. In addition, the number of components extracted by PCA was happened to be 4. Four components were retained by PCA according to the eigenvalues. The 4-factor results would be more comparable to the PCA results than the 3-factor and 5-factor results. Therefore, 4-factor run were used in both 2009 and 2010.

Appendix D:
PMF Outputs

D.1 PMF outputs in 2009

Table D.1: Base run summary of Case 1 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	5262.3	5794.36	Yes	540
2	5262.26	5794.39	Yes	438
3	**5262.21**	**5794.39**	**Yes**	**292**
4	5262.32	5794.36	Yes	555
5	5262.35	5794.38	Yes	472
6	5262.35	5794.41	Yes	556
7	5262.39	5794.42	Yes	410
8	5262.46	5794.35	Yes	444
9	5262.22	5794.39	Yes	392
10	5262.33	5794.37	Yes	491
11	5262.56	5794.41	Yes	345
12	5262.36	5794.35	Yes	353
13	5262.35	5794.36	Yes	567
14	5262.39	5794.42	Yes	557
15	5262.44	5794.36	Yes	485
16	5262.32	5794.42	Yes	453
17	5262.34	5794.39	Yes	470
18	5262.47	5794.34	Yes	354
19	5262.43	5794.36	Yes	647
20	5262.46	5794.28	Yes	243

Table D.2: Factor profiles (concentrations of species, µg/m^3 for non-mercury species) from base run #3 (convergent run) of Case 1.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.05	0.08	0.19	1.05
GOM (ng/m^3)	0.62	0.00	0.00	1.39
PBM (ng/m^3)	0.00	0.97	0.27	2.11
PM	1.24	0.21	0.50	1.01
O_3	3.02	6.67	8.11	45.62
SO_2	0.01	0.35	0.01	0.05
HNO_3	0.08	0.03	0.00	0.04
Ca^{2+}	0.01	0.00	0.02	0.02
K^+	0.01	0.00	0.02	0.02
Na^+	0.05	0.02	0.43	0.00
Mg^{2+}	0.01	0.00	0.05	0.00
Cl^-	0.00	0.00	0.54	0.00
NO_3^-	0.05	0.05	0.09	0.03
NH_4^+	0.17	0.02	0.02	0.03
SO_4^{2-}	0.58	0.06	0.16	0.10

Table D.3: Regression diagnostics of base run #3 for Case 1.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.59	0.00	0.28	0.05	0.84
GOM	0.00	0.09	0.00	0.23	0.29	0.00
PBM	0.00	0.39	0.00	0.57	0.10	0.06
PM	1.13	0.60	0.90	0.73	0.07	0.39
O_3	15.10	0.74	7.08	0.79	0.05	0.79
SO_2	0.00	0.98	0.07	0.98	0.22	0.00
HNO_3	0.06	0.51	0.06	0.81	0.12	0.02
Ca^{2+}	0.03	0.25	0.01	0.50	0.09	0.16
K^+	0.02	0.46	0.02	0.44	0.15	0.00
Na^+	0.00	0.99	0.05	0.99	0.04	0.98
Mg^{2+}	0.00	1.00	0.01	0.98	0.12	0.03
Cl^-	0.04	0.87	0.12	0.97	0.04	0.98
NO_3^-	0.11	0.33	0.09	0.74	0.10	0.09
NH_4^+	0.00	0.97	0.05	0.95	0.05	0.74
SO_4^{2-}	-0.05	1.02	0.21	0.92	0.06	0.66

Table D.4: Factor profiles (concentrations of species, µg/m3 for non-mercury species) from base run #20 (convergent run) of Case 2.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.05	0.08	0.19	1.05
GOM (ng/m^3)	0.62	0.00	0.00	1.39
PBM (ng/m^3)	0.00	0.97	0.27	2.11
PM	1.24	0.21	0.50	1.01
O_3	3.02	6.67	8.11	45.62
SO_2	0.01	0.35	0.01	0.05
HNO_3	0.08	0.03	0.00	0.04
Ca^{2+}	0.01	0.00	0.02	0.02
K^+	0.01	0.00	0.02	0.02
Na^+	0.05	0.02	0.43	0.00
Mg^{2+}	0.01	0.00	0.05	0.00
Cl^-	0.00	0.00	0.54	0.00
NO_3^-	0.05	0.05	0.09	0.03
NH_4^+	0.17	0.02	0.02	0.03
SO_4^{2-}	0.58	0.06	0.16	0.10

Table D.5: Regression diagnostics of base run #20 for Case 2.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.57	0.00	0.17	0.10	0.00
GOM	0.00	0.05	0.00	0.08	0.31	0.00
PBM	0.00	0.32	0.00	0.33	0.19	0.00
PM	1.27	0.56	0.92	0.65	0.02	0.99
O_3	17.99	0.65	10.51	0.59	0.06	0.10
SO_2	0.01	0.96	0.07	0.98	0.22	0.00
HNO_3	0.05	0.60	0.07	0.79	0.08	0.02
Ca^{2+}	0.04	0.21	0.01	0.43	0.10	0.00
K^+	0.02	0.49	0.01	0.47	0.10	0.00

Table D.5 – Continued

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
Na^+	0.00	1.01	0.06	0.99	0.09	0.00
Mg^{2+}	0.00	1.00	0.01	0.97	0.11	0.00
Cl^-	0.04	0.82	0.12	0.95	0.06	0.20
NO_3^-	0.14	0.20	0.07	0.56	0.12	0.00
NH_4^+	0.02	0.90	0.06	0.96	0.02	0.98
SO_4^{2-}	0.16	0.80	0.33	0.90	0.06	0.09

Table D.6: Base run summary of Case 3 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	12569.5	13961.1	Yes	725
2	12570.8	13960.7	Yes	491
3	12569.7	13961.1	Yes	866
4	12571.1	13960.8	Yes	609
5	12570.5	13960.9	Yes	571
6	12569.3	13961.1	Yes	429
7	12570.5	13961.1	Yes	792
8	12570.1	13961.1	Yes	511
9	12570.1	13961	Yes	538
10	12570.6	13960.8	Yes	551
11	12569.4	13961.2	Yes	567
12	**12569.2**	**13961.3**	**Yes**	**677**
13	12569.7	13961	Yes	643
14	12570.1	13961	Yes	914
15	12571.1	13961	Yes	558
16	12569.3	13961.1	Yes	621
17	12570.9	13961.1	Yes	686
18	12570.9	13960.9	Yes	361
19	12569.8	13961	Yes	765
20	12569.4	13961.2	Yes	789

Table D.7 Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #12 (convergent run) of Case 3.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.07	1.02	0.16	0.03
GOM (ng/m^3)	0.25	1.35	0.00	0.01
PBM (ng/m^3)	0.00	1.88	0.26	0.74
PM	1.20	1.33	0.45	0.17
O_3	5.33	40.15	7.27	5.86
SO_2	0.01	0.05	0.00	0.32
HNO_3	0.10	0.03	0.00	0.03
Ca^{2+}	0.01	0.03	0.01	0.00
K^+	0.01	0.02	0.01	0.00
Na^+	0.02	0.01	0.37	0.03
Mg^{2+}	0.00	0.00	0.04	0.00
Cl^-	0.00	0.00	0.42	0.00
NO_3^-	0.02	0.07	0.07	0.03
NH_4^+	0.23	0.01	0.03	0.01
SO_4^{2-}	0.84	0.00	0.17	0.05

Table D.8: Regression diagnostics of base run #12 for Case 3.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.54	0.00	0.15	0.10	0.00
GOM	0.00	0.05	0.00	0.09	0.32	0.00
PBM	0.00	0.34	0.00	0.34	0.19	0.00
PM	1.28	0.56	0.93	0.65	0.03	0.91
O_3	18.08	0.65	10.60	0.58	0.07	0.07
SO_2	0.01	0.96	0.07	0.98	0.22	0.00
HNO_3	0.05	0.60	0.07	0.79	0.08	0.02
Ca^{2+}	0.04	0.21	0.01	0.43	0.10	0.00
K^+	0.02	0.49	0.01	0.47	0.10	0.00
Na^+	0.00	1.01	0.06	0.99	0.09	0.01
Mg^{2+}	0.00	1.00	0.01	0.97	0.11	0.00
Cl^-	0.04	0.82	0.12	0.95	0.06	0.21
NO_3^-	0.14	0.20	0.07	0.56	0.12	0.00
NH_4^+	0.02	0.89	0.06	0.96	0.02	0.98
SO_4^{2-}	0.16	0.79	0.33	0.90	0.06	0.11

Table D.9: Base run summary of Case 4 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	5737.46	6335.91	Yes	323
2	5737.61	6335.93	Yes	781
3	5737.49	6335.94	Yes	583
4	5737.54	6335.95	Yes	656
5	5737.45	6335.93	Yes	585
6	5737.6	6335.93	Yes	480
7	**5737.38**	**6335.9**	**Yes**	**297**
8	5737.5	6335.9	Yes	443
9	5737.61	6335.92	Yes	601
10	5737.54	6335.93	Yes	653
11	5737.5	6335.92	Yes	749
12	5737.6	6335.97	Yes	667
13	5737.53	6335.92	Yes	599
14	5737.61	6335.95	Yes	599
15	5737.4	6335.96	Yes	502
16	5737.5	6335.96	Yes	631
17	5737.55	6335.95	Yes	547
18	5737.44	6335.93	Yes	538
19	5737.46	6335.95	Yes	265
20	5737.51	6335.89	Yes	556

Table D.10: Factor profiles (concentrations of species, µg/m^3 for non-mercury species) from base run #7 (convergent run) of Case 4.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.19	1.09	0.10	0.02
PM	0.47	1.12	1.21	0.15
O_3	8.18	50.36	3.72	4.81
SO_2	0.00	0.06	0.01	0.36
HNO_3	0.00	0.04	0.08	0.03
Ca^{2+}	0.01	0.03	0.01	0.00

Table D.10 – Continued

Species	Factor 1	Factor 2	Factor 3	Factor 4
K^+	0.02	0.01	0.01	0.00
Na^+	0.45	0.00	0.05	0.02
Mg^{2+}	0.05	0.00	0.01	0.00
Cl^-	0.54	0.00	0.00	0.00
NO_3^-	0.09	0.03	0.06	0.05
NH_4^+	0.02	0.03	0.17	0.02
SO_4^{2-}	0.16	0.13	0.59	0.04

Table D.11: Regression diagnostics of base run #7 for Case 4.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.59	0.00	0.25	0.04	0.87
PM	1.16	0.59	0.87	0.71	0.06	0.43
O_3	15.01	0.75	7.40	0.78	0.05	0.74
SO_2	-0.01	1.03	0.08	0.98	0.21	0.00
HNO_3	0.06	0.51	0.05	0.80	0.11	0.01
Ca^{2+}	0.03	0.24	0.01	0.48	0.09	0.08
K^+	0.02	0.47	0.02	0.45	0.13	0.00
Na^+	0.00	1.00	0.06	0.99	0.05	0.73
Mg^{2+}	0.00	1.00	0.01	0.98	0.10	0.03
Cl^-	0.04	0.87	0.12	0.97	0.03	1.00
NO_3^-	0.12	0.32	0.10	0.71	0.10	0.03
NH_4^+	0.01	0.95	0.04	0.96	0.04	0.83
SO_4^{2-}	-0.03	0.99	0.20	0.92	0.05	0.65

Table D.12: Base run summary of Case 5 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	4830.67	5289.54	Yes	623
2	4830.64	5289.5	Yes	622
3	4830.72	5289.56	Yes	554
4	**4830.47**	**5289.55**	**Yes**	**487**
5	4830.74	5289.58	Yes	635
6	4830.62	5289.56	Yes	641
7	4830.68	5289.54	Yes	682
8	4830.74	5289.52	Yes	633
9	4830.64	5289.56	Yes	521
10	4830.64	5289.58	Yes	530
11	4830.66	5289.56	Yes	547
12	4830.71	5289.57	Yes	734
13	4830.73	5289.54	Yes	534
14	4830.62	5289.5	Yes	391
15	4830.69	5289.52	Yes	618
16	4830.65	5289.52	Yes	449
17	4830.67	5289.54	Yes	588
18	4830.67	5289.58	Yes	472
19	4830.53	5289.48	Yes	233
20	4830.8	5289.54	Yes	538

Table D.13: Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #4 (convergent run) of Case 5.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.05	0.08	0.18	1.07
RM (ng/m^3)	0.65	1.18	0.02	4.00
PM	1.31	0.11	0.51	1.04
O_3	2.34	6.92	7.76	46.51
SO_2	0.01	0.36	0.00	0.06
HNO_3	0.08	0.03	0.00	0.04
Ca^{2+}	0.01	0.00	0.01	0.02
K^+	0.01	0.00	0.02	0.02
Na^+	0.05	0.02	0.43	0.00
Mg^{2+}	0.01	0.00	0.05	0.00
Cl^-	0.00	0.00	0.54	0.00
NO_3^-	0.05	0.05	0.09	0.03
NH_4^+	0.18	0.01	0.02	0.03
SO_4^{2-}	0.62	0.00	0.17	0.11

Table D.14: Regression diagnostics of base run #4 for Case 5.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.59	0.00	0.29	0.04	0.97
RM	0.00	0.31	0.00	0.48	0.17	0.00
PM	1.14	0.59	0.90	0.73	0.07	0.42
O_3	14.08	0.75	7.26	0.78	0.05	0.88
SO_2	-0.02	1.05	0.10	0.98	0.23	0.00
HNO_3	0.06	0.53	0.05	0.83	0.12	0.03
Ca^{2+}	0.03	0.25	0.01	0.50	0.09	0.12
K^+	0.02	0.45	0.02	0.44	0.14	0.00
Na^+	0.00	1.00	0.05	0.99	0.03	1.00
Mg^{2+}	0.00	1.01	0.01	0.98	0.11	0.05
Cl^-	0.04	0.87	0.11	0.97	0.03	0.99
NO_3^-	0.11	0.34	0.09	0.74	0.10	0.07
NH_4^+	0.00	0.97	0.05	0.95	0.05	0.90
SO_4^{2-}	-0.04	1.01	0.20	0.93	0.06	0.71

Table D.15: Base run summary of Case 6 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	5525.95	6034.06	Yes	484
2	5525.98	6034.06	Yes	459
3	5525.99	6034.13	Yes	489
4	5525.94	6034.1	Yes	471
5	5525.97	6034.07	Yes	633
6	5525.99	6033.99	Yes	253
7	5525.93	6034.07	Yes	455
8	5526.01	6034.04	Yes	399
9	5525.94	6034.14	Yes	451
10	**5525.9**	**6034.02**	**Yes**	**439**
11	5525.96	6034.07	Yes	382
12	5526	6034.06	Yes	288
13	5525.93	6034.08	Yes	466
14	5525.98	6034.04	Yes	565

Table D.15 – Continued

Run #	Q(Robust)	Q(True)	Converged	# Steps
15	5525.94	6034.04	Yes	511
16	5526.01	6034.07	Yes	540
17	5525.94	6034.06	Yes	507
18	5525.98	6034.12	Yes	411
19	5525.99	6034.05	Yes	784
20	5526.01	6034.02	Yes	396

Table D.16: Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #10 (convergent run) of Case 6.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.00	0.11	0.20	1.07
GOM (ng/m^3)	1.15	0.00	0.04	3.09
PBM (ng/m^3)	0.00	1.25	0.62	4.51
PM	1.18	0.21	0.50	1.07
O_3	0.15	8.13	8.65	46.48
SO_2	0.00	0.36	0.01	0.05
HNO_3	0.08	0.03	0.00	0.04
Ca^{2+}	0.01	0.00	0.02	0.02
K^+	0.01	0.00	0.02	0.02
Na^+	0.05	0.02	0.43	0.00
Mg^{2+}	0.01	0.00	0.05	0.00
Cl^-	0.00	0.00	0.54	0.00
NO_3^-	0.05	0.05	0.09	0.03
NH_4^+	0.17	0.02	0.02	0.03
SO_4^{2-}	0.58	0.05	0.16	0.11

Table D.17: Regression diagnostics of base run #10 for Case 6.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	0.58	0.00	0.28	0.05	0.82
GOM	0.00	0.18	0.00	0.33	0.18	0.00
PBM	0.00	0.48	0.00	0.59	0.11	0.04
PM	1.15	0.59	0.90	0.72	0.07	0.39
O_3	15.47	0.73	7.36	0.77	0.06	0.67
SO_2	0.00	0.98	0.06	0.99	0.23	0.00
HNO_3	0.06	0.52	0.05	0.82	0.11	0.03
Ca^{2+}	0.03	0.25	0.01	0.49	0.08	0.32
K^+	0.02	0.46	0.02	0.44	0.15	0.00
Na^+	0.00	0.99	0.05	0.99	0.04	0.98
Mg^{2+}	0.00	1.00	0.01	0.98	0.12	0.03
Cl^-	0.04	0.87	0.12	0.97	0.04	0.98
NO_3^-	0.11	0.33	0.09	0.73	0.10	0.07
NH_4^+	0.00	0.98	0.05	0.95	0.06	0.63
SO_4^{2-}	-0.05	1.02	0.21	0.92	0.06	0.55

D.2 PMF outputs in 2010

Table D.18: Base run summary of Case 7 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	**11165.2**	**14202.1**	**Yes**	**656**
2	11165.9	14201.7	Yes	712
3	11165.6	14202.1	Yes	715
4	11165.9	14201.7	Yes	657
5	11166	14201.7	Yes	494
6	11166.1	14202.8	Yes	774
7	11166.1	14201.5	Yes	694
8	11166.3	14201.6	Yes	422
9	11165.7	14202	Yes	832
10	11165.7	14202	Yes	621
11	11166.2	14203.2	Yes	520
12	11166.4	14201.6	Yes	614
13	11166.4	14203.3	Yes	661
14	11167	14203	Yes	538
15	11166.5	14201.5	Yes	1026
16	11166.1	14203.3	Yes	488
17	11166.3	14201.7	Yes	530
18	11166.1	14201.7	Yes	862
19	11166.2	14202.4	Yes	715
20	11165.7	14201.9	Yes	894

Table D.19: Factor profiles (concentrations of species, µg/m^3 for non-mercury species) from base run #1 (convergent run) of Case7.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.12	0.01	0.14	1.03
GOM (ng/m^3)	0.00	0.24	0.03	0.39
PBM (ng/m^3)	0.11	0.10	0.27	2.08
O_3	4.62	1.16	6.96	50.20
SO_2	0.00	0.18	0.00	0.01
HNO_3	0.00	0.04	0.10	0.02
Ca^{2+}	0.01	0.01	0.00	0.01
K^+	0.01	0.00	0.02	0.01
Na^+	0.33	0.04	0.03	0.00
Mg^{2+}	0.04	0.01	0.01	0.00
Cl^-	0.42	0.00	0.00	0.00
NO_3^-	0.03	0.07	0.03	0.04
NH_4^+	0.01	0.02	0.22	0.00
SO_4^{2-}	0.13	0.07	0.77	0.00

Table D.20: Regression diagnostics of base run #1 for Case 7.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.29	0.00	0.46	0.05	0.36
GOM	0.00	0.29	0.00	0.31	0.20	0.00
PBM	0.00	0.09	0.00	0.13	0.18	0.00
O_3	10.79	0.81	6.96	0.80	0.07	0.12
SO_2	0.07	0.54	0.17	0.48	0.23	0.00
HNO_3	0.05	0.60	0.09	0.67	0.11	0.00
Ca^{2+}	0.03	0.03	0.02	0.06	0.04	0.62
K^+	0.03	0.23	0.03	0.26	0.23	0.00
Na^+	-0.01	1.01	0.08	0.98	0.12	0.00
Mg^{2+}	0.00	0.95	0.02	0.88	0.20	0.00
Cl^-	0.03	0.83	0.11	0.98	0.08	0.04
NO_3^-	0.13	0.16	0.10	0.27	0.12	0.00
NH_4^+	0.14	0.39	0.21	0.52	0.11	0.00
SO_4^{2-}	0.41	0.50	0.64	0.63	0.11	0.00

Table D.21: Base run summary of Case 8 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	13153.3	16302.6	Yes	952
2	13153.5	16300.9	Yes	928
3	13151.2	16366.4	Yes	932
4	13152.6	16305.4	Yes	1013
5	13156.9	16292	Yes	815
6	13153.7	16309.2	Yes	911
7	13152.4	16311.1	Yes	878
8	13152.9	16310.8	Yes	779
9	13153	16310.6	Yes	904
10	13153	16311.6	Yes	610
11	13152.9	16311.3	Yes	686
12	13152.7	16310.8	Yes	962
13	13150.9	16365.9	Yes	811
14	13148.4	16370.5	Yes	470
15	13152.8	16311	Yes	796
16	13151.9	16310.6	Yes	1010
17	13153.1	16305.1	Yes	991
18	13153	16308.9	Yes	981
19	13152.6	16311.2	Yes	803
20	**13148**	**16370.8**	**Yes**	**572**

Table D.22: Factor profiles (concentrations of species, µg/m^3 for non-mercury species) from base run #20 (convergent run) of Case 8.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.03	0.08	0.11	1.07
GOM (ng/m^3)	0.01	0.20	0.00	0.37
PBM (ng/m^3)	0.00	0.11	0.05	2.75
O_3	1.61	4.32	4.14	51.23
SO_2	0.01	0.19	0.00	0.00
HNO_3	0.10	0.03	0.00	0.01
Ca^{2+}	0.00	0.00	0.01	0.02

Table D.22 – Continued

Species	Factor 1	Factor 2	Factor 3	Factor 4
K^+	0.02	0.00	0.01	0.01
Na^+	0.03	0.01	0.29	0.02
Mg^{2+}	0.01	0.00	0.04	0.00
Cl^-	0.00	0.00	0.35	0.00
NO_3^-	0.05	0.03	0.03	0.06
NH_4^+	0.21	0.01	0.01	0.00
SO_4^{2-}	0.75	0.02	0.12	0.01

Table D.23: Regression diagnostics of base run #20 for Case 8.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.26	0.00	0.32	0.08	0.03
GOM	0.00	0.22	0.00	0.23	0.22	0.00
PBM	0.00	0.09	0.00	0.15	0.17	0.00
O_3	10.72	0.79	11.79	0.55	0.11	0.00
SO_2	0.02	0.91	0.05	0.96	0.26	0.00
HNO_3	0.05	0.61	0.08	0.72	0.13	0.00
Ca^{2+}	0.03	0.02	0.01	0.05	0.07	0.05
K^+	0.03	0.22	0.02	0.25	0.26	0.00
Na^+	0.00	1.00	0.07	0.98	0.13	0.00
Mg^{2+}	0.00	0.98	0.02	0.88	0.14	0.00
Cl^-	0.04	0.78	0.10	0.97	0.11	0.00
NO_3^-	0.14	0.14	0.08	0.25	0.14	0.00
NH_4^+	0.14	0.34	0.18	0.49	0.10	0.00
SO_4^{2-}	0.44	0.45	0.57	0.58	0.19	0.00

Table D.24: Base run summary of Case 9 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	12990.9	16292.9	Yes	879
2	**12989.6**	**16292**	**Yes**	**909**
3	12991.4	16290.9	Yes	820
4	12990	16292.5	Yes	1072
5	12991.3	16290	Yes	912
6	12990.1	16292.2	Yes	885
7	12991.2	16289.8	Yes	750
8	12989.7	16291.5	Yes	942
9	12990	16291.9	Yes	1099
10	12992	16290.7	Yes	588
11	12990.9	16290.3	Yes	870
12	12992.1	16290.7	Yes	992
13	12991.6	16290.7	Yes	942
14	12990	16291.4	Yes	670
15	12991.5	16290.7	Yes	911
16	12990	16292.5	Yes	1040
17	12991	16291.3	Yes	884
18	12990.7	16292.8	Yes	1076
19	12989.7	16292.2	Yes	887
20	12990.4	16291.7	Yes	961

Table D.25: Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #2 (convergent run) of Case 9.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.04	0.00	0.11	1.16
GOM (ng/m^3)	0.00	0.25	0.00	0.32
PBM (ng/m^3)	0.00	0.00	0.06	2.89
O_3	2.09	0.65	4.22	55.52
SO_2	0.00	0.17	0.00	0.00
HNO_3	0.09	0.04	0.00	0.02
Ca^{2+}	0.00	0.01	0.01	0.02
K^+	0.02	0.00	0.01	0.01
Na^+	0.02	0.04	0.29	0.02
Mg^{2+}	0.01	0.01	0.04	0.00
Cl^-	0.00	0.00	0.35	0.00
NO_3^-	0.02	0.08	0.03	0.04
NH_4^+	0.20	0.02	0.01	0.00
SO_4^{2-}	0.72	0.07	0.12	0.01

Table D.26: Regression diagnostics of base run #2 for Case 9.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.27	0.00	0.41	0.05	0.40
GOM	0.00	0.27	0.00	0.27	0.22	0.00
PBM	0.00	0.08	0.00	0.15	0.17	0.00
O_3	11.89	0.78	8.81	0.69	0.10	0.00
SO_2	0.08	0.47	0.15	0.42	0.26	0.00
HNO_3	0.05	0.56	0.09	0.62	0.15	0.00
Ca^{2+}	0.03	0.03	0.02	0.05	0.05	0.41
K^+	0.03	0.24	0.02	0.26	0.26	0.00
Na^+	-0.01	1.01	0.07	0.98	0.15	0.00
Mg^{2+}	0.00	0.98	0.02	0.88	0.14	0.00
Cl^-	0.04	0.78	0.10	0.97	0.12	0.00
NO_3^-	0.13	0.18	0.09	0.28	0.18	0.00
NH_4^+	0.12	0.40	0.19	0.54	0.09	0.00
SO_4^{2-}	0.37	0.53	0.59	0.65	0.18	0.00

Table D.27: Base run summary of Case 10 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	10632.2	13629.6	Yes	905
2	10631	13629.3	Yes	537
3	10631	13629.5	Yes	644
4	10632.4	13629.1	Yes	621
5	10631.6	13629.8	Yes	646
6	10631.7	13628.7	Yes	834
7	10631.2	13629	Yes	814
8	10630.9	13629.4	Yes	696
9	10631.6	13630.7	Yes	710
10	10632	13628.7	Yes	698
11	10631.4	13628.8	Yes	673
12	10631	13629.3	Yes	706
13	10631.2	13629.1	Yes	964
14	10631.5	13630.2	Yes	611

Table D.27 – Continued

Run #	Q(Robust)	Q(True)	Converged	# Steps
15	10631.2	13629.3	Yes	820
16	10631.5	13628.8	Yes	719
17	**10630.9**	**13629.2**	**Yes**	**767**
18	10631.5	13629	Yes	756
19	10631.1	13629.1	Yes	797
20	10631.1	13629.4	Yes	820

Table D.28: Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #17 (convergent run) of Case 10.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	1.03	0.15	0.01	0.12
O_3	50.00	7.00	1.31	4.60
SO_2	0.01	0.00	0.18	0.00
HNO_3	0.02	0.10	0.04	0.00
Ca^{2+}	0.01	0.00	0.01	0.01
K^+	0.01	0.02	0.00	0.01
Na^+	0.00	0.03	0.04	0.33
Mg^{2+}	0.00	0.01	0.01	0.04
Cl^-	0.00	0.00	0.00	0.42
NO_3^-	0.04	0.03	0.07	0.03
NH_4^+	0.00	0.22	0.02	0.01
SO_4^{2-}	0.00	0.77	0.07	0.13

Table D.29: Regression diagnostics of base run #17 for Case 10.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.31	0.00	0.47	0.05	0.45
O_3	10.43	0.81	7.17	0.79	0.07	0.08
SO_2	0.07	0.54	0.17	0.48	0.23	0.00
HNO_3	0.05	0.60	0.09	0.67	0.11	0.00
Ca^{2+}	0.03	0.03	0.02	0.06	0.05	0.51
K^+	0.03	0.23	0.03	0.26	0.24	0.00
Na^+	-0.01	1.01	0.08	0.98	0.12	0.00
Mg^{2+}	0.00	0.95	0.02	0.88	0.21	0.00
Cl^-	0.03	0.83	0.11	0.98	0.08	0.03
NO_3^-	0.13	0.16	0.10	0.27	0.12	0.00
NH_4^+	0.14	0.39	0.21	0.53	0.11	0.00
SO_4^{2-}	0.41	0.50	0.64	0.63	0.11	0.00

Table D.30: Base run summary of Case 11 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	10851	13855.3	Yes	885
2	10851	13855.1	Yes	792
3	**10850**	**13855.7**	**Yes**	**816**
4	10850.9	13855.9	Yes	579
5	10850.6	13855.4	Yes	726
6	10851.2	13855.1	Yes	704
7	10850.9	13854.3	Yes	506
8	10851.5	13854.6	Yes	492
9	10850.5	13855.2	Yes	649

Table D.30 – Continued

Run #	Q(Robust)	Q(True)	Converged	# Steps
10	10850.8	13857.5	Yes	639
11	10850.3	13855.2	Yes	839
12	10851.2	13855	Yes	673
13	10851.2	13855.4	Yes	838
14	10850.2	13855.4	Yes	748
15	10851.3	13856.8	Yes	474
16	10851.2	13856.5	Yes	698
17	10850.3	13855.4	Yes	657
18	10850.4	13855.3	Yes	678
19	13587.9	17924.4	Yes	593
20	10851.2	13856.1	Yes	720

Table D.31: Factor profiles (concentrations of species, μg/m^3 for non-mercury species) from base run #3 (convergent run) of Case 11.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.15	0.12	0.01	1.03
RM (ng/m^3)	0.30	0.09	0.29	2.90
O_3	6.99	4.61	1.18	50.20
SO_2	0.00	0.00	0.18	0.01
HNO_3	0.10	0.00	0.04	0.02
Ca^{2+}	0.00	0.01	0.01	0.01
K^+	0.02	0.01	0.00	0.01
Na^+	0.03	0.33	0.04	0.00
Mg^{2+}	0.01	0.04	0.01	0.00
Cl^-	0.00	0.42	0.00	0.00
NO_3^-	0.03	0.03	0.07	0.04
NH_4^+	0.22	0.01	0.02	0.00
SO_4^{2-}	0.77	0.13	0.07	0.00

Table D.32: Regression diagnostics of base run #3 for Case 11.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.31	0.00	0.46	0.05	0.50
RM	0.00	0.15	0.00	0.19	0.18	0.00
O_3	10.19	0.82	7.10	0.79	0.08	0.06
SO_2	0.07	0.54	0.17	0.48	0.23	0.00
HNO_3	0.05	0.60	0.09	0.67	0.11	0.00
Ca^{2+}	0.03	0.03	0.02	0.06	0.05	0.53
K^+	0.03	0.23	0.03	0.26	0.24	0.00
Na^+	-0.01	1.01	0.08	0.98	0.12	0.00
Mg^{2+}	0.00	0.95	0.02	0.88	0.21	0.00
Cl^-	0.03	0.83	0.11	0.98	0.08	0.04
NO_3^-	0.13	0.16	0.10	0.27	0.12	0.00
NH_4^+	0.14	0.39	0.21	0.53	0.11	0.00
SO_4^{2-}	0.41	0.50	0.64	0.63	0.11	0.00

Table D.33: Base run summary of Case 12 (bold indicates the run used for further analysis).

Run #	Q(Robust)	Q(True)	Converged	# Steps
1	11488.7	14572.9	Yes	811
2	**11488.2**	**14573.2**	**Yes**	**569**
3	11488.5	14573.1	Yes	654
4	11489	14574.7	Yes	579
5	11488.8	14573	Yes	552
6	11488.8	14572.7	Yes	602
7	11489.3	14572.5	Yes	527
8	11489.3	14572.9	Yes	520
9	11489.2	14572.7	Yes	769
10	11488.6	14573.2	Yes	761
11	11489.2	14573.9	Yes	531
12	11488.3	14573.2	Yes	799
13	11489.2	14573.6	Yes	764
14	11489.3	14573.5	Yes	865
15	11488.4	14573.2	Yes	760
16	11489.3	14572.5	Yes	930
17	11489	14572.7	Yes	548
18	11489.3	14572.5	Yes	796
19	11489.1	14573	Yes	893
20	11489.4	14572.7	Yes	1011

Table D.34: Factor profiles (concentrations of species, $\mu g/m^3$ for non-mercury species) from base run #2 (convergent run) of Case 12.

Species	Factor 1	Factor 2	Factor 3	Factor 4
GEM (pg/m^3)	0.12	0.01	0.15	1.03
GOM (ng/m^3)	0.01	0.23	0.12	1.10
PBM (ng/m^3)	0.39	0.43	0.80	5.15
O_3	4.66	1.16	7.06	50.02
SO_2	0.00	0.17	0.00	0.01
HNO_3	0.00	0.04	0.10	0.02
Ca^{2+}	0.01	0.01	0.00	0.01
K^+	0.01	0.00	0.02	0.01
Na^+	0.33	0.04	0.03	0.00
Mg^{2+}	0.04	0.01	0.01	0.00
Cl^-	0.42	0.00	0.00	0.00
NO_3^-	0.03	0.07	0.03	0.04
NH_4^+	0.01	0.02	0.22	0.00
SO_4^{2-}	0.13	0.07	0.77	0.00

Table D.35: Regression diagnostics of base run #2 for Case 12.

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
GEM	0.00	1.19	0.00	0.44	0.05	0.51
GOM	0.00	0.33	0.00	0.42	0.12	0.00
PBM	0.00	0.24	0.00	0.25	0.04	0.82
O_3	13.04	0.77	6.55	0.80	0.06	0.30
SO_2	0.07	0.54	0.17	0.48	0.23	0.00
HNO_3	0.05	0.60	0.09	0.67	0.11	0.00
Ca^{2+}	0.03	0.03	0.02	0.06	0.04	0.71
K^+	0.03	0.23	0.03	0.26	0.23	0.00

Table D.35 – Continued

Species	Intercept	Slope	SE	R^2	KS Test Stat	KS Test P Value
Na^+	-0.01	1.01	0.08	0.98	0.12	0.00
Mg^{2+}	0.00	0.95	0.02	0.89	0.20	0.00
Cl^-	0.03	0.83	0.11	0.98	0.08	0.05
NO_3^-	0.13	0.16	0.10	0.27	0.12	0.00
NH_4^+	0.14	0.39	0.21	0.52	0.11	0.00
SO_4^{2-}	0.42	0.50	0.64	0.63	0.11	0.00

D.3 PMF Obs/Pred Scatter Plot and Obs/Pred Time series in 2009

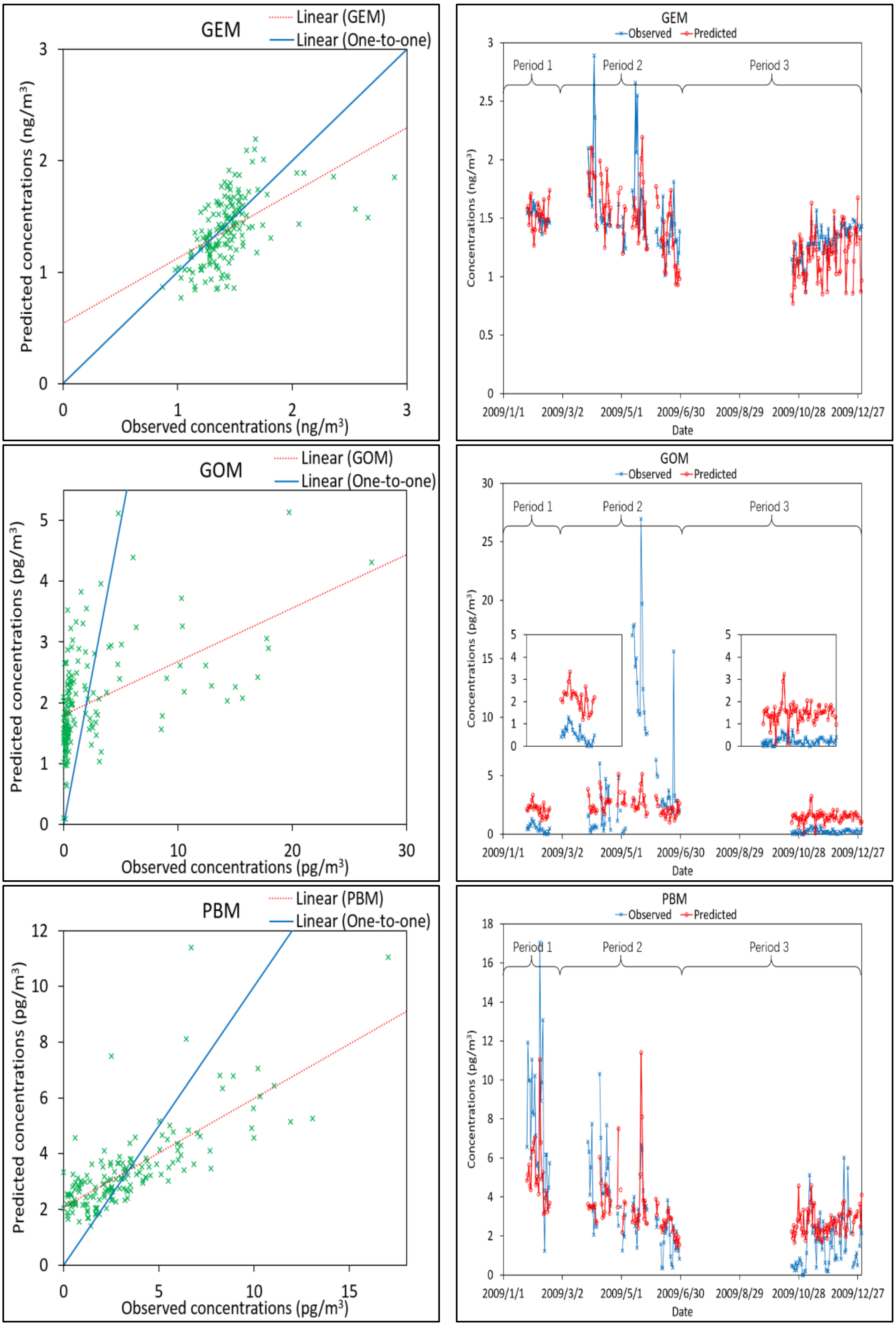

Figure D.1: Obs/Pred scatter plot and Obs/Pred time series in Case 1 in 2009

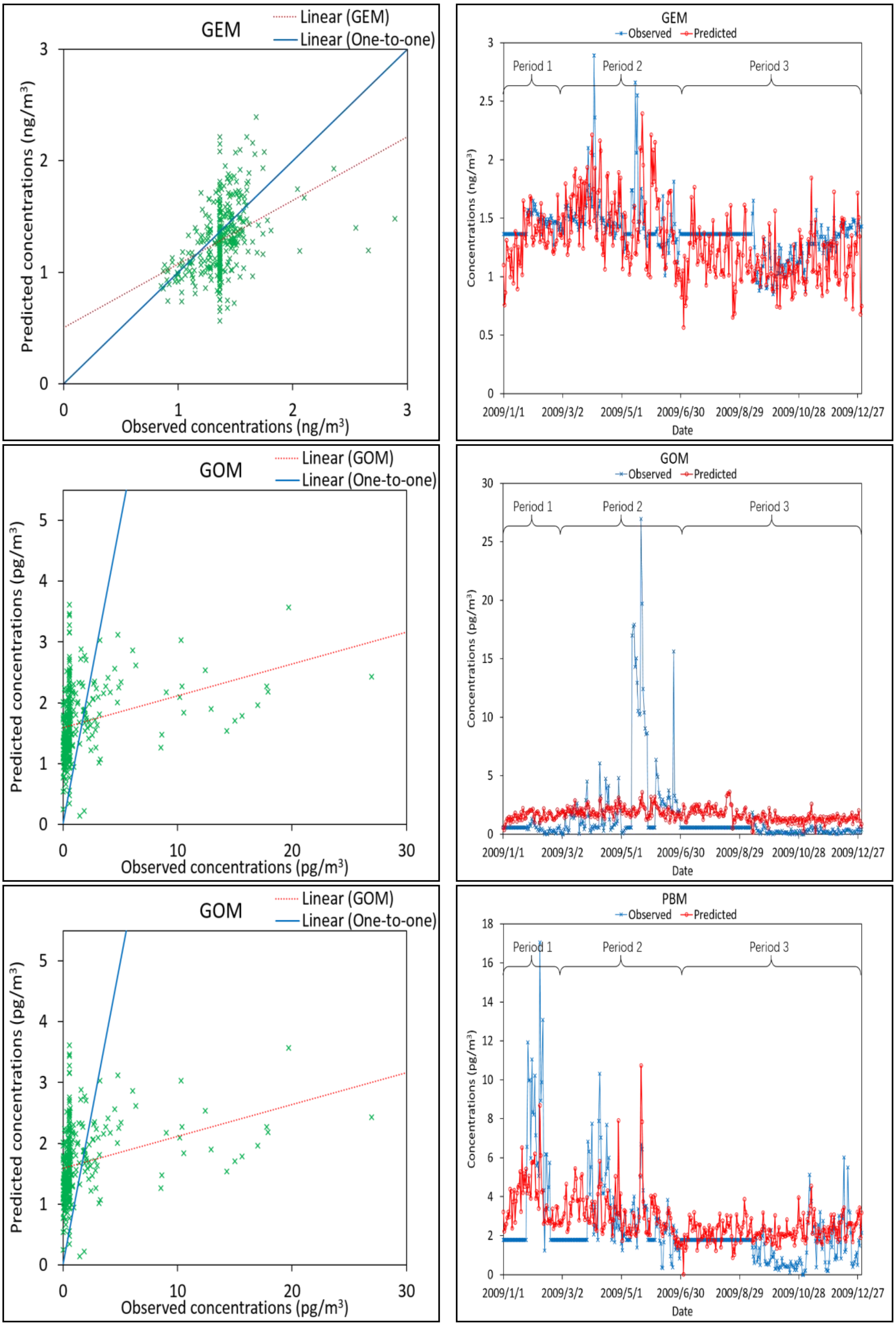

Figure D.2: Obs/Pred scatter plot and Obs/Pred time series in Case 2 in 2009

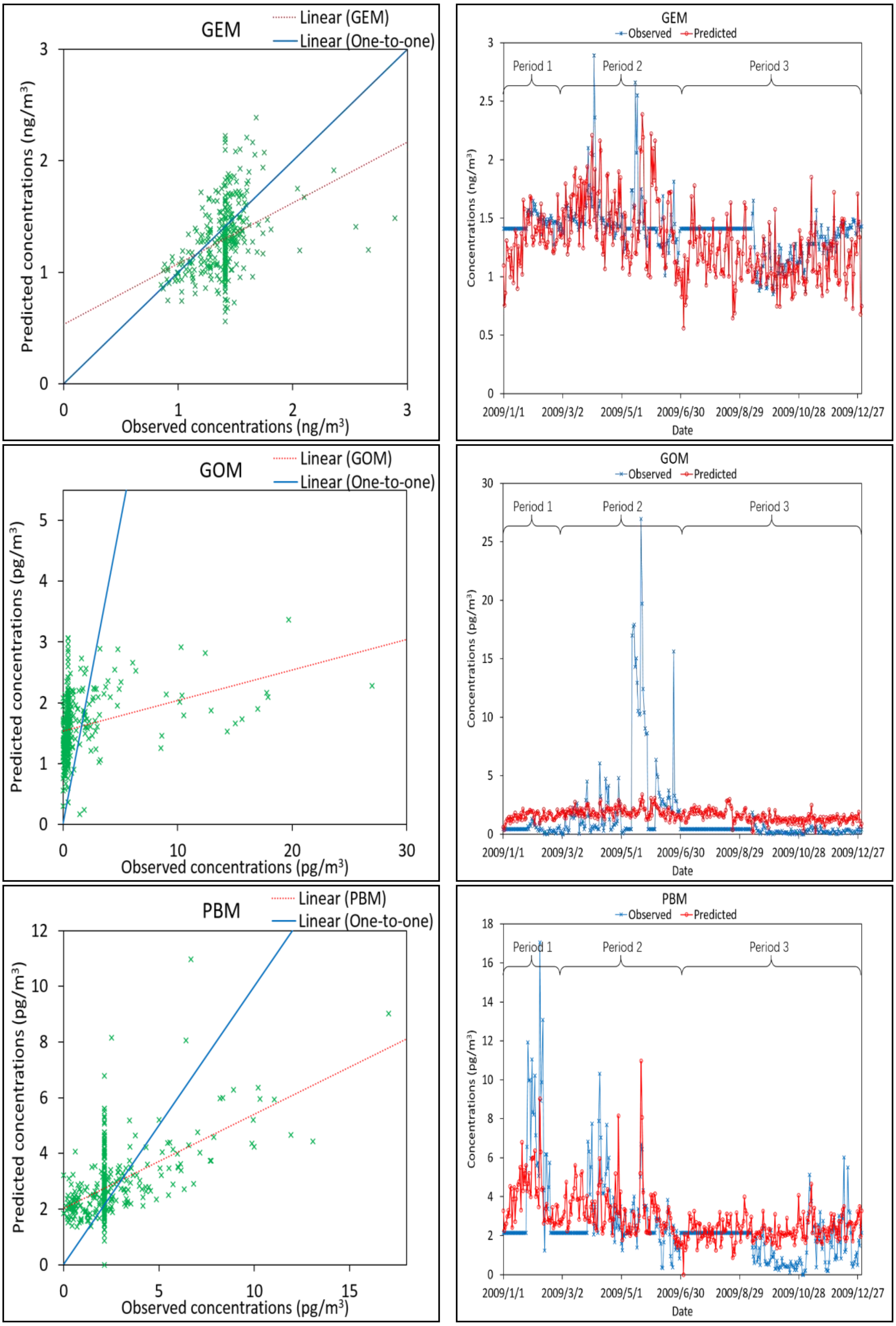

Figure D.3: Obs/Pred scatter plot and Obs/Pred time series in Case 3 in 2009

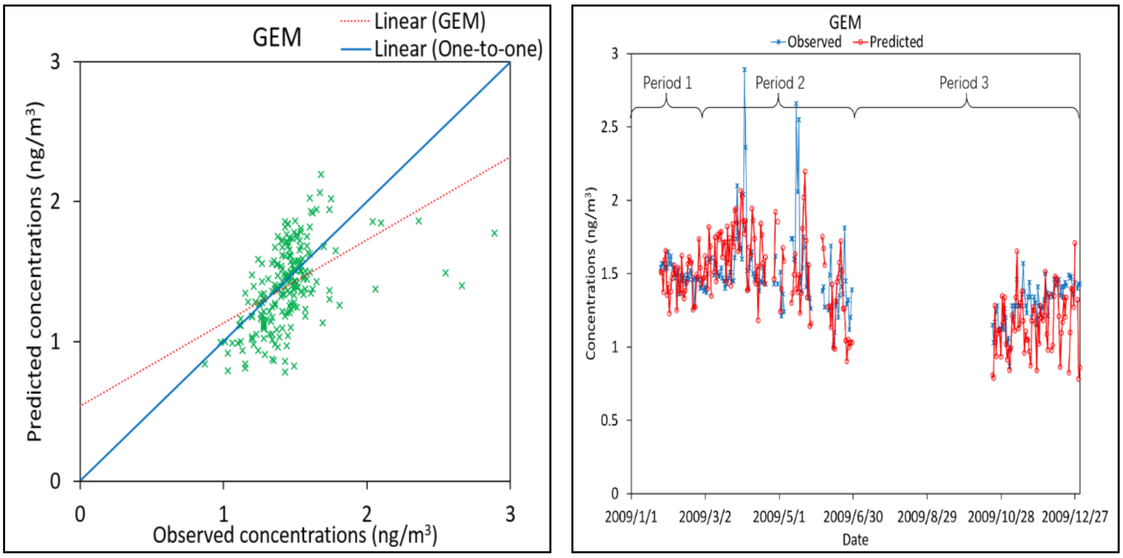

Figure D.4: Obs/Pred scatter plot and Obs/Pred time series in Case 4 in 2009

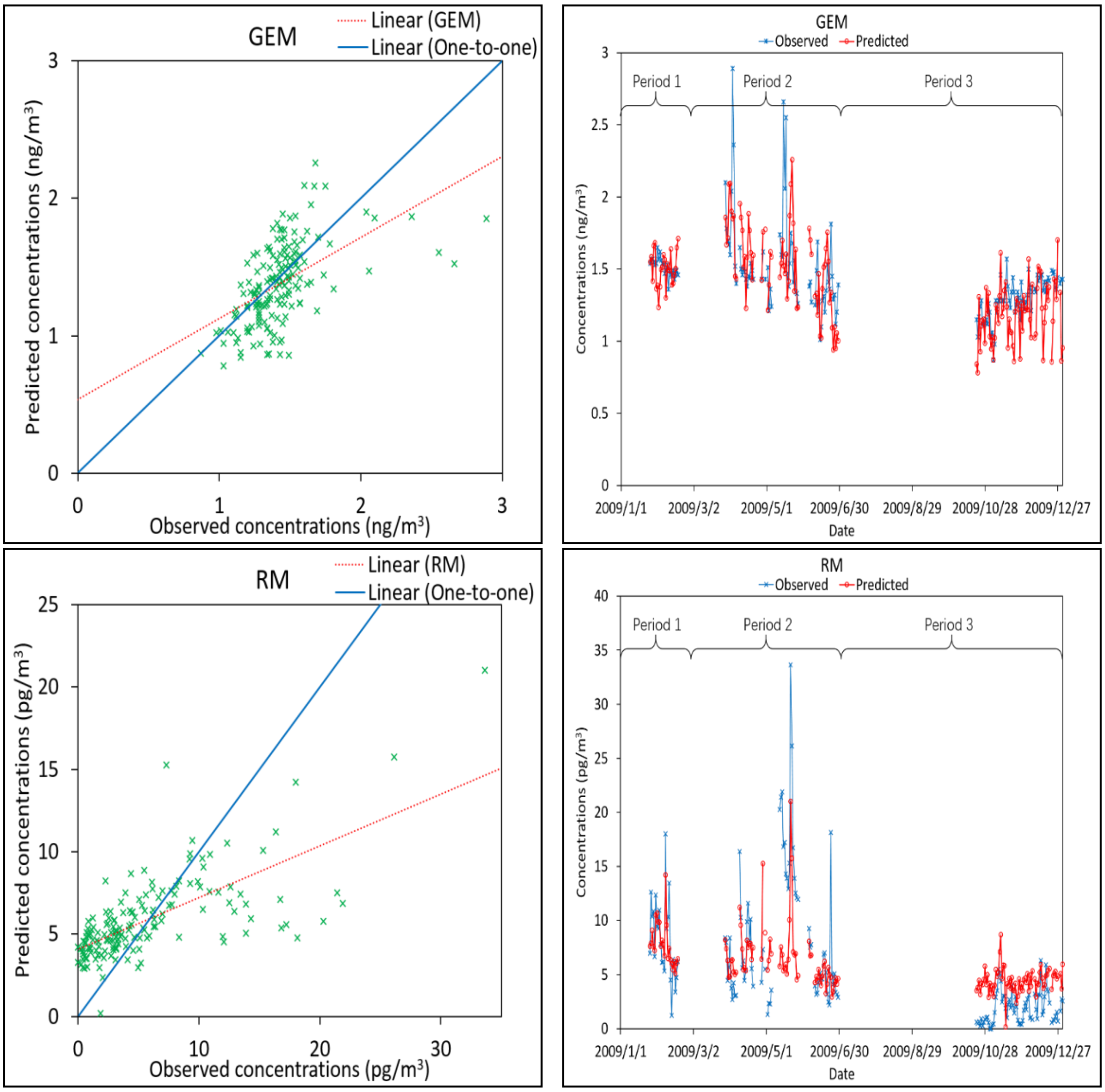

Figure D.5: Obs/Pred scatter plot and Obs/Pred time series in Case 5 in 2009

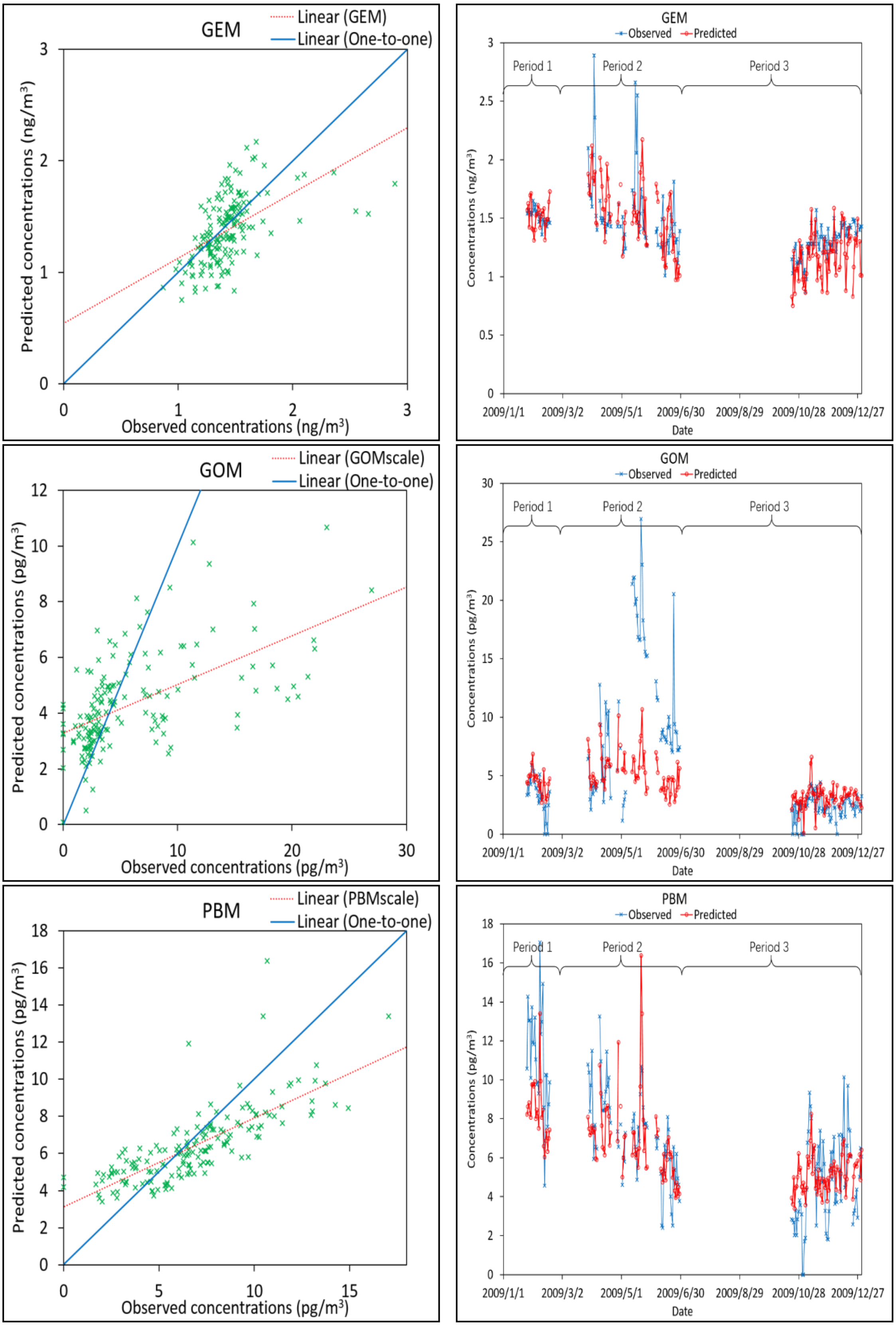

Figure D.6: Obs/Pred scatter plot and Obs/Pred time series in Case 6 in 2009

D.4 PMF Obs/Pred Scatter Plot and Obs/Pred Time Series in 2010

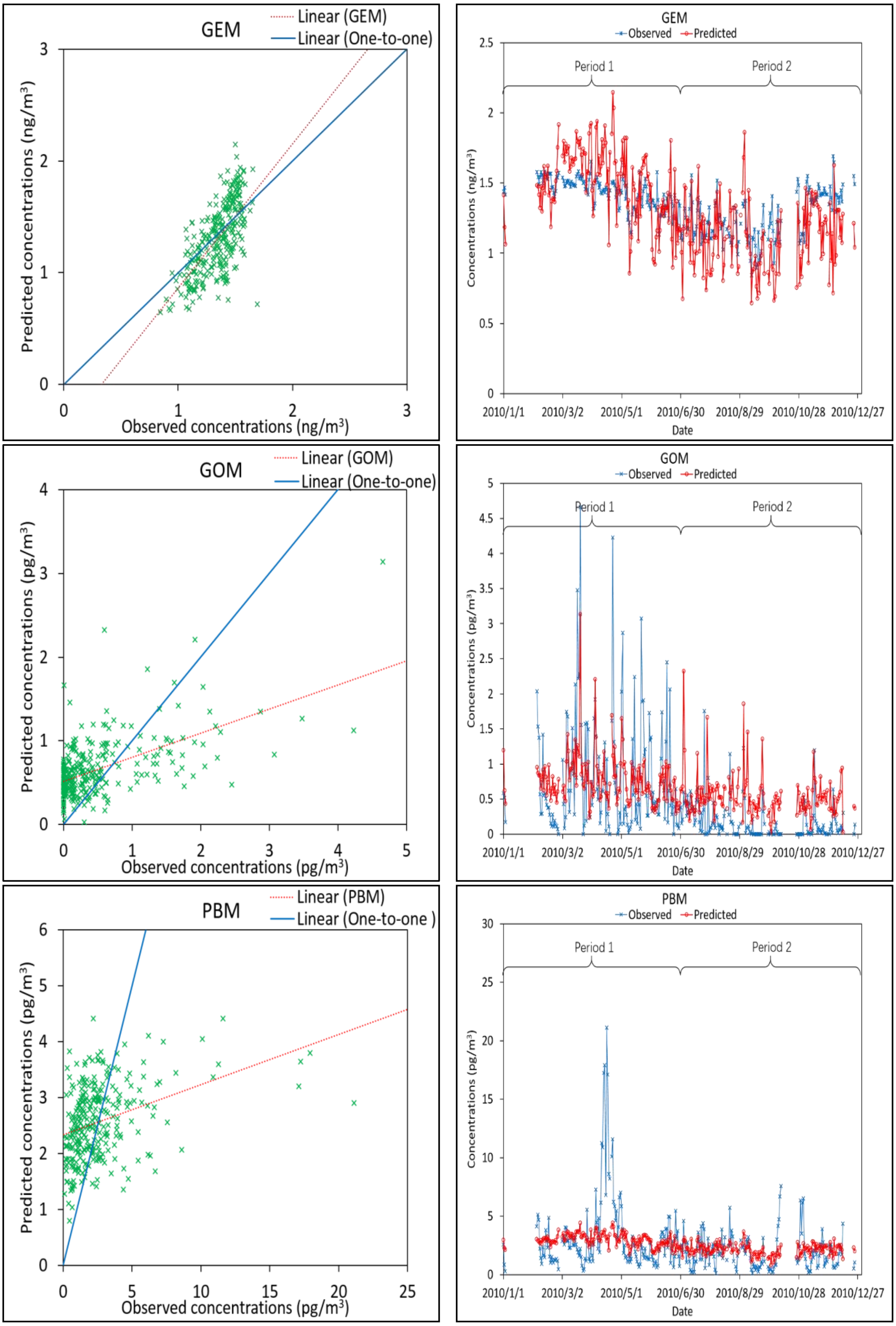

Figure D.7: Obs/Pred scatter plot and Obs/Pred time series in Case 7 in 2010

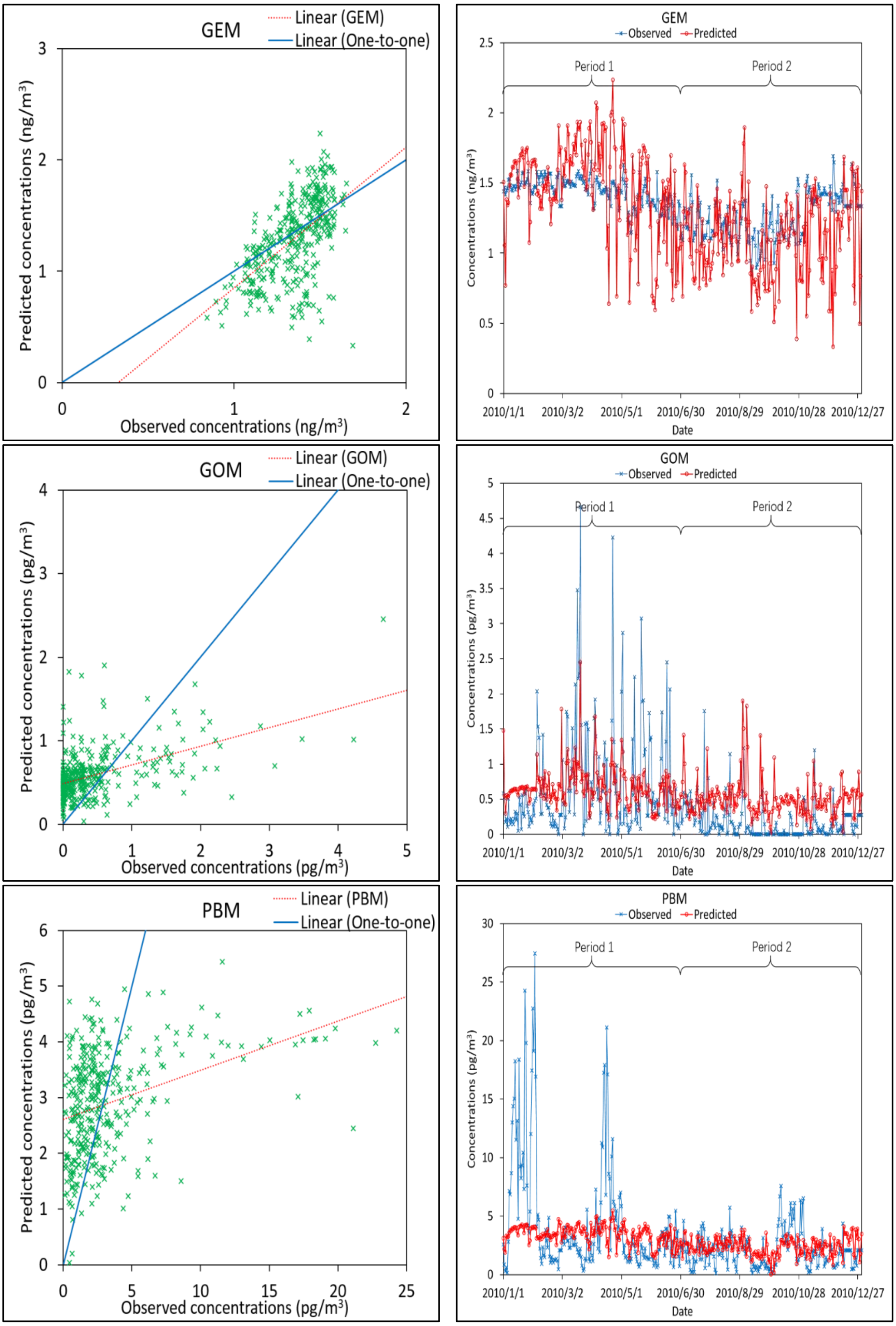

Figure D.8: Obs/Pred scatter plot and Obs/Pred time series in Case 8 in 2010

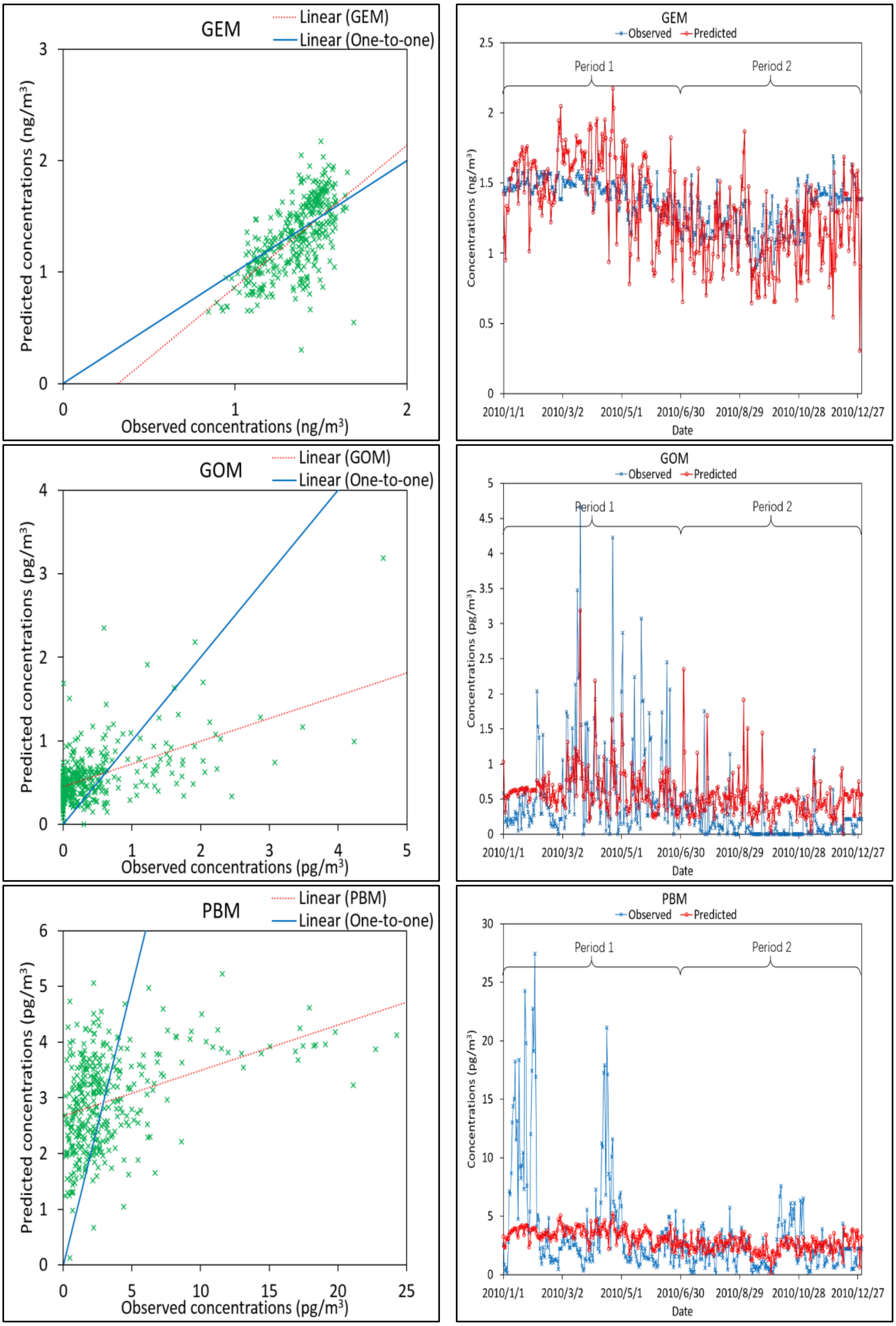

Figure D.9: Obs/Pred scatter plot and Obs/Pred time series in Case 9 in 2010

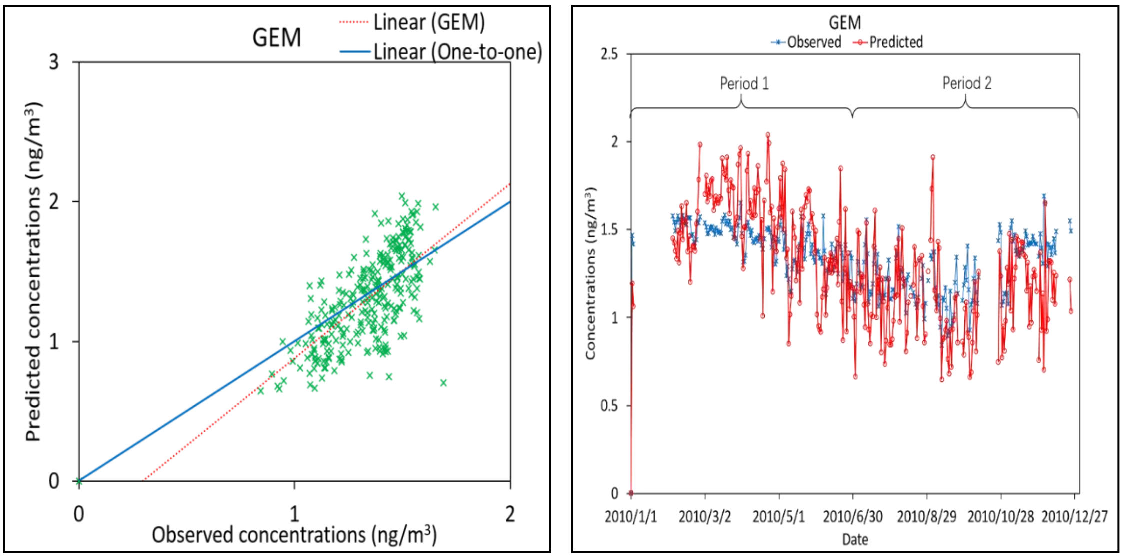

Figure D.10: Obs/Pred scatter plot and Obs/Pred time series in Case 10 in 2010

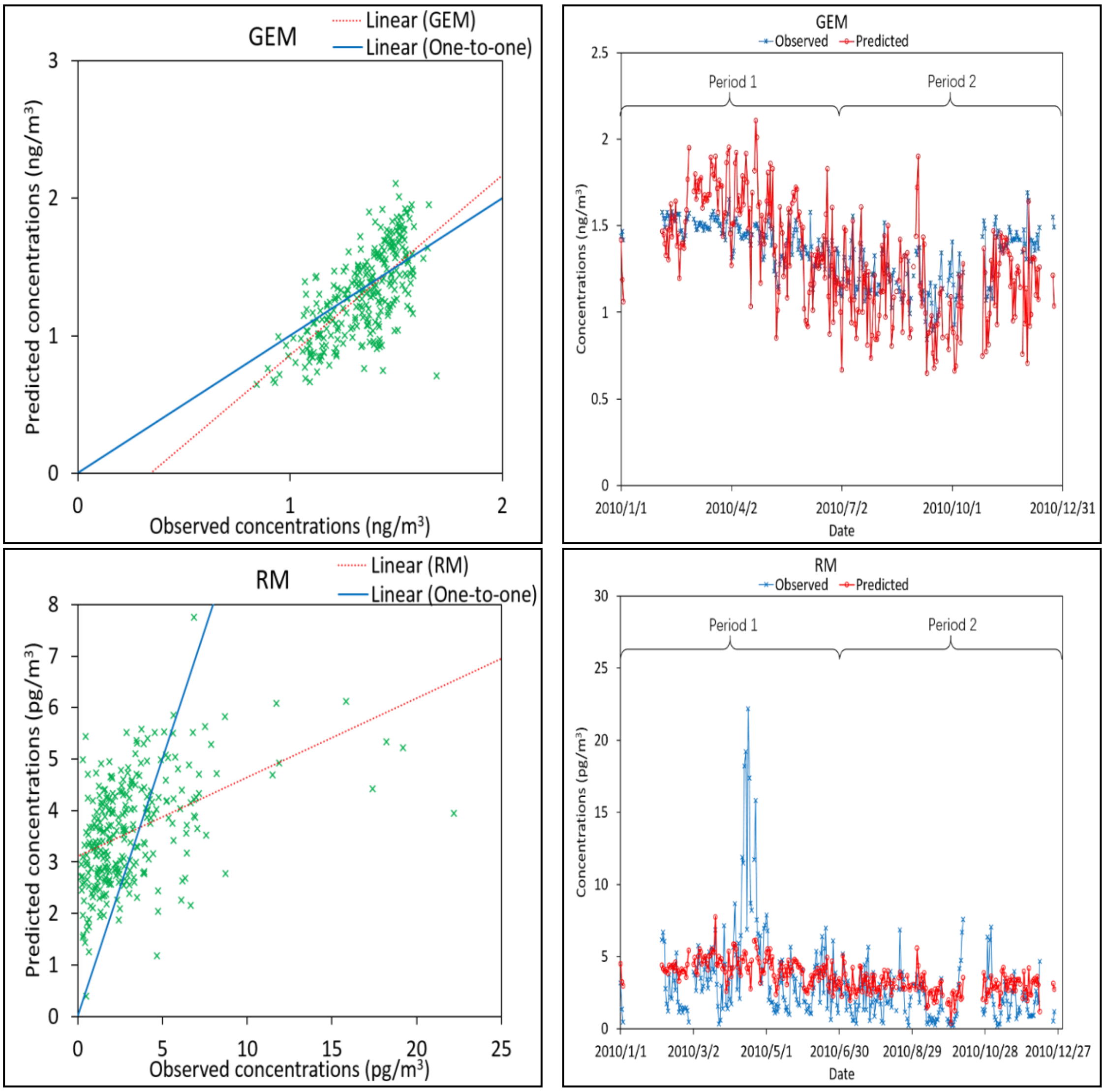

Figure D.11: Obs/Pred scatter plot and Obs/Pred time series in Case 11 in 2010

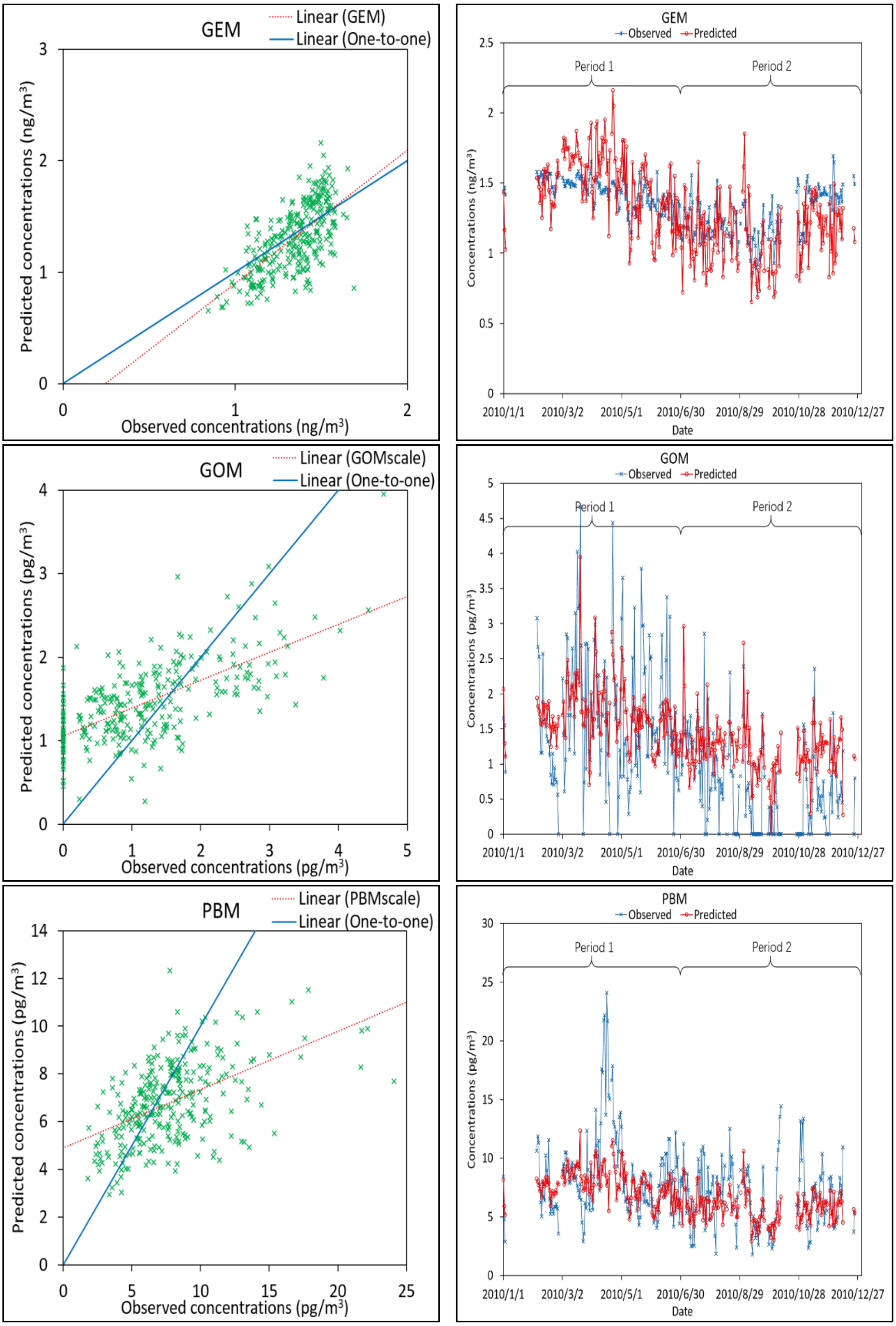

Figure D.12: Obs/Pred scatter plot and Obs/Pred time series in Case 12 in 2010

Appendix E:
PCA outputs

E.1 PCA outputs in 2009

Table E.1 Rotated component matrix of Case 13 in 2009

Factor Number	PC1	PC2	PC3	PC4
GEM	0.01	0.02	0.86	0.27
GOM	0.21	-0.11	0.26	0.84
PBM	0.63	0.03	0.50	-0.33
PM	0.80	0.25	0.10	0.21
O_3	0.50	0.08	0.70	0.05
SO_2	0.88	0.10	0.11	-0.05
HNO_3	0.86	0.00	0.07	0.34
Ca^{2+}	0.59	0.39	0.09	0.45
K^+	0.29	0.70	-0.19	0.33
Na^+	0.18	0.97	0.08	-0.08
Mg^{2+}	0.25	0.95	0.09	-0.01
Cl^-	-0.04	0.97	0.05	-0.12
NO_3^-	0.73	0.48	0.06	0.02
NH_4^+	0.92	0.07	0.09	0.12
SO_4^{2-}	0.86	0.22	0.19	0.11

Table E.2 Component score coefficient matrix of Case 13 in 2009

Factor Number	PC1	PC2	PC3	PC4
GEM	-0.18	0.03	0.61	0.15
GOM	-0.09	-0.03	0.09	0.64
PBM	0.14	-0.05	0.28	-0.40
PM	0.15	0.00	-0.07	0.06
O_3	0.01	-0.01	0.43	-0.08
SO_2	0.22	-0.06	-0.06	-0.17
HNO_3	0.18	-0.08	-0.11	0.16
Ca^{2+}	0.05	0.07	-0.06	0.29
K^+	-0.01	0.19	-0.19	0.26
Na^+	-0.06	0.28	0.06	-0.07
Mg^{2+}	-0.05	0.27	0.05	-0.02
Cl^-	-0.11	0.30	0.08	-0.07
NO_3^-	0.14	0.07	-0.07	-0.07
NH_4^+	0.22	-0.07	-0.10	-0.03
SO_4^{2-}	0.17	-0.02	-0.01	-0.03

\

Table E.3 Rotated component matrix of Case 14 in 2009

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM	0.13	-0.04	0.05	0.80	0.16
GOM	0.25	-0.13	0.64	0.40	-0.29
PBM	0.59	0.03	-0.47	0.34	-0.24
PM	0.81	0.23	0.17	0.10	-0.17
O_3	0.47	0.08	-0.13	0.72	-0.27
SO_2	0.86	0.09	-0.20	0.06	-0.14
HNO_3	0.87	-0.03	0.21	0.11	-0.15
Ca^{2+}	0.60	0.38	0.33	0.19	-0.20
K^+	0.36	0.66	0.39	-0.12	0.11
Na^+	0.21	0.96	-0.09	0.05	0.00
Mg^{2+}	0.28	0.95	-0.03	0.07	-0.04
Cl^-	-0.03	0.98	-0.10	0.03	0.00
NO_3^-	0.76	0.45	-0.04	-0.01	-0.01
NH_4^+	0.94	0.04	0.08	0.07	-0.05
SO_4^{2-}	0.88	0.19	0.05	0.17	-0.05
Temperature	0.02	-0.02	0.94	-0.07	0.07
Relative humidity	-0.26	0.04	-0.15	-0.20	0.79
Wind speed	-0.16	0.32	0.10	0.52	0.49
Precipitation	-0.07	-0.09	0.09	0.21	0.79

Table E.4 Component score coefficient matrix of Case 14 in 2009

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM	-0.05	-0.03	0.01	0.46	0.09
GOM	-0.05	-0.04	0.32	0.22	-0.17
PBM	0.09	-0.04	-0.27	0.14	-0.06
PM	0.15	0.00	0.06	-0.05	0.01
O_3	-0.03	0.00	-0.09	0.39	-0.14
SO_2	0.20	-0.06	-0.13	-0.08	0.06
HNO_3	0.19	-0.09	0.07	-0.06	0.04
Ca^{2+}	0.05	0.07	0.15	0.04	-0.08
K^+	0.05	0.16	0.19	-0.13	0.08
Na^+	-0.05	0.27	-0.05	0.01	-0.04
Mg^{2+}	-0.04	0.27	-0.02	0.02	-0.05
Cl^-	-0.11	0.30	-0.04	0.03	-0.07
NO_3^-	0.16	0.06	-0.05	-0.11	0.09
NH_4^+	0.24	-0.08	0.00	-0.10	0.12
SO_4^{2-}	0.19	-0.03	-0.01	-0.02	0.09
Temperature	-0.01	0.00	0.48	-0.06	0.03
Relative humidity	0.10	-0.02	-0.08	-0.13	0.48
Wind speed	-0.07	0.09	0.05	0.32	0.23
Precipitation	0.09	-0.08	0.03	0.09	0.49

E.1 PCA outputs in 2010

Table E.5 Rotated component matrix of Case 15 in 2010

Factor Number	PC1	PC2	PC3	PC4
GEM	-0.06	0.12	0.79	-0.20
GOM	0.05	-0.04	0.71	0.33
PBM	0.00	-0.10	0.48	0.04
O_3	0.00	0.05	0.91	0.06
SO_2	0.11	0.06	0.13	0.89
HNO_3	0.34	-0.14	0.02	0.82
Ca^{2+}	0.89	0.05	0.04	-0.05
K^+	0.77	0.10	-0.09	0.12
Na^+	0.00	0.99	-0.01	0.03
Mg^{2+}	0.34	0.93	0.01	0.00
Cl^-	-0.05	0.98	-0.04	-0.10
NO_3^-	0.79	0.21	0.09	0.12
NH_4^+	0.94	-0.10	-0.03	0.18
SO_4^{2-}	0.90	-0.01	-0.03	0.26

Table E.6 Component score coefficient matrix of Case 15 in 2010

Factor Number	PC1	PC2	PC3	PC4
GEM	0.03	0.02	0.38	-0.18
GOM	-0.02	-0.01	0.30	0.15
PBM	0.01	-0.04	0.22	-0.02
O_3	0.01	0.01	0.41	-0.03
SO_2	-0.12	0.07	-0.01	0.58
HNO_3	-0.03	-0.02	-0.05	0.49
Ca^{2+}	0.27	-0.03	0.04	-0.19
K^+	0.20	0.00	-0.03	-0.04
Na^+	-0.05	0.35	-0.02	0.08
Mg^{2+}	0.05	0.31	0.00	0.00
Cl^-	-0.05	0.34	-0.02	0.01
NO_3^-	0.21	0.04	0.05	-0.05
NH_4^+	0.25	-0.07	-0.01	-0.04
SO_4^{2-}	0.22	-0.03	-0.02	0.03

Table E.7 Rotated component matrix of Case 16 in 2010

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM	-0.03	0.03	0.87	0.10	-0.05
GOM	0.05	-0.06	0.50	-0.51	0.38
PBM	0.01	-0.08	0.29	-0.62	-0.11
O_3	0.02	-0.01	0.87	-0.19	0.20
SO_2	0.11	0.07	0.08	-0.11	0.84
HNO_3	0.33	-0.14	-0.10	-0.07	0.82
Ca^{2+}	0.89	0.04	0.03	0.00	-0.04
K^+	0.77	0.10	-0.13	-0.04	0.13
Na^+	0.00	0.99	0.06	0.04	0.02
Mg^{2+}	0.34	0.92	0.06	0.04	0.00
Cl^-	-0.05	0.97	0.05	0.08	-0.09
NO_3^-	0.80	0.20	0.06	-0.14	0.09
NH_4^+	0.94	-0.09	-0.07	-0.02	0.17
SO_4^{2-}	0.89	0.00	-0.08	-0.01	0.26
Temperature	0.26	-0.19	-0.52	0.08	0.27
Relative humidity	-0.12	0.07	-0.21	0.74	-0.33
Wind speed	-0.06	0.26	0.52	0.57	-0.06
Precipitation	-0.03	-0.08	0.24	0.76	-0.03

Table E.8 Component score coefficient matrix of case 16 in 2010

Factor Number	PC1	PC2	PC3	PC4	PC5
GEM	0.04	-0.05	0.36	0.08	-0.05
GOM	-0.03	-0.01	0.17	-0.17	0.15
PBM	0.03	-0.02	0.10	-0.32	-0.19
O_3	0.02	-0.04	0.34	-0.03	0.06
SO_2	-0.11	0.07	-0.01	0.07	0.54
HNO_3	-0.04	-0.01	-0.06	0.10	0.49
Ca^{2+}	0.27	-0.04	0.06	0.00	-0.18
K^+	0.20	0.01	-0.03	0.00	-0.04
Na^+	-0.05	0.35	-0.04	-0.03	0.07
Mg^{2+}	0.06	0.31	-0.01	-0.02	0.00
Cl^-	-0.05	0.34	-0.03	-0.02	0.01
NO_3^-	0.22	0.04	0.04	-0.06	-0.09
NH_4^+	0.25	-0.07	0.02	0.03	-0.05
SO_4^{2-}	0.22	-0.03	0.00	0.04	0.03
Temperature	0.02	-0.04	-0.20	0.07	0.16
Relative humidity	0.02	-0.01	-0.05	0.31	-0.08
Wind speed	0.01	0.03	0.23	0.30	0.05
Precipitation	0.03	-0.09	0.14	0.40	0.09

VITA AUCTORIS

NAME: Yanyin Liao

PLACE OF BIRTH: Shanghai, China

YEAR OF BIRTH: 1991

EDUCATION: Jinshan High School, Shanghai, China
2006-2009

Tongji University, Shanghai China
2009-2014 B.Sc (Environmental Engineering)

University of Windsor, Ontario, Canada
2014-2016 M.A.Sc. (Environmental Engineering)