

Interactive comment on “Potential sources and processes affecting speciated atmospheric mercury at Kejimikujik National Park, Canada” by Xiaohong Xu et al.

Response to Reviewers #2 and #3 comments

We appreciate both reviewers' constructive comments which helped us to improve the manuscript. Our point-by-point responses are provided below (in blue). The manuscript has been revised to reflect comments and suggestions by all three reviewers. Track change, yellow highlight (changes in tables and figures), or blue (added Figure 1 & SI sections) were used in the marked-up manuscript uploaded online.

Anonymous Referee #2

Received and published: 8 November 2016

This study used two-year Hg measurements (Tekran) with other air pollutants at Kejimikujik national park in Canada, and applied factor analysis (PMF) and principle component analysis to understand Hg sources and its related atmospheric processes. Overall, this is a well written article and easy to be followed paper. A very similar paper was published couple years ago; however, I understand the authors applied PMF as an additional analysis, and investigated how model setting impacts receptor modeling.

There are couple things I would like suggest to the authors to look into detail:

1) Wang et al., 2013 Chemosphere and Huang et al., 2010 ES&T have compared results from PCA and PMF using Hg related concentrations at Rochester, NY using similar data set. PCA and PMF comparisons using aerosol data have been discussed in detail in previous studies (Paatero and Tapper, 1994; Environmetrics, 1994).

These three papers have been included in the revised manuscript, “Comparisons of results of receptor models for PM source apportionment have been reported, e.g. Paatero and Tapper (1994), Viana et al. (2008), Belis et al. (2013), and Gibson et al. (2015). To date, PCA and PMF have been applied to atmospheric Hg and other air pollutants in Toronto (Canada) (Cheng et al., 2009) and in Rochester, New York (USA) (Huang et al., 2010; Wang et al., 2013). However, both the Toronto and Rochester studies lacked a thorough comparison of the PMF and PCA results.”

2) This study and Cheng et al., 2013 are using similar data set with similar results. What is new that we can learn from this study? After reading the abstract, I think the one new thing to the global Hg research group is the difference between 2009 and 2010. I suggest the authors should focus on these important things instead of repeating what we already knew or has been published on journals. I suggest a minor revision before ACP can accept this article. The specific comments are listed below: In abstract, the authors focus on comparison of result from different models; however, the title looks more like a straight source paper, suggest to modify either the title or abstract.

Good suggestion about the title, it has been revised as “Potential sources and processes affecting speciated atmospheric mercury at Kejimikujik National Park, Canada: comparison of receptor models and data treatment methods.”

The same dataset was used in a PCA study (Cheng et al., 2013), while this paper focuses on PMF and comparison between PMF and PCA results. Furthermore, the variables used, treatment of missing data and number of components are different. 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 SO₂, HNO₃, 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. Those are presented in Table 4. All these differences could result in the differences in the PCA results. In short, the comparison of the results suggests that the PCA results are sensitive to the input parameters. In terms of the differences in the results, four components were extracted in Cheng et al. (2013)’s study in 2009. Three out of four components, including Combustion/Industrial Source, Gas-particle partitioning of Hg, and Gas-phase Oxidation of Hg, were similar as the components in Case 09-C&M. 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 was 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 10-C&M. In a PM₁₀ source apportionment study using PMF and PCA at three European sites (Cesari et al., 2016), the authors reported that PCA results are more sensitive to the air contaminants present as input variables. In the revised manuscript, a reference has been added (Liao, 2016) for a detailed comparison of PCA results in this study and that in Cheng et al. (2013).

After reading this paper, one selling point is both model can capture the significant reduction of Hg and SO₂ from 2009 to 2010. However, it is not mentioned in the abstract.

Agree, the following sentence was added in section 3.1, “Moreover, the long term effects of emission reductions on Hg concentrations and source contributions should be investigated.”

Line 46-64, the most important difference between PMF (as a factor analysis) and PCA is the different concepts of these two receptor models, PMF constrains factor loadings and factor scores to nonnegative values and thereby minimizes the ambiguity caused by rotating factors. I suggest the authors dig this into detail and include the information there.

Agree, in the revised manuscript, we have referenced more papers for detailed comparison of the PMF and PCA approaches in the Introduction section. “Various receptor models have been used

to identify the sources and processes affecting air pollutant levels. Strengths and weaknesses of some receptor models have been reported previously (e.g. Watson et al., 2008, Viana et al., 2008; Belis et al., 2013).”

Line 60-62, Wang et al., 2013 Chemosphere and Huang et al., 2010 ES&T have done the comparison between PMF and PCA using Hg data.

In the revised manuscript, we have included both studies “Comparisons of results of receptor models for PM source apportionment have been reported, e.g. Paatero and Tapper (1994), Viana et al. (2008), Belis et al. (2013), and Gibson et al. (2015). To date, PCA and PMF have been applied to atmospheric Hg and other air pollutants in Toronto (Canada) (Cheng et al., 2009) and in Rochester, New York (USA) (Huang et al., 2010; Wang et al., 2013). However, both the Toronto and Rochester studies lacked a thorough comparison of the PMF and PCA results.”

Line 62-64, PMF has been applied to aerosol and evaluated in plenty previous studies, Belis et al., 2013 is a good article to start.

Agree, the article of Belis et al. (2013) has been included in the revised manuscript, see reply to comment 1).

Line 120-123, many people using GOM and PBM to do advanced statistical analysis, the biggest problem is how to handle missing and BMDL data. I look into table 1, a large portion of GOM/PBM is missing or BMDL. I understand that is the limitation of using statistical modeling on Hg data, but it will skew data distribution significantly.

We agree with the reviewer that when a large portion of GOM/PBM is missing or BMDL, it will skew the data distribution significantly. Data treatment techniques such as those used in this study may not improve the data distribution but may lead to improved model performance.

Line 142-143, after reading the entire paper, I still don't fully understand these cases.

Please see Tables 3 and 4 for a summary of the various cases which describes the different input variables and treatment of missing data.

Line 165, the authors used manufacture method detection limit. However, this can vary with locations and time, can the authors also talk about real MDL for Tekran system at this site?

We agree with the reviewer that study period and location specific MDLs would be useful. However, such MDLs for GOM and PBM cannot be accurately determined at this site and the majority of other monitoring sites using the Tekran system. This is primarily due to the lack of calibration standards for GOM and PBM as pointed out by Jaffe et al. (2014) and Gustin et al. (2015) and a lack of technologies to determine individual species of GOM (Jones et al., 2016). Therefore we used manufacture method detection limits for GOM and PBM for the purpose of identifying concentrations values with large uncertainties.

Line 166, why MDL for RM is 4 ng m⁻³? MDL is defined as 3 standard deviation of blanks, that could be the upper bound of MDL for RM, but if you look into distribution sum, that might be lower.

Added in the revised manuscript, “For RM, the MDL was assumed to be 4 pg/m³, which is a summation of MDLs of GOM and PBM (2 pg/m³each).”

Line 181, how did the authors select number of factor in PMF? In general, we look into Q and the variation of Q and number of factor.

We agree with the reviewer that more analysis of the PMF model outputs would be useful. In fact, the analyses of Q, IM and IS versus the number of factors were conducted. Similarly, the percent concentrations reconstructed by all factors were monitored for each of the three Hg forms. However, these analyses were not included in the submitted manuscript. A brief description of how the optimal number of factors was determined is now included in the Methods section. Detailed analysis is presented as Supplemental Information (SI), which support the stability of PMF runs, and justify the final solution and the number of factors chosen.

2) Line 238, Is this possible only due to biomass burning + soil emissions? We will see high ozone with biomass combustion, and it matches to all these increase for GEM, GOM, PBM, ozone. Does this happen in summer or winter, if you look into detail time series factor profiles, the authors should be able to figure this out.

An examination of the time series factor profiles revealed that model-reproduce K⁺, O₃ and GEM, GOM, PBM concentrations (in this factor) were rather smooth without any episodes of high O₃, K⁺, and Hg forms identified. The relatively stable patterns of K⁺ and GEM suggest re-emission of GEM while GOM was high in spring with elevated K⁺, O₃, indicating enhanced photochemistry. In the revised manuscript, we have added, “An examination of the time series factor profiles revealed that model-reproduced K⁺, O₃ and GEM, GOM, PBM concentrations (in this factor) were rather smooth. The impact of biomass burning seems to be small in this factor due to a lack of high K⁺, O₃, and Hg concentration periods or episodes identified. The relatively stable patterns of K⁺ and GEM suggest re-emission of GEM while GOM was high in spring with elevated O₃, indicating enhanced photochemistry.”

3) Line 302, I don't suggest using these analyses to predict GOM and PBM concentrations, as discussed above, a large portion of GOM and PBM is missing and BMDL.

Agree, we have replaced the word “predicted” with “model-reproduced” or “reproduced” throughout the revised manuscript.

4) Table 7, in the column title they are Case 10, but I think they should be 9.

The reviewer is correct; it should be “Case 9” in the column heading.

Anonymous Referee #3

Received and published: 10 November 2016

The receptor models for source apportionment of atmospheric mercury are of great importance. This study applied PMF and PCA on the data of speciated mercury and other tracers from a coastal observation site. Different methods of data processing were conducted for comparison. The comparison between PMF and PCA as well as between the two monitoring years was also performed. Advantages and disadvantages of the two receptor models were discussed. Overall, it is an important exploration of receptor models applying to atmospheric mercury studies. Elaborations on some key points are still needed. Therefore, I suggest the manuscript be accepted for the publication on Atmospheric Chemistry and Physics after major revision. Here are some specific comments:

1. Lines 46–64: This paragraph could use more literatures. Although the authors have reviewed the receptor model studies on atmospheric mercury in their previous paper (Cheng et al., 2015), examples on the applications of PMF and PCA are still needed in the introduction of this paper, not limited to atmospheric mercury. For example, Gibson et al. (2015) compared the four receptor models for PM_{2.5} source apportionment in Halifax. Some models could be more suitable for PM_{2.5} than for mercury. The authors could provide more proof on the merits and drawbacks of PMF and PCA when applied to atmospheric mercury.

Agree, we have included more papers in the revised manuscript, including the following: “Various receptor-based models have been used to identify the sources and processes affecting air pollutant levels. Strengths and weaknesses of some receptor models have been reported previously (e.g. Watson et al., 2008, Viana et al., 2008; Belis et al., 2013).” “Comparisons of results of receptor models for PM source apportionment have been reported, e.g. Paatero and Tapper (1994), Viana et al. (2008), Belis et al. (2013), and Gibson et al. (2015). To date, PCA and PMF have been applied to atmospheric Hg and other air pollutants in Toronto (Canada) (Cheng et al., 2009) and in Rochester, New York (USA) (Huang et al., 2010; Wang et al., 2013). However, both the Toronto and Rochester studies lacked a thorough comparison of the PMF and PCA results.”

2. Line 56: How do the authors define “qualitative” here? Aren’t the loadings of the PCA method quantitative? To my understanding, PMF describes the contributions of one parameter in different factors, while PCA describes the contributions of different parameters in one PC. The quantitative contribution of each PC to the receptor can be reflected by the “variance explained” (in Table 7 and 8).

Agree, the sentence has been rephrased as “PCA can only provide qualitative assessment of sources/processes but it cannot determine the source contributions to pollutant concentrations (Hopke, 2015).”. This is because the PCA loadings are not the same as the source contributions in PMF. In PCA, the loadings reflect the correlation coefficients between the variables and component, which are used to qualitatively assign the components to sources. Furthermore, the variance explained by each component in PCA is also not equivalent to source contributions to receptor measurements; it is only a measure of how well the component can explain the

variability in the dataset. The PMF factor contributions actually quantify the contribution of each source to atmospheric pollutant concentrations.

3. Section 2.1: A map of the observation site with locations of the emission sources listed in Table S1 and a brief description of the meteorological conditions would be useful. This information could be referred to in the discussion part to verify the results from the receptor models.

A map (Figure 1) has been included, also added statistics of meteorological parameters in Tables 1-2 with the following statements, “The weather conditions were similar in the two years, with an annual mean relative humidity of 88% and 87% in 2009 and 2010 respectively, moderate wind speeds (4.7 km/h and 4.4 km/h), but a higher precipitation amount (1597 mm/yr vs. 1480 mm/yr) and a lower temperature (6.6°C vs. 8.1°C) in 2009 than 2010.”

Table 1. General statistics of daily air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise noted) and meteorological parameters in 2009.

Compound	Percent of missing values	Method detection limit (MDL)	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation	Coefficient of variability (%)
GEM (ng/m^3)	31%	0.1	0%	1.37	1.41	1.39	0.26	18.7
GOM (pg/m^3)	32%	2	78%	0.57	0.42	1.77	3.70	209
PBM (pg/m^3)	41%	2	48%	1.78	2.15	2.81	2.72	96.8
PM	20%	1	9%	2.71	2.91	3.44	2.49	72.4
O ₃	0%	4.3	0%	59.4	62.1	62.4	19.1	30.6
SO ₂	3%	0.002	0%	0.20	0.22	0.40	0.51	128
HNO ₃	3%	0.05	12%	0.13	0.12	0.19	0.22	116
Ca ²⁺	1%	0.002	0%	0.05	0.05	0.06	0.04	66.7
K ⁺	1%	0.029	17%	0.04	0.03	0.04	0.03	75.0
Na ⁺	1%	0.05	9%	0.25	0.30	0.43	0.47	109
Mg ²⁺	1%	0.0004	2%	0.04	0.04	0.06	0.06	100
Cl ⁻	1%	0.046	23%	0.19	0.23	0.46	0.64	139
NO ₃ ⁻	1%	0.06	9%	0.18	0.17	0.28	0.39	139
NH ₄ ⁺	1%	0.001	0%	0.19	0.17	0.28	0.32	114
SO ₄ ²⁻	1%	0.05	0%	0.78	0.76	1.14	1.27	111
Total ions	1%	-	-	2.13	2.05	2.76	2.23	81
Temperature (°C)	0%	-	-	-	7.31	6.64	9.28	140
Relative humidity (%)	0%	-	-	-	87.5	84.5	12.0	14
Wind speed (m/s)	0%	-	-	-	4.33	4.70	2.39	51
Precipitation (mm/day)	3%	-	-	-	0.60	4.50	10.0	222

Table 2. General statistics of daily air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise noted) and meteorological parameters in 2010, MDL same as in Table 1.

Compound	Percent of missing values	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation	Coefficient of variability (%)
GEM (ng/m ³)	4%	0%	1.34	1.38	1.35	0.17	12.6
GOM (pg/m ³)	4%	96%	0.27	0.21	0.44	0.64	145
PBM (pg/m ³)	4%	46%	2.08	2.20	3.40	4.13	121
O ₃	1%	0%	62.2	63.4	64.5	16.6	25.7
SO ₂	19%	1%	0.10	0.13	0.23	0.31	135
HNO ₃	19%	25%	0.10	0.10	0.18	0.22	122
Ca ²⁺	19%	0%	0.04	0.04	0.07	0.13	186
K ⁺	19%	46%	0.04	0.03	0.06	0.07	117
Na ⁺	19%	16%	0.20	0.24	0.40	0.53	133
Mg ²⁺	19%	0 %	0.03	0.04	0.05	0.06	120
Cl ⁻	19%	27%	0.14	0.15	0.46	0.83	180
NO ₃ ⁻	19%	21%	0.14	0.13	0.25	0.36	144
NH ₄ ⁺	19%	0%	0.16	0.15	0.30	0.57	190
SO ₄ ²⁻	19%	0%	0.69	0.64	1.11	1.65	149
Total ions	19%	-	1.89	1.80	2.71	2.95	109
Temperature (°C)	0%	-	-	8.57	8.13	8.92	110
Relative humidity (%)	0%	-	-	86.8	84.5	12.6	15
Wind speed (m/s)	0%	-	-	3.63	4.37	3.09	71
Precipitation (mm/day)	2%	-	-	0.60	4.15	9.71	234

4. Lines 113–114: Is there any specific reason why the authors averaged the original data to daily values? If the original data is hourly or 3-hr, it should be possible to obtain 3-hr, 6-hr or 12-hr averages, which could result in a larger database for PMF and PCA. Isn't it better?

The Hg data were 3-hrs. However, the concentrations of SO₂ and HNO₃, PM_{2.5} (2009 only), and particulate SO₄²⁻, NO₃⁻, Mg²⁺, Cl⁻, K⁺, Ca²⁺, NH₄⁺, and Na⁺ were daily values. Thus, hourly or 3-hr concentrations of GEM, GOM, PBM, O₃ and meteorological data were averaged into daily values because PMF and PCA require the same interval for all input variables.

5. Line 145: The expression “resultant PMF result” seems repetitive.

Agree, the word “resultant” has been removed.

6. Lines 192–193: Since the PCA analysis has already been conducted in Cheng et al. (2013), I think the current title of the manuscript is inappropriate. It could give the readers misimpression that this is partially repeated from the previous study. To my understanding, the methodology of

this study is the novelty of this paper. Therefore, it is better to embody the methodology in the title.

Agree, the title has been revised as “Potential sources and processes affecting speciated atmospheric mercury at Kejimikujik National Park, Canada: comparison of receptor models and data treatment methods”.

7. Line 203: Have the authors checked the inter-correlations between any two of the major PCs? Varimax is an orthogonal rotation method, which requires the PCs to be independent on each other. This validation process for the applicability of the Varimax rotation could be mentioned here.

We have checked the correlations between PCs using an oblique rotation method (direct oblimin) instead of Varimax rotation. The correlations between factors in Case 09-C, Case 10-C and Case 10-C&M are all below the Tabachnick and Fidell threshold of 0.32 (Tabachnick and Fidell, 2007) indicating the solution remains nearly orthogonal. The correlation between factor combustion/industrial emission and factor Hg wet deposition had a correlation of -0.33. However, the factor loadings between direct oblimin and varimax rotation were not very different. Therefore, we kept the original varimax results and added an explanation of using an oblique rotation method to verify the inter-correlations between the components in the revised paper.

8. Table 5 and Figure 1-2: NO₃ in Table 5 should be NO₃⁻. All the “+” and “-” signs cannot be omitted in Figure 1-2. NO₃ and NO₃⁻ stand for different compounds.

Table 5 and Figures 1-2 have been modified as suggested.

9. Line 212: From the context (Lines 267–268), Combustion Emission include both coal combustion and biomass burning? It is better to mention it here. Does open biomass burning or wildfires included in F1?

We agree with the reviewer that it should be clarified early on that combustion emissions include both coal combustion and biomass burning. However there are not enough pollutant markers to distinguish between the various types of combustion sources, such as open biomass burning from wildfires. In the revised manuscript, we have added “Combustion Emission includes fuel combustion and biomass burning. The small contributions of Ca⁺ (19%) and K⁺ (22%) suggest a minor impact of biomass burning.”

10. Line 232: Can the authors specify what types of sources could be Industrial Sulfur? Non-ferrous metal melting? What could be the possible Industrial Sulfur sources in this region?

Added as suggested, “As shown in Table S1, point sources of industrial sulfur in the province of Nova Scotia include tire production, engineered wood production, food industry, and universities. Coal-fired power plants and metal production are major sources of sulfur; however there are no combustion sources close to the sampling site. These sources are located in eastern

U.S., which could be transported to the site. Mobile sources of sulfur are ships and vessels from nearby ports (Cheng et al., 2013).”

11. Line 238: The authors mentioned biomass combustion in this part while the name of Factor 3 is Photochemical Process and Re-emission of Hg. Why is it necessary for the biomass combustion to be related to Re-emission? Is it possible that F1 is composed of coal combustion and controlled biomass combustion which are usually mixed from regional sources while F3 is composed of mineral dust and open biomass burning/wildfires which are usually mixed in long-range transport?

We agree with the reviewer the need to clarify. For F1 (Combustion Emission), the following sentences have been added: “Combustion Emission includes fuel combustion and biomass burning. The small contributions of Ca^+ (19%) and K^+ (22%) suggest a minor impact of biomass burning.” For F3, we have added the following, “An examination of the time series factor profiles revealed that model-reproduced K^+ , O_3 and GEM, GOM, PBM concentrations (in this factor) were rather smooth. The impact of biomass burning seems to be small in this factor due to a lack of high K^+ , O_3 , and Hg concentration period or episodes identified. The relatively stable patterns of K^+ and GEM suggest re-emission of GEM while GOM was high in spring with elevated O_3 , indicating enhanced photochemistry.”

12. Table 6: The performance of 2009 GOM and 2010 PBM is poor to me. I don’t think the previous discussion linked to these two parts can be validated. Is it possible to improve the model performance by using the 3-hr or 6-hr averages instead of daily averages to increase the size of the database?

In this study, the model performance was evaluated using a number of indexes, include scaled residual plot to evaluate distribution of residuals, Obs/Pred scatter plot to evaluate overall model-measurement agreement, Obs/Pred time series to visualize the model’s ability to reproduce monitored concentrations, the Pred/Obs ratios and the annual Predmean/Obsmean ratios to quantify agreement between predicted and observed Hg concentrations on day-to-day and annual basis, respectively. Those analyses indeed suggested that the model performance of 2009 GOM and 2010 PBM was poor.

We agree with the reviewer that it could be possible to improve the model performance by using the 3-hr or 6-hr averages instead of daily averages to increase the size of the database. Unfortunately, the concentrations of SO_2 and HNO_3 , $\text{PM}_{2.5}$ (2009 only), and particulate SO_4^{2-} , NO_3^- , Mg^{2+} , Cl^- , K^+ , Ca^{2+} , NH_4^+ , and Na^+ were daily values. Thus, hourly O_3 and meteorological data, as well as 3-hr concentrations of GEM, GOM, PBM were averaged into daily values because PMF and PCA require the same interval for all input variables.

References (papers used in this response but not listed in the reference section of the revised manuscript)

Jaffe, D. A., Lyman, S., Amos, H. M., Gustin, M. S., Huang, J., Selin, N. E., Levin, L., Schure, A. t., Mason, R. P., Talbot, R., Rutter, A., Finley, B., Jaeglé, L., Shah, V., McClure, C., Ambrose, J., Gratz, L., Lindberg, S., Weiss-Penzias, P., Sheu, G.-R., Feddersen, D., Horvat, M.,

Dastoor, A., Hynes, A. J., Mao, H., Sonke, J. E., Slemr, F., Fisher, J. A., Ebinghaus, R., Zhang, Y., and Edwards, G.: Progress on Understanding Atmospheric Mercury Hampered by Uncertain Measurements, *Environ. Sci. Technol.*, 48, 7204-7206, 10.1021/es5026432, 2014.

Jones, C. P., Lyman, S. N., Jaffe, D. A., Allen, T., and O'Neil, T. L.: Detection and quantification of gas-phase oxidized mercury compounds by GC/MS, *Atmospheric Measurement Techniques*, 9, 2195-2205, 10.5194/amt-9-2195-2016, 2016.

Tabachnick, B. G., and Fidell, L. S.: *Using Multivariate Statistics*, 5th ed., Pearson Allyn & Bacon, Upper Saddle River, NJ, 2007.

Potential sources and processes affecting speciated atmospheric mercury at

**Kejimikujik National Park, Canada: comparison of receptor models and data
treatment methods**

Xiaohong Xu^{1*}, Yanying Liao¹, Irene Cheng², Leiming Zhang^{2*}

¹ Department of Civil and Environmental Engineering, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, N9B 3P4, Canada

² Air Quality Research Division, Science and Technology Branch, Environment and Climate Change Canada, 4905 Dufferin Street, Toronto, Ontario, M3H 5T4, Canada

Corresponds to Xiaohong Xu (xxu@uwindsor.ca) or Leiming Zhang (leiming.zhang@canada.ca)

1 **Abstract:** Source apportionment analysis was conducted with Positive Matrix
2 Factorization (PMF) and Principal Component Analysis (PCA) methods using
3 concentrations of speciated mercury (Hg), i.e., gaseous elemental mercury (GEM),
4 gaseous oxidized mercury (GOM), and particulate-bound mercury (PBM), and other
5 air pollutants collected at Kejimikujik National Park, Nova Scotia, Canada in 2009 and
6 2010. The results were largely consistent between the two years for both methods. The
7 same four source factors were identified in each year using PMF method. In both
8 years, factor Photochemistry and Re-emission had the largest contributions to
9 atmospheric Hg, while the contributions of Combustion Emission and Industrial
10 Sulfur varied slightly between the two years. Four components were extracted with air
11 pollutants only in each year using PCA method. Consistency between the results of
12 PMF and PCA include, 1) most or all PMF factors overlapped with PCA components,
13 2) both methods suggest strong impact of photochemistry, but little association
14 between ambient Hg and sea salt, 3) shifting of PMF source profiles and source
15 contributions from one year to another was echoed in PCA. Inclusion of
16 meteorological parameters led to identification of an additional component - Hg Wet
17 Deposition in PCA, while it did not affect the identification of other components.

18 The PMF model performance was comparable in 2009 and 2010. Among the
19 three Hg forms, the agreements between ~~model-reproduced~~ ~~predicted~~ and observed
20 annual mean concentrations were excellent for GEM, very good for PBM and
21 acceptable for GOM. However, on daily basis, the agreement was very good for GEM,
22 but poor for GOM and PBM. Sensitivity tests suggest that increasing sample size by
23 imputation is not effective in improving model performance, while reducing the
24 fraction of concentrations below method detection limit, by either scaling GOM and
25 PBM to higher concentrations or combining them to reactive mercury, is effective.
26 Most of the data treatment options considered had little impact on the source
27 identification/contribution.

28 1. Introduction

29 Atmospheric mercury (Hg) exists in the form of gaseous elemental Hg (GEM) and
30 oxidized Hg, the latter can be in gaseous phase (gaseous oxidized Hg - GOM) or
31 associated with particulate matter (particulate - bound Hg - PBM). Identification of
32 major sources and processes affecting ambient levels of different Hg forms will help

33 mitigate the risks of Hg pollution. Atmospheric Hg can be produced from
34 anthropogenic activities, natural events and re-emission of previously deposited Hg,
35 the latter two are sometimes grouped together as natural emission sources (Gustin et
36 al., 2008; Pirrone et al., 2010; UNEP, 2013; Gaffney and Marley, 2014; Zhang et al.,
37 2016). Natural events consist of volatilization from the ocean, volcanic eruption,
38 geothermal activities, and weathering of Hg-containing minerals (Pirrone et al., 2010;
39 Gaffney and Marley, 2014). Small scale or artisanal gold mining, mining and smelting,
40 and coal combustion are the three major anthropogenic sources (UNEP, 2013; Zhang
41 et al., 2016). Some of the dry and wet deposited PBM and GOM will be reduced to
42 GEM in soil, water, and vegetation surfaces where Hg will be re-emitted in the form
43 of GEM to the atmosphere (Gaffney and Marley, 2014). However, the contributions of
44 each source and process to a given receptor site are affected by many factors including
45 proximity to sources and weather conditions.

46 Various receptor-based models have been used to identify the sources and
47 processes affecting air pollutant levels. Strengths and weaknesses of some receptor
48 models have been reported previously (e.g. Viana et al., 2008; Watson et al., 2008;
49 Belis et al., 2013).ambient Hg levels (Cheng et al., 2015). Among these, Positive
50 Matrix Factorization (PMF) and Principal Component Analysis (PCA) are two
51 commonly used methods. PMF method provides quantitative source profiles and
52 source contributions. The resultant source profiles could aid future studies in factor
53 interpretation. Another strength of PMF is input variable screening and provision of
54 model performance measures. The users could specify uncertainty values for each
55 variable in each sample to reduce the impact of measurements with high uncertainties
56 on the final results (US EPA, 2014a; Hopke, 2016). However, in order to derive
57 profiles, PMF requires a large number of air pollutants which are often unavailable. In
58 contrast, PCA can only provide qualitative assessment of sources/processes; however
59 it cannot determine the source contributions to pollutant concentrations (Hopke, 2015).
60 One advantage of PCA over PMF is its capability of allowing inclusion of
61 meteorological parameters as input, enabling the assessment of the effects of weather
62 conditions on ambient ~~Hg~~ concentrations of e.g. Hg (Cheng et al., 2015). Therefore, it
63 is beneficial to conduct source apportionment of atmospheric Hg using both PMF and
64 PCA.

65 Comparisons of results of receptor models for PM source apportionment have
66 been reported, e.g. Paatero and Tapper (1994), Viana et al. (2008), Belis et al. (2013),
67 and Gibson et al. (2015). To date, PCA and PMF have been applied to atmospheric Hg
68 and other air pollutants in Toronto (Canada) (Cheng et al., 2009) and in Rochester,
69 New York (USA) (Huang et al., 2010; Wang et al., 2013). However, both the Toronto
70 and Rochester studies lacked a thorough comparison of the PMF and PCA results. To
71 date, only one study used this combined approach (Cheng et al., 2009), yet it lacked a
72 thorough comparison of the results. Furthermore, the ability of receptor models to
73 reproduce the observed concentrations should be assessed in order to gauge the model
74 performance (Henry, 1991; Viana et al., 2008; [Belis et al., 2015a](#)), which has been
75 rarely reported in the literature.

76 The overall objective of this study is to identify the factors affecting ambient Hg
77 concentrations at a receptor site using PMF and PCA approaches. The specific
78 objectives are to, (1) identify the factors affecting ambient Hg concentrations using
79 PCA and PMF model; (2) summarize the similarity and differences in PMF factors
80 and PCA components; (3) evaluate the PMF model performances by Hg forms; (4)
81 investigate the impact of including meteorological parameters on PCA results, and (5)
82 assess the sensitivity of PMF results and performance to different treatment of
83 missing data and low concentration values of speciated Hg.

85 **2. Method**

86 **2.1 Study site**

87 The study site is located in Kejimikujik (KEJ) National Park (44.32°N; 65.2°W;
88 elevation: 170 m), Nova Scotia, Canada. The KEJ site is one of the first speciated Hg
89 sites operated by Environment Canada outside the Arctic. This site was selected
90 primarily because of the bioaccumulation issues at this area. Studies have found that
91 common loons in Kejimikujik National Park had the highest mean blood Hg
92 concentrations in northeastern United States and Southeastern Canada (Evers et al.,
93 2007). Similarly, a 1996/97 survey found that yellow perch and common loons from
94 Kejimikujik National Park and National Historic Site (Nova Scotia) had the highest
95 blood Hg concentrations across North America. A 2006/07 follow up study on yellow
96 perch observed on average a 29% increase in 10 out of 16 lakes, although

97 anthropogenic emission from North America decreased between the mid-90s to the
98 mid-2000s (Wyn [et al.](#), 2010).

99 The sampling site was surrounded by forests on a flat terrain. It was
100 approximately 50 km away from the nearest coast, 120 km southwest of Halifax, and
101 relatively remote from anthropogenic air emissions. A search of the National Pollutant
102 Release Inventory (NPRI, Environment Canada, 2016) yielded seven Nova Scotia
103 facilities reporting Hg air releases in both 2009 and 2010 ([Figure 1](#)). Four of them
104 were electric power generation stations, the other three were a refinery, a cement plant,
105 and a university. The provincial annual air emissions of Hg were 147.5 kg and 90.3 kg
106 in 2009 and 2010, respectively (Table S1). The two largest Hg emitters were Lingan
107 Power Generating Station (2009-2010 average: 71 kg/yr) and Trenton Power
108 Generating Station (26 kg/yr), located 450 km and 250 km from the sampling site,
109 respectively. The nearest anthropogenic Hg sources (Dalhousie University, Halifax:
110 0.17 kg/yr, Imperial oil, Dartmouth Refinery: 2.8 kg/yr) were 140 km northeast of the
111 sampling site. In addition to Hg sources, the nearby NPRI (Environment Canada,
112 2016) combustion/industrial sources were a biomass-fueled power station and tire
113 production factory located approximately 50 km east/southeast of the KEJ site (Table
114 S1).

115

116 **2.2 Data collection**

117 GEM, GOM and PBM concentrations were collected from 2009 to 2010 using
118 Tekran[®] instruments (Models 1130/1135/2537) at 3-hour intervals. Hourly
119 concentrations of ground level ozone (O₃) and meteorological parameters
120 (temperature, relative humidity, wind speed, and precipitation amount), as well as
121 daily concentrations of SO₂ and HNO₃, PM_{2.5} (2009 only), and particulate SO₄²⁻, NO₃⁻,
122 Mg²⁺, Cl⁻, K⁺, Ca²⁺, NH₄⁺, and Na⁺ were also collected at KEJ site. Detailed
123 information of data collection can be found in Cheng et al. (2013).

124 Hourly or 3-hr concentrations of GEM, GOM, PBM, O₃ and meteorological data
125 were averaged into daily values because PMF and PCA require the same interval for
126 all input variables. All daily values were the same as those used in a PCA study by
127 Cheng et al. (2013). The general statistics of the daily concentrations [and](#)
128 [meteorological parameters](#) are listed in Table 1 and Table 2 for year 2009 and year

2010, respectively. The total aerosol mass characterized in 2009 accounted for 80% of the PM mass. The weather conditions were similar in the two years, with an annual mean relative humidity of 88% and 87% in 2009 and 2010 respectively, moderate wind speeds (4.7 km/h and 4.4 km/h), but a higher precipitation amount (1597 mm/yr vs. 1480 mm/yr) and a lower temperature (6.6°C vs. 8.1°C) in 2009 than 2010. The number of missing daily concentrations ranged from 0% (ozone, 2010) to 41% (PBM, 2009), which are excluded from PMF or PCA. Among the three Hg forms, GEM had the fewest values below the Method Detection Limit (MDL), while GOM had the largest percentages of concentrations below MDL, followed by PBM, in both years. The variability, as indicated by coefficient of variability, was low for GEM but much higher for GOM and PBM.

2.3 Model setup and case design

Detailed description of the theory of PMF and PCA methods can be found in Cheng et al. (2015). Model set up and case design are described below.

PMF

EPA PMF5.0 (US EPA, 2014b) was used in this study. The 12 cases investigated are listed in Table 3. Two approaches were employed in PMF modeling to handle missing values. The first approach is listwise deletion. Listwise deletion excludes all the records having one or more missing values, resulting in a complete data matrix as required in PMF. However, it may cause a large reduction of the dataset when one of the pollutants has many missing values or several pollutants have missing values at different time periods. In environmental studies, this approach may lead to biased results because listwise deletion benefits the records with high concentrations when below MDL values are flagged as missing (Huang et al., 1999). The second method is imputation, which increases the sample size in PMF. Hedberg et al. (2005) found that the relative error of factor profiles decreased as the sample size increased. In this study, geometric mean and median imputation were used to minimize the undue influence of extreme values as in Pekey et al. (2004). The effects of imputation ~~were~~ investigated in Cases 09+Mean, 10+Mean, 09+Median, and 10+Median (Table 3).

Cases 09+RM, 10+RM, 09-RM, and 10-RM (Table 3) were devised to investigate the effects of excluding or combining GOM and PBM into reactive mercury (RM) on

161 | the ~~resultant~~-PMF results compared with the full dataset. Uncertainties of GOM and
162 PBM measurements are considered high (Gustin et al., 2015). It has been reported that
163 GOM may be collected on the PBM filter thus GOM concentrations could be biased
164 low (Lynam and Keeler, 2005). Therefore, combining GOM and PBM to RM may
165 reduce the uncertainties (Cheng et al., 2016). RM was calculated by summing GOM
166 and PBM when both forms of Hg are detected.

167 In Case 09ScaleRM and Case 10ScaleRM, a variable scaling factor was used to
168 increase the GOM and PBM concentrations:

$$169 \quad \text{scaling factor} = \sqrt{\frac{\max(x)}{x_i}} \quad (1)$$

170

171 where x_i is the concentration of GOM or PBM in the i^{th} sample. The scaling factor is
172 large when the concentration is low, and vice versa, but the maximum concentration is
173 unchanged.

174 Equation-based uncertainties (US EPA, 2014a) were used in this study, expressed
175 as:

176

$$177 \quad \text{Uncertainty} = \frac{5}{6} \times MDL, \text{ when concentration} \leq MDL$$
$$178 \quad \text{Uncertainty} = \sqrt{(\text{Error Fraction} \times \text{concentration})^2 + (0.5 \times MDL)^2}, \quad (2)$$
$$179 \quad \text{when concentration} > MDL$$

177

178 The MDLs used in this study are 0.1 ng/m³, 2 pg/m³, and 2 pg/m³ for GEM, GOM
179 and PBM, respectively (Tekran Inc., 2010). For RM, the MDL was assumed to be 4
180 | pg/m³, which is a summation of MDLs of GOM and PBM (2 pg/m³ each). The error
181 fractions were assumed to be 15% of concentrations for Hg forms and 10% of
182 concentrations for other compounds. This is because most of the measured GOM and
183 PBM concentrations have low concentrations near or below MDL as seen in Tables
184 1-2; thus have large uncertainties as pointed out by Croghan and Egeghy (2003).
185 Following Polissar et al. (1998), constant uncertainties (100%, 200% and 1000% of
186 the mean/median for GEM, PBM and GOM, respectively) were used for imputed Hg
187 concentrations, based on the uncertainty distributions of the below MDL values in the
188 two base cases. This is to down weight the imputed values.

189 The so called “total variable” (e.g. PM) was not used because this study focused
190 on speciated Hg and input variables also include both PM ions and gaseous pollutants.

191 No variables or samples were excluded after input data screening to reflect all
192 observations. No variables were down-weighted, with the exception of imputed values,
193 because runs with and without GOM and PBM categorized as “weak” led to similar
194 results. Other PMF input parameters include: the number of runs was set to 20 to
195 enable stability evaluation, and the best run was used; the number of the starting seed
196 was set to 25.

197 PMF outputs used in this study include source profiles, model performances and
198 factor contributions. Different numbers of factors were also analyzed and the
199 four-factor results had the best interpretability (Liao, 2016). Therefore, fFour factors
200 were retained in each case. Detailed analysis is presented as Supplemental
201 Information (SI), which support the stability of PMF runs, and justify the final
202 solution and the number of factors chosen. The factors were interpreted based on the
203 comparison of the major variables ($\geq 25\%$) in each of the four factors to markers and
204 source profiles in the literature, taking into consideration NPRI emission sources.

205 Various methods have been employed to evaluate receptor models’ performance
206 (e.g. Belis et al., 2015a, 2005b; Cesari et al., 2016). In this study, stability indexes of
207 model runs, scaled residual plot, Obs/Pred scatter plot and Obs/Pred time series were
208 used to evaluate the model performances for speciated Hg. The impact of each data
209 treatment method on PMF results was assessed, taking into consideration
210 interpretability of the factors and model performance of the three Hg forms.

211 212 **PCA**

213 The PCA source apportionment analysis using speciated Hg in 2009 and 2010
214 was already conducted in another study (Cheng et al., 2013). In this study, different
215 cases were investigated, as listed in Table 4. Briefly, all compounds were included to
216 enable comparison with PMF results (Case 2009 and Case 2010), instead of removing
217 some air pollutants as in Cheng et al. (2013) due to a lack of correlation between those
218 air pollutants and atmospheric Hg. Pairwise deletion of missing values in Cheng et al.
219 (2013) was replaced with listwise deletion to be consistent with the PMF model input
220 which must be a complete data matrix. As shown in Table 4, there is a requirement of

221 sample size in order to obtain statistically stable source apportionment results (Henry
222 et al., 1984; Thurston and Spengler, 1985). Our dataset meet the more restrictive
223 requirement by Thurston and Spengler (1985) in both years, by a margin of 90-300 in
224 2009 and 216-300 in 2010 (Tables 3-4).

225 The PCA runs were conducted using SPSS 22.0 (IBM Corp., USA). Cases
226 09-C&M and Case 10-C&M were included to evaluate the effects of weather
227 conditions on factor identification. The dimensions of the reference cases in PMF
228 model and PCA are the same. After including the meteorological parameters in PCA
229 input, the dimensions of the input data are slightly smaller. The components with
230 eigenvalues greater than 1 were retained for further analysis, following the Kaiser
231 Criterion (Kaiser, 1960). An oblique rotation method was used to verify the
232 inter-correlations between the components. Principal components after Varimax
233 rotation were interpreted by comparing the major variables (loadings > 0.25) of the
234 component with the outcomes of other studies, and by checking NPRI sources in the
235 region (Table S1).

236

237 3. Results and discussion

238 3.1 PMF - base cases

239 In this section, only the two base cases, Case 2009 and Case 2010 are considered.

240 *PMF sources*

241 Table 5 and Figures ~~24-32~~ present percent concentration of each pollutant
242 apportioned to each of the four factors. Factor 1 was named Combustion Emission
243 due to large contributions of SO_4^{2-} (64%) and HNO_3 (54%) and a moderate
244 contribution of GOM (31%) (Table 5). Combustion Emission includes fuel
245 combustion and biomass burning. The small contributions of Ca^+ (19%) and K^+ (22%)
246 in this factor suggest a minor impact of biomass burning. SO_2 and NO_x are precursors
247 of SO_4^{2-} and HNO_3 , respectively. These precursors are from combustion sources and
248 probably oxidized during the transport from sources to receptor sites (Liu et al., 2007).
249 The presence of GOM is consistent with the combustion emission which is one of the
250 GOM sources (Carpi, 1997). There ~~was~~ were little NH_3 emissions from point sources
251 near the study site (Table S1). Thus, the presence of NH_4^+ (71%) should be related to
252 the transport and transformation of NH_3 from agriculture emissions as well as other

253 physical and chemical processes (e.g., aqueous phase chemistry, condensational
254 growth, droplet evaporation) producing NH_4^+ (Zhang et al., 2008; Pitchford et al.,
255 2009). In this factor, the molar ratio of NH_4^+ to SO_4^{2-} is 1.7, although some observed
256 profiles having ratios greater than 2 (Lee et al., 1999). Ratios less than 2 suggest
257 insufficient amount of NH_3 to neutralize H_2SO_4 thus H_2SO_4 will react with other
258 compounds to form sulfate (Pavlovic et al., 2006; Zhang et al., 2008). The moderate
259 contribution of PM (42%) is consistent with the presence of particulate SO_4^{2-} and
260 NH_4^+ . Also, SO_4^{2-} accounted for over 50% of PM mass (Table 1). In addition to a lack
261 of major combustion facilities nearby (Table S1), a strong correlation between SO_4^{2-}
262 and NH_4^+ (Tables S2-S3) also suggests formation of secondary aerosols. Therefore,
263 this factor suggests transported plumes instead of fresh emissions.

264 Factor 2 was assigned to Industrial Sulfur. The major variables PBM and SO_2 are
265 indicators of coal combustion (Huang et al., 2010). The minor contributions of HNO_3
266 and NO_3^- also suggest combustion sources because their precursor, NO_x , is mainly
267 released by combustion sources (Liu et al., 2007). However, there were no
268 combustion sources emitting Hg compounds near the KEJ site in 2009 (Table S1).
269 Therefore, this factor is more likely related to industrial sources in the region. As
270 shown in Table S1, point sources of industrial sulfur in the province of Nova Scotia
271 include tire production, engineered wood production, food industry, and universities.
272 Coal-fired power plants and metal production are major sources of sulfur; however
273 there are no combustion sources close to the sampling site. These sources are located
274 in eastern U.S., which could be transported to the site. Mobile sources of sulfur are
275 ships and vessels from nearby ports (Cheng et al., 2013).

276 Factor 3 was named Photochemical Process and Re-emission of Hg due to the
277 high contributions of ozone (72%), GEM (76%), GOM (69%), PBM (63%), and
278 moderate contributions of Ca^{2+} (45%) and K^+ (37%). The high contribution of ozone
279 indicates an ozone rich environment, resulting in oxidation of GEM to GOM and the
280 sequential formation of PBM (Pal and Ariva, 2004; Liu et al., 2007). Although results
281 of recent studies show that the reaction rate of Hg and ozone has large uncertainties,
282 the oxidation of Hg by bromine is very fast (Goodsite et al., 2004). The KEJ site is
283 near the Atlantic, making the oxidation of Hg by bromine applicable. The presence of
284 K^+ is related to soil emission or biomass burning (Andersen et al., 2007), while Ca^{2+}

285 is related to soil/crustal. The site is located in Kejimikujik National Park. Therefore, it
286 is under the impact of soil emission, emission from the nearby biomass-fired power
287 station (Table S1), and transported biomass combustion. It was estimated that
288 re-emission of Hg from biomass burning and land surfaces contributed 13% and 34%
289 of the global re-emission budget, respectively (Pirrone et al., 2010). Thus, the high
290 contribution of GEM may be attributable to the re-emission of GEM. The emission
291 from soil and biomass combustion was also identified in the PCA study at this site
292 (Cheng et al., 2013). An examination of the time series factor profiles revealed that
293 model-reproduced K⁺, O₃ and GEM, GOM, PBM concentrations (in this factor) were
294 rather smooth. The impact of biomass burning seems to be small in this factor due to a
295 lack of high K⁺, O₃, and Hg concentration periods or episodes identified. The
296 relatively stable patterns of K⁺ and GEM suggest re-emission of GEM while GOM
297 was high in spring with elevated O₃, indicating enhanced photochemistry.

298 Factor 4 has high contributions of Cl⁻ (100%), Mg²⁺ (82%) and Na⁺ (86%) and
299 moderate contributions of Ca²⁺ (31%), K⁺ (39%) and NO₃⁻ (40%). The presence of
300 Na⁺, Mg²⁺, and Cl⁻ indicates marine aerosols because these elements are rich in sea
301 water (Huang et al., 1999). The strong correlations among these three compounds
302 (≥ 0.89 , Tables S2-S3) also suggest a common source. As the sampling site is located
303 near the Atlantic, the presence of marine aerosols is reasonable. Major production
304 pathways of NO₃⁻ include reaction of HNO₃ with NH₃, sea salt and soil dust
305 (Pakkanen, 1996). In this factor, the NO₃⁻ is probably related to the reaction of HNO₃
306 and sea salt. Thus, this factor was named Sea Salt.

307 As seen in Table 5 and Figures 24-32, the same four factors were identified in
308 year 2009 and 2010. The profiles of each factor were also largely consistent between
309 the two years. Factor 1 in 2010 is similar to the factor named Combustion Emission in
310 Case 2009. However, this factor lacks PM (not available in 2010) and has a higher
311 contribution from K⁺, which may relate to biomass burning. This factor is assigned to
312 the same name as in 2009 because the presence of SO₄²⁻ and HNO₃ is enough to
313 identify combustion process (Liu et al., 2007). It should be noted that this factor has a
314 much smaller contribution of GOM than in 2009. This may be due to a large reduction
315 in SO₂ emissions (2.42 million tons or 32% reduction) from coal-fired power plants
316 across the United States between 2008 and 2010 (US EPA, 2011). Large reductions in

317 Hg (-39%) and SO₂ (-35%) emissions also occurred in Nova Scotia between 2009 and
318 2010, as seen in Table S1. However, reduction in Hg emissions is only reflected on
319 GOM (-75%), while GEM decreased a little and PBM increased slightly. Moreover,
320 the long term effects of emission reductions on Hg concentrations and source
321 contributions should be investigated.

322 The major variables of factor 2 are also similar to those of the factor Industrial
323 Sulfur in Case 2009. However, this factor has a moderate contribution of GOM
324 instead of PBM in 2009. Factor 3 has similar major variables as the factor named
325 Photochemistry and Re-emission in Case 2009. Factor 4 is dominated by Cl⁻ (100%),
326 Na⁺ (83%) and Mg²⁺ (75%). This factor was named Sea Salt as in Case 2009.

327 ***PMF source contributions***

328 The PMF factor contributions of the two base cases are presented in Table S4
329 (Case 2009) and Table S5 (Case 2010). In both years, factor Photochemistry and
330 Re-emission had the largest contributions to GEM (averaging 77% and 79% in 2009
331 and 2010, respectively), GOM (70% and 67%), and PBM (69% and 80%) among all
332 four factors. In other words, ambient Hg concentrations at the KEJ site were
333 dominated by photochemistry and re-emission of Hg. Industrial Sulfur had moderate
334 contributions to GOM (average, 29%) in 2010 instead of PBM in 2009 (21%).
335 Combustion Emission contributed 25% of GOM in 2009 but 11% each of GEM and
336 PBM in 2010. The factor Sea Salt only had minor contribution to GEM (14% in 2009
337 and 9% in 2010) and PBM (<10% in both years). This is not unexpected because
338 GEM is likely to be oxidized to GOM by the *in situ* photochemical process under the
339 bromine-rich environment (Obrist et al., 2011). However, this factor has no
340 contribution to GOM because it was estimated that >80% of GOM in the marine
341 boundary layer is absorbed by sea salt aerosols and it is sequentially deposited onto
342 the earth's surface where evasion occurs (Holmes et al., 2009).

343 ***PMF model performance***

344 Among the three Hg forms, GEM had the best performances in terms of scaled
345 (i.e. standardized) residual because it had normal distribution and fewer absolute
346 values of scaled residual greater than 3 in both years (Case 2009 and Case 2010, Table
347 6). Table 6 also lists the coefficient of determination (R²) and the slope of the
348 regression line for speciated Hg in Obs/Pred scatter plot (Figures ~~S51~~-~~S62~~), to

349 evaluate the overall model-measurement agreement. Between the two years, the
350 agreement was better with GEM in 2010 and PBM in 2009 because of higher R^2
351 values and slope closer to 1. The low values of R^2 and slope in both years indicate the
352 agreement was poor for GOM.

353 The Obs/Pred time series of the three Hg forms reveal the model's ability to
354 reproduce the observational concentrations on a day-to-day basis. In Case 2009, the
355 | Obs/Pred time series (Figure S73) were split into three time periods by the data gaps,
356 | January to February (period 1), March to July (period 2), and October to December
357 | (period 3). GEM had better performances than the other two forms because the peak
358 | values were reproduced by the model in all three periods. However, the modeled
359 | values in period 3 are too low compared to observed concentrations, leading to a
360 | lower R^2 (Table 6). The performance for PBM is better than GOM because the
361 | ~~model-reproduced~~~~predicted~~ concentrations tracked the observed concentrations well
362 | in period 2. However, PBM concentrations were underestimated and overestimated by
363 | the model in period 1 and period 3, respectively. The GOM concentrations were not
364 | reproduced well with unmatched peak values in period 2, and there was a clear
365 | separation of observed and ~~model-reproduced~~~~predicted~~ trend lines in periods 1 and 3,
366 | leading to over prediction.

367 | In Case 2010, the time series (Figure S84) were split into two periods,
368 | January-June (period 1) and July-December (period 2), based on a clearly visible
369 | overestimation of GOM concentrations in the second period. The ~~reproduced~~~~predicted~~
370 | GEM concentrations tracked the trend of observations well in both periods but with
371 | more fluctuations. The model was unable to reproduce high GOM concentrations in
372 | period 1. For PBM, the ~~reproduced~~~~predicted~~ concentration was rather flat, missing
373 | completely the high concentration episode in spring 2010.

374 The model-measurement agreement was further quantified with the ratios of
375 | ~~reproduced~~~~predicted~~ to observed concentrations (Pred/Obs ratio, Figure 43). In both
376 | years, the ~~reproduced~~~~predicted~~ GEM agreed well with the observed concentrations as
377 | supported by the small range of Pred/Obs ratios (0.56-1.32 in 2009, 0.42-1.43 in 2010)
378 | and mean ratios approaching 1 (0.97 and 0.98). On an annual basis, the observed
379 | GEM concentrations were also well reproduced because the ratios of

380 | ~~reproduced~~~~predicted~~ to observed annual means (annual Predmean/Obsmean) were
381 | almost 1 (0.97 and 0.98) (Tables S4-S5). Compared with GOM, PBM had better
382 | agreement between the ~~reproduced~~~~predicted~~ and observed concentrations with a
383 | smaller range of Pred/Obs ratios (0.40-13.4 and 0.14-18.3 vs. 0.13-53 and 0-193) and
384 | mean ratios closer to 1 (2.09 and 1.98 vs. 5.89 and 4.44). In spite of large sample to
385 | sample variability in the Pred/Obs ratios, the model performance was very good for
386 | PBM (annual Predmean/Obsmean ratio of 1.03 and 1) and reasonable for GOM (0.86
387 | and 1.34) in reproducing annual means.

388 | *Comparison between PMF in year 2009 and 2010*

389 | Overall, the interpretability of the factors was similar in the two years. The same
390 | factors were observed in 2009 and 2010, and most factor contributions were highly
391 | consistent between the two years. Among the three Hg forms, PMF reproduced GEM
392 | concentrations well in both years. Possible reasons of poor performance on PBM and
393 | GOM include PMF uncertainties for modeling pollutants that undergo various
394 | transformation processes, unlike the modeling of only aerosols. PMF does not account
395 | for chemical reactions that may occur as the air mass travels from source to receptor.
396 | Another likely reason is lower concentration levels and much higher percentages of
397 | readings below MDL (Tables 1-2) leading to large uncertainties. However, the
398 | differences in sample size (161 in 2009 vs. 290 in 2010) and fractions of below MDL
399 | values (Tables 1-2) alone could not explain the mixed results of poor performance
400 | on GOM in 2009 and PBM in 2010. Further examination of time series (Figures S73
401 | and S84) suggests that the reduced performance could also be attributable to high
402 | concentration episodes in GOM in 2009 and PBM in 2010. The impact of Hg data
403 | treatment on PMF results was investigated and the results are presented in section 3.4.

404

405 | **3.2 PCA components**

406 | *Case 09-C*

407 | The component loadings of Case 09-C are presented in Table 7. PC1 was named
408 | Combustion/industrial Emission due to positive loadings of PBM, PM, O₃, SO₂,
409 | HNO₃, Ca²⁺, K⁺, NO₃⁻, NH₄⁺, and SO₄²⁻. Most major compounds except O₃ were also
410 | found in a component named “transport of combustion and industrial emissions” in
411 | another PCA study using the same dataset (Cheng et al., 2013). The high loadings of

412 secondary pollutants HNO_3 , NO_3^- , and SO_4^{2-} indicate the factor represents transport of
413 combustion/industrial emission because their precursors (NO_x and SO_2) are mainly
414 emitted by combustion/industrial sources (Liu et al., 2007). The precursors may be
415 oxidized during the transport process. The moderate loading of O_3 is also related to
416 the transport of combustion emission because the precursors of O_3 (NO_x and VOC)
417 are emitted from mobile and stationary combustion sources. Ammonia is likely related
418 to the transport of agriculture emissions and reaction of NH_3 and H_2SO_4 or HNO_3
419 (Pichford et al., 2009).

420 PC2 has high loadings of Na^+ , Mg^{2+} , Cl^- , and K^+ and moderate loadings of
421 Ca^{2+} . Those compounds indicate marine aerosols (Huang et al., 1999). The moderate
422 loading of NO_3^- is likely due to the reaction of HNO_3 and sea salt (Pakkanen, 1996).
423 As in the PMF factor interpretation, the identification of component Sea Salt is
424 relevant because the monitoring site is near the Atlantic.

425 PC3 has positive loadings of GEM, GOM, PBM and O_3 . The positive loadings on
426 O_3 and GOM indicate the photochemical production of GOM (Huang et al., 2010).
427 The positive loading of GEM is somewhat unexpected because the photochemical
428 production of GOM consumes GEM thus leading to opposite signs of GEM and GOM
429 (e.g. Huang et al., 2010). However, daily average concentrations were used in this
430 study instead of two-hour means in Huang et al. (2010). The daily GEM and GOM
431 were indeed positively correlated ($r=0.37$ in 2009, Table S2; 0.31 in 2010, Table S3).
432 Using the same dataset, Cheng et al. (2013) conducted further analysis on O_3
433 concentrations and %GOM/TGM (TGM=GEM+GOM) ratios. The ratio is indicative
434 of the degree of oxidation. The results showed that the %GOM/TGM ratio increased
435 with O_3 when O_3 concentrations were greater than 40 ppb, suggesting gas phase
436 oxidation of Hg at this coastal site. Therefore, this factor was named Photochemical
437 Production of GOM.

438 PC4 represents Gas-particle Partitioning of Hg. The negative loading of PBM and
439 the positive loading of GOM indicate the partition process. The positive loadings of
440 Ca^{2+} and K^+ suggest soil aerosols (Cheng et al., 2012) which could be abundant at the
441 Kejimikujik National Park.

442 Three out of four components (Combustion/industrial Emission, Photochemical
443 Production of GOM and Gas-particle Partitioning of Hg) have significant association

444 with ambient Hg concentrations at the site, while Sea Salt has little.

445 *Case 09-C&M*

446 Five principal components are extracted when meteorological data were included
447 in PCA (Case 09-C&M, Table 7). The loadings in PC1-PC4 are similar with the
448 loadings of PC1, PC2, PC4, PC3 in Case 09-C, respectively. Thus the names of those
449 four components were retained. The inclusion of meteorological parameters resulted
450 in small loadings of relative humidity (-0.26) in PC1 and wind speed (0.32) in PC2, as
451 well as a moderate loading of wind speed (0.52) in PC4. A large loading of
452 temperature (0.94) was observed in PC3. The opposite signs of temperature and PBM
453 are consistent with the gas-particle partitioning process because low temperatures
454 favor the formation of PBM (Rutter and Schauer, 2007). The lack of GEM in PC3
455 (Case 09-C&M) did not affect the identification of this factor, because the partitioning
456 of GEM onto particles is much weaker than that of GOM (Liu et al., 2007).

457 PC5 was derived mostly from meteorological variables. The negative loading of
458 GOM and positive loadings of relative humidity and precipitation suggest removal of
459 GOM by dew, cloud and precipitation (Cheng et al., 2013). The loading of GOM is
460 small, nonetheless consistent with the wet deposition process because GOM is more
461 easily removed compared to GEM due to its higher water solubility (Gaffney and
462 Marley, 2014). Therefore, this component was named Hg Wet Deposition.

463 Similar to Case 09-C, all components except Sea Salt are associated with ambient
464 Hg concentrations. After the inclusion of meteorological data, each factor contains at
465 least one meteorological parameter. The presence of meteorological variables did not
466 contribute to the determination of the components except a new component Hg wet
467 deposition was identified.

468 *Case 10-C*

469 The component loadings of Case 10-C are listed in Table 8. PC1 was named
470 Combustion Emission. The positive loadings of HNO₃, NO₃⁻ and SO₄²⁻ are indicative
471 of transport of combustion emission because their precursors (NO₂ and SO₂) are
472 mainly released by combustion emissions (Liu et al., 2007). The high positive loading
473 of NH₄⁺ represents transport of agriculture emissions of ammonia which may react
474 with H₂SO₄ and HNO₃ during the transport process (Pitchford et al., 2009). The
475 | positive loadings of Ca²⁺ and K⁺ indicate biomass burning from wildfires or

476 biomass-fueled power station (Andersen et al., 2007).

477 PC2 was named Sea Salt due to high loadings of Na^+ , Mg^{2+} , and Cl^- , because
478 these three compounds are rich in sea water (Huang et al., 1999). PC3 has the same
479 major variables as the component Photochemical Production of GOM in 2009.
480 Therefore, PC3 was also named as such.

481 PC4 was assigned to Industrial Source. The positive loadings of GOM and SO_2
482 indicate coal combustion (Lynam and Keeler, 2006), although no combustion facilities
483 were reported near the KEJ site in 2010 (Table S1). The positive loadings of SO_4^{2-} and
484 HNO_3 are consistent with the transport of industrial emissions which release their
485 precursors, SO_2 and NO_x (Liu et al., 2007). Therefore, this factor was named
486 Industrial Source. Two out of four factors (i.e. Photochemical Production of GOM and
487 Industrial source) have significant association with Hg compounds.

488 *Case 10-C&M*

489 As shown in Table 8, five principal components are extracted in Case 10-C&M.
490 The loadings in PC1-PC3 and PC5 are similar with the loadings of PC1-PC4 in Case
491 10-C, respectively. Thus the names of those four components were retained. The
492 additional negative loading of temperature (-0.52, Table 8) and positive loading of
493 wind speed (0.52, Table 8) in PC3 may indicate colder air flows from the north
494 containing more O_3 and GOM (Cheng et al., 2013). This is reasonable because Hg
495 sources in Nova Scotia were mainly located north of the sampling site (Figure 1). PC4
496 in Case 10-C&M was named Hg Wet Deposition due to negative loadings of GOM
497 and PBM and positive loadings of relative humidity, wind speed and precipitation,
498 similar with PC5 in Case 09-C&M (Table 7). Three out of five components (i.e.
499 Photochemical Production of GOM, Industrial Source, and Hg Wet Deposition) were
500 associated with Hg concentrations. The influence of meteorological data on
501 identification of components was also similar to in 2009. For Case 09-C&M
502 10-C&M, a detailed comparison of PCA results in this study and that in Cheng et al.
503 (2013) can be found in Liao (2016).

504 *Comparison between PCA in year 2009 and 2010*

505 In each year, four components were extracted in PCA with air pollutants only.
506 The two common factors between the two years are Photochemical Production of
507 GOM and Sea Salt. The former has a strong association with Hg compounds, while

508 the latter has little. Component Gas-particle Partitioning of Hg was only identified in
509 2009, likely due to a lower percentage of PBM readings <MDL than those in 2010
510 (Table 9, Case 2009 and 2010). It is also consistent with strong correlations between
511 temperature as well as GOM and PBM ($r=0.46$ and -0.43 , Table S2) in 2009 but
512 non-significant or weak correlations ($r=-0.04$, and -0.16 , Table S3) in 2010.

513 The component Combustion/industrial Emission in 2009 affected PBM and SO₂
514 levels. It was split into two components in 2010, Combustion Emission and Industrial
515 Source. The former was no longer strongly associated with any of the three Hg forms,
516 while the latter was associated with GOM and SO₂. This is probably due to the
517 reduction of coal combustion in Canada and the USA, evident by lower provincial Hg
518 (reduction of 39%) and SO₂ emissions (-35%) in 2010 (Table S1). The reductions in
519 GEM, GOM, and SO₂ concentrations at the KEJ site were 3%, 75%, and 43%
520 respectively in 2010 (Tables 1-2). The shifting of PBM & SO₂ relationship in 2009 to
521 GOM & SO₂ in 2010 is sustained by a strong correlation between PBM and SO₂
522 ($r=0.63$, Table S2) in 2009, but little correlation ($r=0.06$) accompanied by a moderate
523 correlation between GOM and SO₂ ($r=0.30$) (Table S3) in 2010. The shift is also
524 consistent with the PMF results where Industrial Sulfur accounted for 21% of PBM in
525 2009 (Table S4) but 29% of GOM in 2010 (Table S5).

526 In both years, inclusion of meteorological parameters did not affect the
527 identification of the four factors from air concentrations. However, relative humidity
528 and precipitation yielded an additional component named Hg Wet Deposition.

529 Overall, the PCA results were largely consistent between the two years, in terms
530 of the number of components, impact of meteorological parameters, and major
531 processes associated with ambient Hg. The changing emissions/concentrations and the
532 resultant correlations among monitored air pollutants from one year to another are
533 reflected in the limited shifting of variable loadings.

534

535 3.3 Comparison of PMF and PCA results

536 The PCA loadings and the factor profiles as well as factor contributions in PMF
537 model have very different meanings. In PCA, variables with large loading indicate
538 their correlation or association with that component derived from all samples. In PMF,
539 presence of variables in profiles indicates their contribution to that source/process

540 derived from all samples, while the contribution values are further quantified in
541 source contribution tables of each sample. Therefore, a direct comparison between the
542 PMF and PCA results is not feasible. However, the similarities and differences in the
543 major sources/processes identified by each approach, chemical markers in each factor
544 profile or component, and the impact/association of factors/components on Hg could
545 reveal strength and weakness of each method.

546 A comparison of Table 5 and Tables 7-8 (cases with air concentrations only)
547 shows that Na^+ , Cl^- , and Mg^{2+} are markers of Sea Salt in both PMF and PCA.
548 Similarly, GEM, GOM, PBM and O_3 indicate Photochemistry. Both methods suggest
549 strong contribution to or association between Hg compounds and photochemistry, but
550 weak with Sea Salt. Both methods identified combustion and industrial sources, while
551 the variables in factors/components differed to some extent. Furthermore, combustion
552 and industrial were separate sources in PMF in both years and in PCA in 2010, but
553 combined as one component in PCA in 2009. Overall, PMF profiles are more
554 consistent between the two years, while the PCA loadings are more sensitive to
555 correlation among variables. However, the shift of PBM & SO_2 to GOM & SO_2
556 loadings in PCA between the two years is consistent with the shift of those two pairs
557 in Combustion & Industrial Sulfur profiles/contributions in PMF. On the other hand,
558 Gas-particle Partitioning of Hg was only recognized in PCA. This is because the
559 identification of this factor relies on negative association between PBM and GOM
560 (Table 7), but such association is not reflected in PMF due to its non-negative nature.
561 This is one of the limitations of PMF. Furthermore, the inclusion of meteorological
562 conditions in PCA enables identification of a new component related ~~with-to~~ weather
563 conditions. The good agreement between PMF and PCA outputs is consistent with a
564 comparison of receptor models in PM source appointment (Viana et al., 2007; [Cesari](#)
565 [et al., 2016](#)). A common weakness of PCA and PMF is the suggestiveness of
566 factors/components. Other techniques, such as back trajectories, have been used in
567 previous studies to verify some factors (Cheng et al., 2015). Overall, when
568 accompanied by model performance evaluation, PMF results are with more
569 confidence.

570

571 3.4 Sensitivity of PMF results to data treatment

572 **3.4.1 Year 2009**

573 *Case 09+mean & Case 09+median*

574 | The factor profiles of the six PMF cases in 2009 are displayed in Figure 24. In
575 | Case 09+mean and Case 09+median, all four factors have similar profiles as in Case
576 | 2009. Compare with the base case, factor 3 (Photochemistry and Re-emission of Hg)
577 | has a higher contribution by NO_3^- , however it is common to observe NO_3^- from soil
578 | emissions (Parmar et al., 2001). GOM has a much smaller contribution in factor 1
579 | (Combustion Emission) (Figure 24, Table S4). This is likely because the correlation
580 | coefficients between GOM, NH_4^+ and SO_4^{2-} become insignificant after imputation
581 | (Table S6). Consequently, GOM is not strongly related to that factor which is
582 | dominated by NH_4^+ and SO_4^{2-} . Changing correlation among variables is a
583 | shortcoming of imputation (Huang et al., 1999).

584

585 *Case 09+RM & Case 09-RM*

586 | As shown in Figure 24 and Table S4, by combining GOM and PBM into RM,
587 | RM replaced PBM instead of GOM in related factors as major variables with similar
588 | contributions. This is because the median concentration of PBM is approximately 5
589 | time of the median concentration of GOM (Table 9). Once these two forms are
590 | combined to RM, the variance of RM is dominated by PBM. The presences of other
591 | compounds including GEM in factor profiles/contributions in these two cases are
592 | similar to those in Case 2009.

593 *Case 09ScaleRM*

594 | The factor profiles were similar to those in Case 2009 (Figure 24). The same can
595 | be said about factor contributions to speciated Hg (Table S4).

596 *Performance*

597 | Case 09-RM, Case 09+RM and Case 09ScaleRM have similar performances with
598 | Case 2009, on distribution of scaled residuals (Table 6). Imputation (Case 09+mean
599 | and Case 09-median) worsened the performance because the scaled residuals are
600 | concentrated near zero for gaseous Hg.

601 In terms of the coefficients of determination (R^2) and the slopes of the regression
602 line for speciated Hg in Obs/Pred scatter plot (Table 6, Figure S54), imputation (Case
603 09+mean and Case 09+median) deteriorated [the](#) PMF performance compared to the
604 base case. This is not unexpected because the use of a constant imputation value
605 reduced the variance in observed concentrations (Table 9). The similar performances
606 on GEM in Case 2009, Case 09+RM, Case 09-RM, and Case 09ScaleRM indicate
607 combining, excluding, or scaling GOM and PBM, respectively, did not affect the
608 performance on GEM. The performances on RM are similar to that of PBM in Case
609 2009 because the RM concentrations are dominated by PBM. Using scaling factors to
610 increase GOM and PBM concentrations resulted in better performances on those two
611 forms than in the base case. This is attributable to a significant reduction in percent of
612 concentrations below MDL (Table 9).

613 The changes in model performance are more evident in the observed and
614 ~~reproduced/predicted~~ time series (Figure S73). Compared with the base case,
615 imputation led to more fluctuation in the ~~reproduced/predicted~~ GEM values, thus
616 slightly worse. RM had better model-measurement agreement than GOM or PBM as
617 individual compound. The agreement was also improved by scaling GOM or PBM.
618 The peak values (PBM in period 1 and both forms in period 2) were better reproduced
619 and the over prediction in period 3 with low concentrations was greatly corrected.

620 Compared with the base case, the distributions of the ratios of
621 ~~reproduced/predicted~~ to observed Hg concentrations and the ratio of
622 ~~reproduced/predicted~~ to observe annual means changed little for GEM among the six
623 cases (Figure 43, Table S4). Scaling GOM and PBM improved model-measurement
624 agreement of those two forms, evident by a much narrower range and a shift toward
625 smaller values in the distribution of ratios.

626 3.4.2 Year 2010

627 *Case 10+mean & Case 10+median*

628 Factor profiles (Figure 23) and contributions (Table S5) after imputation have
629 minor changes compared to those in Case 2010. However, less changes ~~were~~-was
630 observed with the use of median imputation. The smaller er deviations after imputations
631 is probably because only a small fraction (4%) of Hg concentrations ~~were~~-was missing

632 in 2010 than in 2009 (31-41%). Although HNO₃, SO₂, and inorganic ions have up to
633 19% missing values (Table 2), the correlations between each of the three Hg forms
634 and other compounds changed little (Table S7).

635 *Case 10+RM & Case 10-RM*

636 | The impact of combining or removing GOM and PBM (Figure 23, Table S5) is
637 the same as in 2009. The dominance of PBM in RM is stronger in 2010 with the ratio
638 of median PBM to median GOM concentration being approximately 10 (Table 9).

639 Overall, excluding or combining GOM and PBM did not affect the source
640 | identification in PMF model in both years (Figures 24 and 32). However, the
641 identification of the factors relying on GOM or PBM only (e.g. gas-particle
642 partitioning of Hg) may be affected after combining or excluding GOM and PBM.
643 In this study, such factors were not encountered in PMF. Nonetheless, excluding or
644 combining GOM and PBM did affect the source contributions. After combining GOM
645 and PBM, factors contributing to GOM only (Combustion Emission, 2009; Industrial
646 Sulfur, 2010, Table 10) did not contribute to any Hg forms, and the factor contributing
647 to PBM only (Industrial Sulfur, 2009) was contributing to RM, due to dominance of
648 PBM in RM. In both years, using three Hg forms instead of GEM only led to more Hg
649 sources/processes identified. Therefore, monitoring speciated Hg could help us better
650 understand Hg cycling.

651 *Case 10ScaleRM*

652 The factor profiles and contributions of Case 10ScaleRM are similar to those in
653 | Case 2010 (Figure 23, Table S5). A noticeable deviation is the much smaller
654 contribution by GOM in factor 2 compared to Case 2010. However, factor 2 was still
655 assigned to Industrial Sulfur because of the presence of SO₂ and NO₃⁻.

656 *Performance*

657 Firstly, the distribution of scaled residuals as well as R² value and the slope of the
658 | regression line for speciated Hg in Obs/Pred scatter plot were evaluated for the six
659 cases (Table 6, Figure S62). Similar to 2009, the comparable performances observed
660 in Case 2010, Case 10-RM, Case 10+RM, and Case 10ScaleRM indicate that the
661 model performance on GEM is insensitive to excluding, scaling, or combining GOM

662 and PBM to RM. Case 10ScaleRM also has the best performances on GOM and PBM
663 among all the cases in 2010. Unlike in 2009, the negative impact of imputation was
664 smaller when median value was used, compared with the mean imputation.

665 Secondly, in the observed and ~~reproduced~~~~predicted~~ time series (Figure S84),
666 imputation resulted in more severe fluctuation in ~~reproduced~~~~predicted~~ GEM
667 concentration as in 2009, but less so when median values were used. Scaling of GOM
668 or PBM also improved the reproducibility of day-to-day variability in the observed
669 values, owing to a large reduction in concentrations below MDL (Table 9). Among the
670 6 cases, the most significant change is in PBM with imputation. There were additional
671 high concentration episodes in early 2010 when imputation of non-Hg compounds
672 brought back Hg concentrations otherwise removed by listwise deletion in the base
673 case, leading to increased standard deviation (Table 9). Those peaks were completely
674 missed by the model, leading to deteriorated agreement.

675 Finally, the distributions of the ratios of ~~reproduced~~~~predicted~~ to observe Hg
676 concentrations and the ratio of ~~reproduced~~~~predicted~~ to observe annual means changed
677 little among the first five cases in 2010 (Figure 43 and Table S5). The exceptions are
678 under prediction of the annual mean of PBM in the two imputation cases and over
679 prediction for RM. Compared with the base case, the distribution of ratios for GOM
680 and PBM became narrower and shifted toward smaller values, but leading to under
681 prediction of PBM.

682 3.4.3 Comparison of 2009 and 2010 among different data treatments

683 The different characteristics of Hg forms led to different impact of data treatment
684 on model results and performances in the two years. Imputation using geometric mean
685 and median values led to minor changes in factor profiles in both years, with more
686 variations in contributions of Hg forms in 2009 but non-mercury compounds in 2010.
687 This is likely because the Hg and non-Hg compounds were missing at a larger
688 percentage in 2009 and 2010, respectively. The lack of significant impact is likely due
689 to already high sample to compound ratios (161 samples/15 compounds in 2009, 290
690 samples/14 compounds in 2010, Tables 1-3). Huang et al. (1999) have reported that
691 mean imputation generally yielded better PMF results than listwise deletion,
692 especially with higher percentage of missing values. Particularly, composition of

693 crustal and marine factors were closer to those of crust and sea water. Imputation
694 resulted in degraded performance on all three Hg forms, but for different reasons. For
695 GEM, it is largely due to more fluctuation than the already over predicted one in the
696 base case in both years. For PBM in 2010, the peak values otherwise removed in
697 listwise deletion (base case) are beyond the model's ability to match. This seems to be
698 a random occurrence and is an uncertainty of imputation. Between geometric mean
699 and median imputations, the impact was similar in both years for each of the three Hg
700 forms. The exception is with median imputation in 2010, there was less deviation in
701 factor profile/contribution from the base case. The reason is unclear because the
702 difference in geometric mean and median was very small for GEM in both years and
703 slightly greater in 2009 for GOM and PBM (Tables 1-2).

704 In both years, some changes in the factor profiles and factor contributions but
705 little changes in model performances were observed in the cases excluding GOM and
706 PBM. Scaling GOM and PBM or combining them into RM improved
707 model-measurement agreement, suggesting the approach is effective in both years in
708 spite of large percentages of below MDL values (GOM, 78% in 2009 vs. 96% in 2010;
709 PBM, 48% in 2009 vs. 46% in 2010, Tables 1-2). The improvement is largely
710 attributable to reduction in concentrations below MDL (Table 9) which in turn
711 reduced PMF uncertainty expressed in equation (2). Another benefit of using a
712 variable scaling factor is reduced data variability as indicated by smaller coefficients
713 of variation in Table 9. PMF is ~~more likely better at~~ reproducing ~~well~~ compounds with
714 less variability. However, there is little evidence that the scientific uncertainties of
715 scaled GOM and PBM concentrations are indeed reduced from that of the original
716 dataset.

717

718 **4. Conclusions**

719 Source apportionment analysis was conducted with PMF and PCA using
720 concentrations of speciated Hg and other air pollutants collected at KEJ site in 2009
721 and 2010. Year 2010 was characterized by reduced Hg and SO₂ emissions compared
722 with 2009. However, GOM is more sensitive to the decrease in Hg emissions while
723 GEM and PBM are not, underscoring the benefits of speciated Hg measurements. It

724 was found that consideration of emission inventories and correlation among air
725 pollutants is useful in factor/component interpretation.

726 Using PMF, the nature of each of the four factors identified was the same in 2009
727 and 2010. In both years, ambient concentration of all three Hg forms at the KEJ site
728 were dominated by contributions from factor Photochemistry and Re-emission, and
729 the contribution by Sea Salt was the smallest. However, slight variations between the
730 two years were observed in the contributions by the other two factors (Combustion
731 Emission, Industrial Sulfur).

732 Good agreement was found between PMF and PCA results. In each year, four
733 components were extracted in PCA with air pollutants only. Three or four of them
734 overlapped with factors obtained in PMF. PCA results suggest little association
735 between Hg and Sea Salt, consistent with PMF. Furthermore, PMF and PCA had
736 similar shift of source profile/contribution from one year to another, suggesting both
737 methods were able to respond to changing concentration levels, and interrelationships
738 among the air pollutants. In both years, inclusion of meteorological parameters in
739 PCA led to extraction of an additional component Hg Wet Deposition while the
740 identification of other components was not affected. Therefore, PCA is superb to PMF
741 in terms of identifying factors related to atmospheric processes. With regards to
742 atmospheric processes represented by negative correlation among variables, such as
743 Gas-particle Partitioning of Hg (Table 8), PCA is more likely to identify them because
744 component loadings reflect correlations, while it is difficult for PMF because its
745 variable contributions in source profile are all positive.

746 A comprehensive PMF model performance evaluation was conducted for each of
747 the three Hg forms. Between the two years, the model performance was comparable.
748 In both years, the observed daily GEM concentrations were well reproduced, but
749 relatively poor for GOM and PBM. On an annual basis, the model-measurement
750 | agreements of annual mean concentrations were excellent for GEM, very good for
751 PBM and acceptable for GOM.

752 The sensitivity of PMF results and model performance to different approaches of
753 dealing with missing values and concentrations with large uncertainties was
754 investigated. In our study of more than 160 samples with 15 or 14 air pollutants,

755 increasing the sample size by geometric mean or median imputation of missing values
756 is not effective in improving the model performance. With over 70% GOM and over
757 40% PBM concentrations below MDL in our dataset, the impact of large measurement
758 uncertainties in GOM and PBM is much more significant. Scaling GOM and PBM to
759 increase their concentrations or combining them to reactive mercury is effective in
760 improving the model-measurement agreement. The identification of sources/processes
761 and their contributions to speciated Hg are relatively insensitive to any of the data
762 treatment options considered. The exception is that less sources/processes affecting
763 ambient Hg were identified when GOM and PBM were excluded, further underlining
764 the importance of monitoring speciated Hg.

765 The good agreement between PCA and PMF results in both years is encouraging
766 although these two methods bear little resemblance. PMF partitions observed
767 concentrations by solving mass balance equations, while PCA is a data reduction tool
768 to explain majority of variances in the entire dataset with a small number of
769 components. Our observation was made possible by the use of multiple-year dataset.
770 Future studies should conduct more PMF and PCA comparisons to validate our
771 findings.

772 Overall, PMF results are quantitative and with more confidence with model
773 performance evaluation. However, when ancillary air pollutant data are available, it is
774 recommended to carry out both PCA and PMF simulations to verify the
775 sources/processes identified.

776 Our PMF results suggest that PMF has difficulties reproducing daily
777 concentrations of GOM and PBM, because of high concentration episodes and large
778 uncertainties due to low concentrations and large percentage of below MDL values.
779 More attention should be devoted to those issues in future studies.

780

781 **Acknowledgements:** Funding of this project was provided by Environment Canada
782 and National Sciences and Engineering Research Council of Canada. The authors
783 acknowledge John Dalziel and Rob Tordon of Environment Canada for providing
784 mercury data and US EPA for the PMF model used in this study.

785

786 **References**

- 787 Andersen, Z. J., Wahlin, P., Raaschou-Nielsen, O., Scheike, T., and Loft, S.: Ambient
788 particle source apportionment and daily hospital admissions among children
789 and elderly in Copenhagen, *Journal of Exposure Science and Environmental*
790 *Epidemiology*, 17, 625-636, 10.1038/sj.jes.7500546, 2007.
- 791 [Belis, C. A., Karagulian, F., Larsen, B. R., and Hopke, P. K.: Critical review and](#)
792 [meta-analysis of ambient particulate matter source apportionment using](#)
793 [receptor models in Europe, *Atmos. Environ.*, 69, 94-108,](#)
794 [10.1016/j.atmosenv.2012.11.009, 2013.](#)
- 795 [Belis, C. A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B. R., Gerboles, M.,](#)
796 [and Hopke, P. K.: A New Methodology to Assess the Performance and](#)
797 [Uncertainty of Source Apportionment Models in Intercomparison Exercises,](#)
798 [Atmos. Environ., 119, 35-44, 10.1016/j.atmosenv.2015.08.002, 2015a.](#)
- 799 [Belis, C. A., Karagulian, F., Amato, F., Almeida, M., Artaxo, P., \(...\), Hopke, P. K.:](#)
800 [A New Methodology to Assess the Performance and Uncertainty of Source](#)
801 [Apportionment Models II: The Results of Two European Intercomparison](#)
802 [Exercises, *Atmos. Environ.*, 123, 240-250, 10.1016/j.atmosenv.2015.10.068,](#)
803 [2015b.](#)
- 804 Carpi, A.: Mercury from combustion sources: a review of the chemical species
805 emitted and their transport in the atmosphere, *Water, Air, Soil Pollut.*, 98,
806 241-254, 10.1023/A:1026429911010, 1997.
- 807 [Cesari, D., Amato, F., Pandolfi, M., Alastuey, A., Querol, X., and Contini, D.: An](#)
808 [inter-comparison of PM10 source apportionment using PCA and PMF](#)
809 [receptor models in three European sites, *Environment Science and Pollution*](#)
810 [Research, 23, 15133-15148, 10.1007/s11356-016-6599-z, 2016.](#)
- 811 Croghan, C. W., and Egeghy, P. P.: Methods of dealing with values below the limit of
812 detection using SAS, available at:
813 <http://analytics.ncsu.edu/sesug/2003/SD08-Croghan.pdf> (last access: May 30,
814 2016), 2003.
- 815 Cheng, I., Lu, J., and Song, X.: Studies of potential sources that contributed to
816 atmospheric mercury in Toronto, Canada., *Atmos. Environ.*, 43, 6145-6158,
817 10.1016/j.atmosenv.2009.09.008, 2009.
- 818 Cheng, I., Zhang, L., Blanchard, P., Graydon, J. A., and Louis, V. L. S.:
819 Source-receptor relationships for speciated atmospheric mercury at the
820 Remote Experimental Lakes Area, Northwestern Ontario, Canada, *Atmos.*
821 *Chem. Phys.*, 12, 1903-1922, 10.5194/acp-12-1903-2012, 2012.
- 822 Cheng, I., Zhang, L., Blanchard, P., Dalziel, J., Tordon, R., Huang, J., and Holsen, T.
823 M.: Comparisons of mercury sources and atmospheric mercury processes
824 between a coastal and inland site, *J. Geophys. Res-Atmos.*, 118, 2434-2443,
825 10.1002/jgrd.50169, 2013.
- 826 Cheng, I., Xu, X., and Zhang, L.: Overview of receptor-based source apportionment
827 studies for speciated atmospheric mercury, *Atmos. Chem. Phys.*, 15,
828 7877-7895, 10.5194/acp-15-7877-2015, 2015.
- 829 Cheng, I., Zhang, L., and Xu, X.: Impact of measurement uncertainties on receptor
830 modeling of speciated atmospheric mercury, *Scientific Reports*, 6, 20676,
831 10.1038/srep20676, 2016.

832 Environment Canada (EC): National Pollutants Release Inventory (NPRI) Datasets,
833 available at:
834 <https://www.ec.gc.ca/inrp-npri/default.asp?lang=en&n=0EC58C98-1> (last
835 access: May 29, 2016), 2016.

836 Evers, D. C., Han, Y.-J., Driscoll, C. T., Kamman, N. C., Goodale, M. W., Lambert,
837 K. F., Holsen, T. M., Chen, C. Y., Clair, T. A., and Butler, T.: Biological
838 mercury hotspots in the Northeastern United States and Southeastern Canada,
839 *BioScience*, 57, 29-43, 10.1641/B570107, 2007.

840 Gaffney, J. S., and Marley, N. A.: In-depth review of atmospheric mercury: sources,
841 transformations, and potential sinks, *Energy and Emission Control*
842 *Technologies*, 2, 1-21, 10.2147/EECT.S37038, 2014.

843 [Gibson, M. D., Haelssig, J., Pierce, J. R., Parrington, M., Franklin, J. E., Hopper, J. T.,](#)
844 [Li, Z., and Ward, T. J.: A comparison of four receptor models used to quantify](#)
845 [the boreal wildfire smoke contribution to surface PM_{2.5} in Halifax, Nova](#)
846 [Scotia during the BORTAS-B experiment, *Atmos. Chem. Phys.*, 15\(2\), 815–](#)
847 [827, 2015.](#)

848 Goodsite, M. E., Plane, J. M. C., and Skov, H.: A theoretical study of the oxidation of
849 Hg⁰ to HgBr₂ in the troposphere, *Environ. Sci. Technol.*, 38, 1772-1776,
850 10.1021/es034680s, 2004.

851 Gustin, M. S., Lindberg, S. E., and Weisberg, P.J.: An update on the natural sources
852 and sinks of atmospheric mercury, *Applied Geochemistry*, 23, 482-493, 2008.

853 Gustin, M. S., Amos, H. M., Huang, J., Miller, M. B., & Heidecorn, K.: Measuring
854 and modeling mercury in the atmosphere: a critical review, *Atmos. Chem.*
855 *Phys.*, 15(10), 5697-5713. 10.5194/acp-15-5697-2015, 2015.

856 Hedberg, E., Gidhagen, L., and Johansson, C.: Source contributions to PM₁₀ and
857 arsenic concentrations in Central Chile using positive matrix factorization,
858 *Atmos. Environ.*, 39, 549-561, 10.1016/j.atmosenv.2004.11.001, 2005.

859 [Henry, R. C., Lewis, C. W., Hopke, P. K., and Williamson, H. J.: Review of receptor](#)
860 [model fundamentals, *Atmos. Environ.*, 18, 1507-1515,](#)
861 [10.1016/0004-6981\(84\)90375-5 1984.](#)

862 Henry, R. C.: Multivariate receptor models, in: *Receptor modeling for air quality*
863 *management*, 1st ed., edited by: Hopke, P. K., Elsevier Science Publishers,
864 Amsterdam, 1991.

865 Holmes, C. D., Jacob, D. J., Mason, R. P., & Jaffe, D. A.: Sources and deposition of
866 reactive gaseous mercury in the marine atmosphere. *Atmos. Environ.*, 43(14),
867 2278-2285, 10.1016/j.atmosenv.2009.01.051, 2009.

868 [Hopke, P. K.: It is time to drop principal components analysis as a “receptor model”. *J.*](#)
869 [*Atmos. Chem.*, 72, 127-128, 2015.](#)

870 Hopke, P. K.: Review of receptor modeling methods for source apportionment, *J. Air.*
871 *Waste. Manage.*, 66, 237-259, 10.1080/10962247.2016.1140693, 2016.

872 Huang, J., Choi, H.-D., Hopke, P. K., and Holsen, T. M.: Ambient Hg sources in
873 Rochester, NY: results from principle components analysis (PCA) of Hg
874 monitoring network data, *Environ. Sci. Technol.*, 44, 8441-8445,
875 10.1021/es102744j, 2010.

876 Huang, S., Rahn, K. A., and Arimoto, R.: Testing and optimizing two factor-analysis
877 techniques on aerosol at Narragansett, Rhode Island, *Atmos. Environ.*, 33,
878 2169-2185, 10.1016/S1352-2310(98)00324-0, 1999.

- 879 Kaiser, H. F.: The Application of Electronic Computers to Factor Analysis, *Educ.*
880 *Psychol. Meas.*, 20, 141-151, 10.1177/001316446002000116, 1960.
- 881 Lee, E., Chan, C. K., & Paatero, P.: Application of positive matrix factorization in
882 source apportionment of particulate pollutants in Hong Kong. *Atmos. Environ.*,
883 33(19), 3201-3212, 10.1016/S1352-2310(99)00113-2, 1999.
- 884 [Liao, Y.: Analysis of Potential Sources and Processes Affecting Ambient Speciated](#)
885 [Mercury Concentrations at Kejimikujik National Park, Nova Scotia, Master's](#)
886 [thesis, University of Windsor, Windsor, Ontario, Canada. 2016.](#)
- 887 Liu, B., Keeler, G. J., Dvonch, J. T., Barres, J. A., Lynam, M. M., Marsik, F. J., and
888 Morgan, J. T.: Temporal variability of mercury speciation in urban air, *Atmos.*
889 *Environ.*, 41, 1911-1923, 10.1016/j.atmosenv.2006.10.063, 2007.
- 890 Lynam, M. M., and Keeler, G. J.: Artifacts associated with the measurement of
891 particulate mercury in an urban environment: the influence of elevated ozone
892 concentrations, *Atmos. Environ.*, 39, 3081-3088,
893 10.1016/j.atmosenv.2005.01.036, 2005.
- 894 Lynam, M. M., and Keeler, G. J.: Source-receptor relationships for atmospheric
895 mercury in urban Detroit, Michigan, *Atmos. Environ.*, 40, 3144-3155,
896 10.1016/j.atmosenv.2006.01.026, 2006.
- 897 Obrist, D., Tas, E., Peleg, M., Matveev, V., Fain, X., Asaf, D., and Luria, M.:
898 Bromine-induced oxidation of mercury in the mid-latitude atmosphere, *Nat.*
899 *Geosci.*, 4, 22-26, 10.1038/ngeo1018, 2011.
- 900 [Paatero, P., and Tapper, U.: Positive matrix factorization: A non-negative factor](#)
901 [model with optimal utilization of error estimates of data values,](#)
902 [Environmetrics, 5, 111-126, 10.1002/env.3170050203, 1994.](#)
- 903 Pakkanen, T. A.: Study of formation of coarse particle nitrate aerosol, *Atmos.*
904 *Environ.*, 30, 2475-2482, 10.1016/1352-2310(95)00492-0, 1996.
- 905 Pal, B., and Ariya, P. A.: Studies of ozone initiated reactions of gaseous mercury:
906 kinetics, product studies, and atmospheric implications, *Phys. Chem. Chem.*
907 *Phys.*, 6, 572-579, 10.1039/B311150D, 2004.
- 908 Parmar, R. S., Satsangi, G. S., Kumari, M., Lakhani, A., Srivastava, S. S., and Prakash,
909 S.: Study of size distribution of atmospheric aerosol at Agra, *Atmos. Environ.*,
910 35, 693-702, 10.1016/S1352-2310(00)00317-4, 2001.
- 911 Pavlovic, R. T., Nopmongcol, U., Kimura, Y., and Allen, D. T.: Ammonia emissions,
912 concentrations and implications for particulate matter formation in Houston,
913 TX, *Atmos. Environ.*, 40, 538-551, 10.1016/j.atmosenv.2006.04.071, 2006.
- 914 Pekey, H., Karakaş, D., & Bakoglu, M.: Source apportionment of trace metals in
915 surface waters of a polluted stream using multivariate statistical analyses.
916 *Marine Pollution Bulletin*, 49(9), 809-818, 2004.
- 917 Pirrone, N., Cinnirella, S., Feng, X., Finkelman, R.B., Friedli, H.R., Leaner, J., Mason,
918 R., (...), Telmer, K.: Global mercury emissions to the atmosphere from
919 anthropogenic and natural sources. *Atmos. Chem. Phys.*, 10, 5951-5964,
920 2010.
- 921 Pitchford, M. L., Poirot, R. L., Schichtel, B. A., and Malm, W. C.: Characterization of
922 the winter midwestern particulate nitrate bulge, *J. Air Waste Manage.*, 59,
923 1061-1069, 10.3155/1047-3289.59.9.1061, 2009.
- 924 Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C., and Sisler, J. F.: Atmospheric
925 aerosol over Alaska: 2. Elemental composition and sources, *J. Geophys.*
926 *Res-Atmos.*, 103, 19045-19057, 10.1029/98JD01212, 1998.

927 Rutter, A. P., and Schauer, J. J.: The effect of temperature on the gas-particle
928 partitioning of reactive mercury in atmospheric aerosols, *Atmos. Environ.*, 41,
929 8647-8657, 10.1016/j.atmosenv.2007.07.024, 2007.

930 Tekran Inc.: Products. Ambient Air, available at:
931 <http://www.tekran.com/products/ambient-air/overview/> (last access: May 29,
932 2016), 2010.

933 Thurston, G. D., and Spengler, J. D.: A quantitative assessment of source
934 contributions to inhalable particulate matter pollution in metropolitan Boston,
935 *Atmos. Environ.*, 19, 9-25, doi:10.1016/0004-6981(85)90132-5, 1985.

936 United Nations Environmental Programme (UNEP): Global Mercury Assessment
937 2013: Sources, Emissions, Releases and Environmental Transport. UNEP
938 Chemicals Branch, Geneva, Switzerland, available at:
939 <http://www.unep.org/PDF/PressReleases/GlobalMercuryAssessment2013.pdf>
940 (last access: May 30, 2016), 2013.

941 US Environmental Protection Agency (US EPA): Clean Air Markets: 2010 Progress
942 Report Emission, Compliance, and Market Analyses, available at:
943 <https://www.epa.gov/airmarkets/acid-rain-program-historical-reports> (last
944 access June 9, 2016), 2011

945 US Environmental Protection Agency (US EPA): EPA Positive Matrix Factorization
946 (PMF) 5.0 Fundamentals and User Guide, available at:
947 [https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user](https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf)
948 [guide.pdf](https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf) (last access: May 30, 2016), 2014a.

949 US Environmental Protection Agency (US EPA): Positive Matrix Factorization
950 Model (version 5.0), available at:
951 https://www.epa.gov/sites/production/files/2015-03/epa_pmf_5.0_setup.exe
952 (last access: June 6, 2015), 2014b.

953 Viana, M., Pandolfi, M., Minguillón, M. C., Querol, X., Alastuey, A., Monfort, E.,
954 and Celades, I.: Inter-comparison of receptor models for PM source
955 apportionment: case study in an industrial area, *Atmos. Environ.*, 42,
956 3820-3832, 10.1016/j.atmosenv.2007.12.056, 2008.

957 Wang, Y., Huang, J., Hopke, P. K., Rattigan, O. V., Chalupa, D. C., Utell, M. J., and
958 Holsen, T. M.: Effect of the shutdown of a large coal-fired power plant on
959 ambient mercury species, *Chemosphere*, 92, 360-367,
960 10.1016/j.chemosphere.2013.01.024, 2013.

961 Watson, J. G., L. -W. Antony Chen, Chow, J. C., Doraiswamy, P., and Lowenthal, D.
962 H.: Source apportionment: findings from the U.S. supersites program, *J. Air*
963 *Waste Manage.*, 58, 265-288, 10.3155/1047-3289.58.2.265, 2008.

964 Wyn, B., Kidd, K. A., Burgess, N. M., Curry, R. A., and Munkittrick, K. R.:
965 Increasing mercury in yellow perch at a hotspot in Atlantic Canada,
966 Kejimikujik National Park, *Environ. Sci. Technol.*, 44, 9176–9181,
967 10.1021/es1018114, 2010.

968 Zhang L., Vet R., Wiebe A., Mihele C., Sukloff B., Chan E., Moran M., and Iqbal S.:
969 Characterization of the size-segregated water-soluble inorganic ions at eight
970 Canadian rural sites, *Atmos. Chem. Phys.*, 8, 7133-7151, 2008.

971 Zhang L., Wang S., Wu Q., Wang F., Lin C.-J., Zhang L., Hui M., and Hao J.:
972 Mercury transformation and speciation in flue gases from anthropogenic
973 emission sources: A critical review, *Atmos. Chem. Phys.*, 16, 2417-2433,
974 2016.

975 **List of Tables**

976 Table 1. General statistics of air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise
977 noted) and meteorological parameters in 2009.

978 Table 2. General statistics of air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise
979 noted) and meteorological parameters in 2010, MDL same as in Table 1.

980 Table 3. PMF case design with different treatments of speciated Hg data.

981 Table 4. PCA input and set-up.

982 Table 5. Factor profiles (concentration >25%, between 20% and 25% in parenthesis)
983 of Case 2009 and Case 2010.

984 Table 6. PMF model performances on speciated mercury in 2009 and 2010.

985 Table 7. PCA component loadings (>0.25) of Case 09-C and Case 09-C&M.

986 Table 8. PCA component loadings (>0.25) of Case 10-C and Case 10-C&M.

987 Table 9. General statistics of speciated Hg with different data treatment options.

988 Table 10. Impact of combining or excluding GOM and PBM on PMF factor
989 contributions (>15%) to Hg compounds.

990

991 **List of Figures**

992 Figure 1. Map showing the locations of sampling site (▲), the top 19 SO₂ or NO_x point
993 sources (average of 2009 and 2010) (★), and all Hg point sources in 2009 and
994 2010 (○), in Nova Scotia, Canada.

995 Figure 12. PMF source profiles in percent of concentration, 2009.

996 Figure 23. PMF source profiles in percent of concentration, 2010.

997 Figure 34. Box plot of ~~reproduced~~ predicted to observed concentration ratios (upper
998 whisker- upper 25% of the distribution excluding outliers; interquartile range
999 box - middle 50% of the data; horizontal line in the box - median; lower
1000 whisker- lower 25% of the distribution excluding outliers; ⊕ - mean).

1001

Table 1. General statistics of **daily** air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise noted) and **meteorological** parameters in 2009.

Compound	Percent of missing values	Method detection limit (MDL)	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation	Coefficient of variability (%)
GEM (ng/m^3)	31%	0.1	0%	1.37	1.41	1.39	0.26	18.7
GOM (pg/m^3)	32%	2	78%	0.57	0.42	1.77	3.70	209
PBM (pg/m^3)	41%	2	48%	1.78	2.15	2.81	2.72	96.8
PM	20%	1	9%	2.71	2.91	3.44	2.49	72.4
O ₃	0%	4.3	0%	59.4	62.1	62.4	19.1	30.6
SO ₂	3%	0.002	0%	0.20	0.22	0.40	0.51	128
HNO ₃	3%	0.05	12%	0.13	0.12	0.19	0.22	116
Ca ²⁺	1%	0.002	0%	0.05	0.05	0.06	0.04	66.7
K ⁺	1%	0.029	17%	0.04	0.03	0.04	0.03	75.0
Na ⁺	1%	0.05	9%	0.25	0.30	0.43	0.47	109
Mg ²⁺	1%	0.0004	2%	0.04	0.04	0.06	0.06	100
Cl ⁻	1%	0.046	23%	0.19	0.23	0.46	0.64	139
NO ₃ ⁻	1%	0.06	9%	0.18	0.17	0.28	0.39	139
NH ₄ ⁺	1%	0.001	0%	0.19	0.17	0.28	0.32	114
SO ₄ ²⁻	1%	0.05	0%	0.78	0.76	1.14	1.27	111
Total ions	1%	-	-	2.13	2.05	2.76	2.23	81
Temperature (°C)	0%	-	-	-	7.31	6.64	9.28	140
Relative humidity (%)	0%	-	-	-	87.5	84.5	12.0	14
Wind speed (m/s)	0%	-	-	-	4.33	4.70	2.39	51
Precipitation (mm/day)	3%	-	-	-	0.60	4.50	10.0	222

Table 2. General statistics of **daily** air pollutant concentrations (in $\mu\text{g}/\text{m}^3$ unless otherwise noted) and **meteorological parameters** in 2010, MDL same as in Table 1.

Compound	Percent of missing values	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation	Coefficient of variability (%)
GEM (ng/m^3)	4%	0%	1.34	1.38	1.35	0.17	12.6
GOM (pg/m^3)	4%	96%	0.27	0.21	0.44	0.64	145
PBM (pg/m^3)	4%	46%	2.08	2.20	3.40	4.13	121
O ₃	1%	0%	62.2	63.4	64.5	16.6	25.7
SO ₂	19%	1%	0.10	0.13	0.23	0.31	135
HNO ₃	19%	25%	0.10	0.10	0.18	0.22	122
Ca ²⁺	19%	0%	0.04	0.04	0.07	0.13	186
K ⁺	19%	46%	0.04	0.03	0.06	0.07	117
Na ⁺	19%	16%	0.20	0.24	0.40	0.53	133
Mg ²⁺	19%	0 %	0.03	0.04	0.05	0.06	120
Cl ⁻	19%	27%	0.14	0.15	0.46	0.83	180
NO ₃ ⁻	19%	21%	0.14	0.13	0.25	0.36	144
NH ₄ ⁺	19%	0%	0.16	0.15	0.30	0.57	190
SO ₄ ²⁻	19%	0%	0.69	0.64	1.11	1.65	149
Total ions	19%	-	1.89	1.80	2.71	2.95	109
Temperature (°C)	0%	-	-	8.57	8.13	8.92	110
Relative humidity (%)	0%	-	-	86.8	84.5	12.6	15
Wind speed (m/s)	0%	-	-	3.63	4.37	3.09	71
Precipitation (mm/day)	2%	-	-	0.60	4.15	9.71	234

Table 3. PMF case design with different treatments of speciated Hg data.

Case		Input variables (m)	Treatment of missing value	Sample size	
2009	2010			2009	2010
2009 (base case)	2010 (base case)	All compounds (15)	Excluding listwise	161	290
09+Mean	10+Mean	All compounds (15)	Geometric mean imputation	365	365
09+Median	10+Median	All compounds (15)	Median imputation	365	365
09+RM	10+RM	All compounds, but combining GOM & PBM to RM (14)	Excluding listwise	161	290
09-RM	10-RM	All compounds, except GOM & PBM (13)	Excluding listwise	201	290
09ScaleRM	10ScaleRM	All compounds, GOM & PBM scaled (15)	Excluding listwise	161	290

Table 4. PCA input and set-up.

Case	Year	Input variables (m)	Sample size (n)	Required sample size (50+m)	Other settings
09-C	2009	All compounds (15)	161	65	1) Missing value: Listwise deletion 2) Components to keep: eigenvalues >1) 3) Rotation: Varimax 4) Cut-off value for major loadings: 0.25
09-C&M	2009	All compounds and meteorological parameters (19)	159	69	
10-C	2010	All compounds (15)	290	65	
10-C&M	2010	All compounds and meteorological parameters (19)	285	69	

Table 5. Factor profiles (concentration >25%, between 20% and 25% in parenthesis) of Case 2009 and Case 2010.

Compound	2009				2010			
	F1	F2	F3	F4	F1	F2	F3	F4
GEM			76				79	
GOM	31		69			37	59	
PBM		29	63				81	
PM	42		34		-	-	-	-
O ₃			72				80	
SO ₂		82				93		
HNO ₃	54	(21)	(25)		64	26		
Ca ²⁺	(19)		45	31		29	36	(21)
K ⁺	(22)		37	39	51		27	(23)
Na ⁺				86				83
Mg ²⁺				83				75
Cl ⁻				100				100
NO ₃	(25)	(23)		40		41	(23)	
NH ₄ ⁺	71				87			
SO ₄ ²⁻	64				79			
Factor	Combustion emission	Industrial sulfur	Photochemistry & re-emission of Hg	Sea salt	Combustion emission	Industrial sulfur	Photochemistry & re-emission of Hg	Sea salt

Table 6. PMF model performances on speciated mercury in 2009 and 2010.

Hg form	Case	Distribution	Number of scaled residuals greater than 3	Coefficient of determination (R ²)	Slope of regression line
GEM	09	Normal	0	0.28	0.59
	09+mean	Concentrated near zero	5	0.17	0.57
	09+median	Concentrated near zero	5	0.15	0.54
	09+RM	Normal	0	0.29	0.59
	09-RM	Normal	1	0.25	0.59
	09ScaleRM	Normal	0	0.28	0.58
	10	Normal	2	0.46	1.29
	10+mean	Normal	19	0.32	1.26
	10+median	Normal	2	0.41	1.26
	10+RM	Normal	2	0.46	1.31
	10-RM	Normal	2	0.47	1.31
10ScaleRM	Normal	1	0.44	1.19	
GOM	09	Right skewed	17	0.23	0.09
	09+mean	Concentrated near zero, right skewed	17	0.08	0.05
	09+median	Concentrated near zero, right skewed	19	0.09	0.05
	09+RM	-	-	-	-
	09-RM	-	-	-	-
	09ScaleRM	Right skewed	26	0.33	0.18
	10	Narrower	0	0.31	0.29
	10+mean	Narrower	0	0.23	0.22
	10+median	Narrower	0	0.28	0.28
	10+RM	-	-	-	-
	10-RM	-	-	-	-
10ScaleRM	Narrower	0	0.42	0.33	
PBM	09	Normal	5	0.57	0.39
	09+mean	Right skewed	6	0.33	0.32
	09+median	Right skewed	6	0.34	0.34
	09+RM	Right skewed (RM)	8 (RM)	0.48(RM)	0.31(RM)
	09-RM	-	-	-	-
	09ScaleRM	Left skewed	2	0.59	0.48
	10	Right skewed	14	0.13	0.09
	10+mean	Right skewed	28	0.15	0.09
	10+median	Right skewed	29	0.16	0.08
	10+RM	Right skewed (RM)	5	0.19	0.15
	10-RM	-	-	-	-
10ScaleRM	Normal	18	0.25	0.24	

Table 7. PCA component loadings (>0.25) of Case 09-C and Case 09-C&M.

Variable	Case 09-C				Case 09-C&M				
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC5
GEM			0.86	0.27				0.80	
GOM			0.26	0.84			0.64	0.41	-0.29
PBM	0.63		0.50	-0.33	0.59		-0.47	0.34	
PM	0.80				0.81				
O ₃	0.50		0.70		0.47			0.72	-0.27
SO ₂	0.88				0.86				
HNO ₃	0.86			0.34	0.88				
Ca ²⁺	0.59	0.39		0.45	0.60	0.38	0.33		
K ⁺	0.29	0.70		0.33	0.36	0.66	0.39		
Na ⁺		0.97				0.96			
Mg ²⁺		0.95			0.28	0.95			
Cl ⁻		0.97				0.98			
NO ₃ ⁻	0.73	0.48			0.76	0.45			
NH ₄ ⁺	0.92				0.94				
SO ₄ ²⁻	0.86				0.88				
Temperature	-	-	-	-			0.94		
Relative humidity	-	-	-	-	-0.26				0.79
Wind speed	-	-	-	-		0.32		0.52	0.49
Precipitation	-	-	-	-					0.79
Component	Combustion/industrial emission	Sea salt	Photochemical production of GOM	Gas-particle partition of Hg	Combustion/industrial emission	Sea salt	Gas-particle partition of Hg	Photochemical production of GOM	Hg wet deposition
Variance explained	37%	25%	11%	9%	30%	20%	10%	10%	9%

Table 8. PCA component loadings (>0.25) of Case 10-C and Case 10-C&M.

Variable	Case 10-C				Case 10-C&M				
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC5
GEM			0.79				0.87		
GOM			0.71	0.33			0.51	-0.51	0.38
PBM			0.48				0.29	-0.62	
O ₃			0.91				0.87		
SO ₂				0.89					0.84
HNO ₃	0.34			0.83	0.33				0.82
Ca ²⁺	0.89				0.89				
K ⁺	0.77				0.77				
Na ⁺		0.99					0.99		
Mg ²⁺	0.34	0.93			0.34		0.92		
Cl ⁻		0.98					0.97		
NO ₃ ⁻	0.79				0.80				
NH ₄ ⁺	0.94				0.94				
SO ₄ ²⁻	0.90			0.26	0.89				0.26
Temperature	-	-	-	-	0.27		-0.52		0.27
Relative humidity	-	-	-	-				0.74	-0.33
Wind speed	-	-	-	-		0.26	0.52	0.57	
Precipitation	-	-	-	-				0.76	
Component	Combustion emission	Sea salt	Photochemical production of GOM	Industrial source	Combustion emission	Sea salt	Photochemical production of GOM	Hg wet deposition	Industrial source
Variance explained	28%	21%	16%	13%	22%	17%	14%	12%	10%

Table 9. General statistics of speciated Hg with different data treatment options.

a) 2009

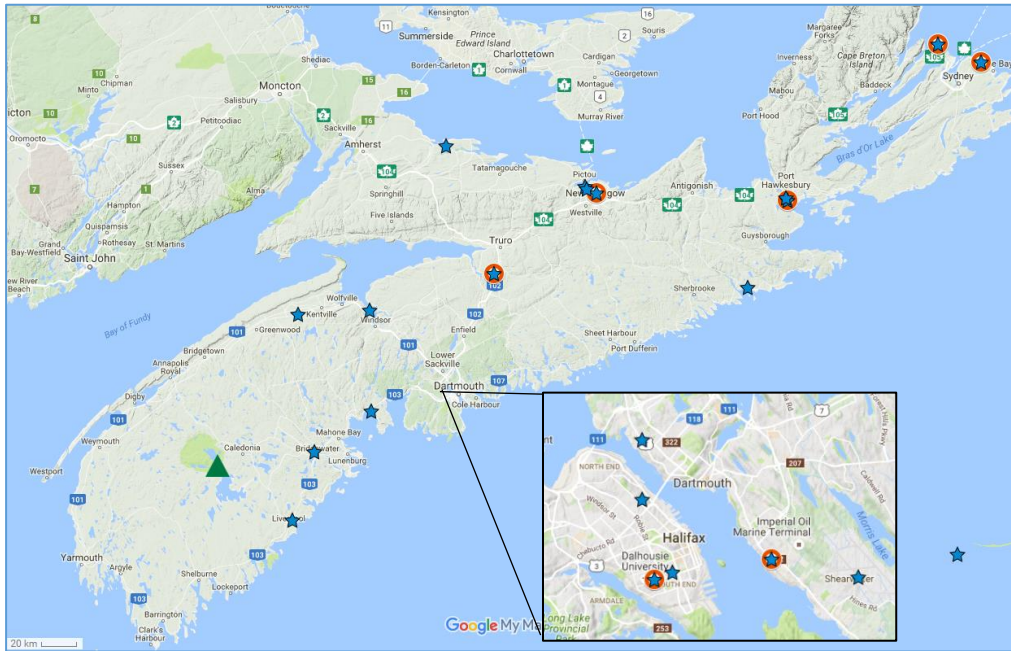
Hg form	Case	Percent of missing values	MDL	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m ³)	09	31%		0%	1.37	1.41	1.39	0.28
	09+mean	0%	0.1	0%	1.37	1.37	1.38	0.22
	09+median	0%		0%	1.38	1.41	1.39	0.22
GOM (pg/m ³)	09	32%		73%	0.57	0.42	1.77	3.98
	09+mean	0%		86%	0.57	0.57	1.39	3.11
	09+median	0%	2	86%	0.51	0.42	1.34	3.12
	09+RM	-		-	-	-	-	-
	09Scale RM	32%		16%	3.91	3.35	5.02	5.49
PBM (pg/m ³)	09	41%	2	37%	1.79	2.15	2.81	2.71
	09+mean	0%	2	70%	1.79	1.79	2.39	2.14
	09+median	0%	2	28%	1.93	2.15	2.53	2.11
	09+RM	42%	4 (RM)	52%	2.73	3.02	4.69	5.56
	09Scale RM	41%	2	4%	5.52	6.05	6.19	3.15

b) 2010, MDL same as in a)

Hg form	Case	Percent of missing values	Percent of values <MDL	Geometric Mean	Median	Mean	Standard deviation
GEM (ng/m ³)	10	4%	0%	1.33	1.37	1.34	0.17
	10+mean	0%	0%	1.34	1.37	1.35	0.16
	10+median	0%	0%	1.34	1.38	1.35	0.17
	10+RM	4%	0%	1.33	1.37	1.34	0.17
	10ScaleRM	4%	0%	1.33	1.38	1.34	0.17
GOM (pg/m ³)	10	4%	96%	0.29	0.26	0.49	0.69
	10+mean	0%	96%	0.27	0.24	0.43	0.63
	10+median	0%	96%	0.27	0.21	0.43	0.63
	10+RM	-	-	-	-	-	-
	10ScaleRM	4%	67%	1.15	1.12	1.40	0.86
PBM (pg/m ³)	10	4%	51%	1.79	1.92	2.59	2.67
	10+mean	0%	44%	2.08	2.12	3.35	4.04
	10+median	0%	44%	2.08	2.20	3.35	4.04
	10+RM	4%	75%	2.16	2.31	3.08	2.95
	10ScaleRM	4%	1%	6.15	6.38	6.75	3.01

Table 10. Impact of combining or excluding GOM and PBM on PMF factor contributions (>15%) to Hg compounds.

Case	Combustion emission	Industrial sulfur	Photochemistry and re-emission	Sea salt
Case 2009	GOM	PBM	GEM, GOM, and PBM	
Case 09+RM		RM	GEM and RM	
Case 09-RM			GEM	
Case 2010		GOM	GEM, GOM, and PBM	
Case 10+RM			GEM and RM	
Case 10-RM			GEM	



1002
 1003
 1004
 1005

Figure 1. Map showing the locations of sampling site (▲), the top 19 SO₂ or NO_x point sources (average of 2009 and 2010) (★), and all Hg point sources in 2009 and 2010 (○), in Nova Scotia, Canada.

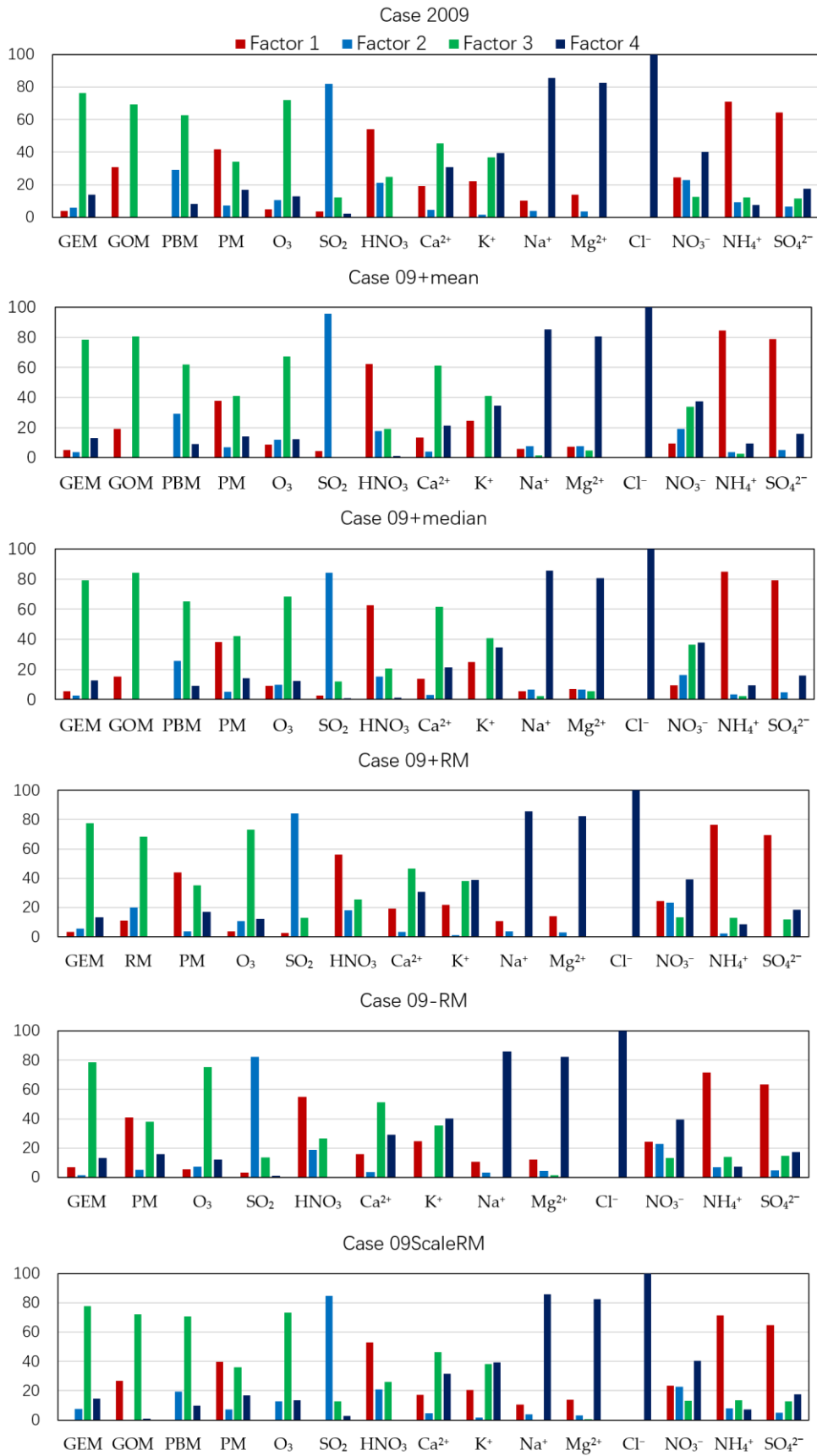


Figure 42. PMF source profiles in percent of concentration, 2009. (subscription/superscription updated)

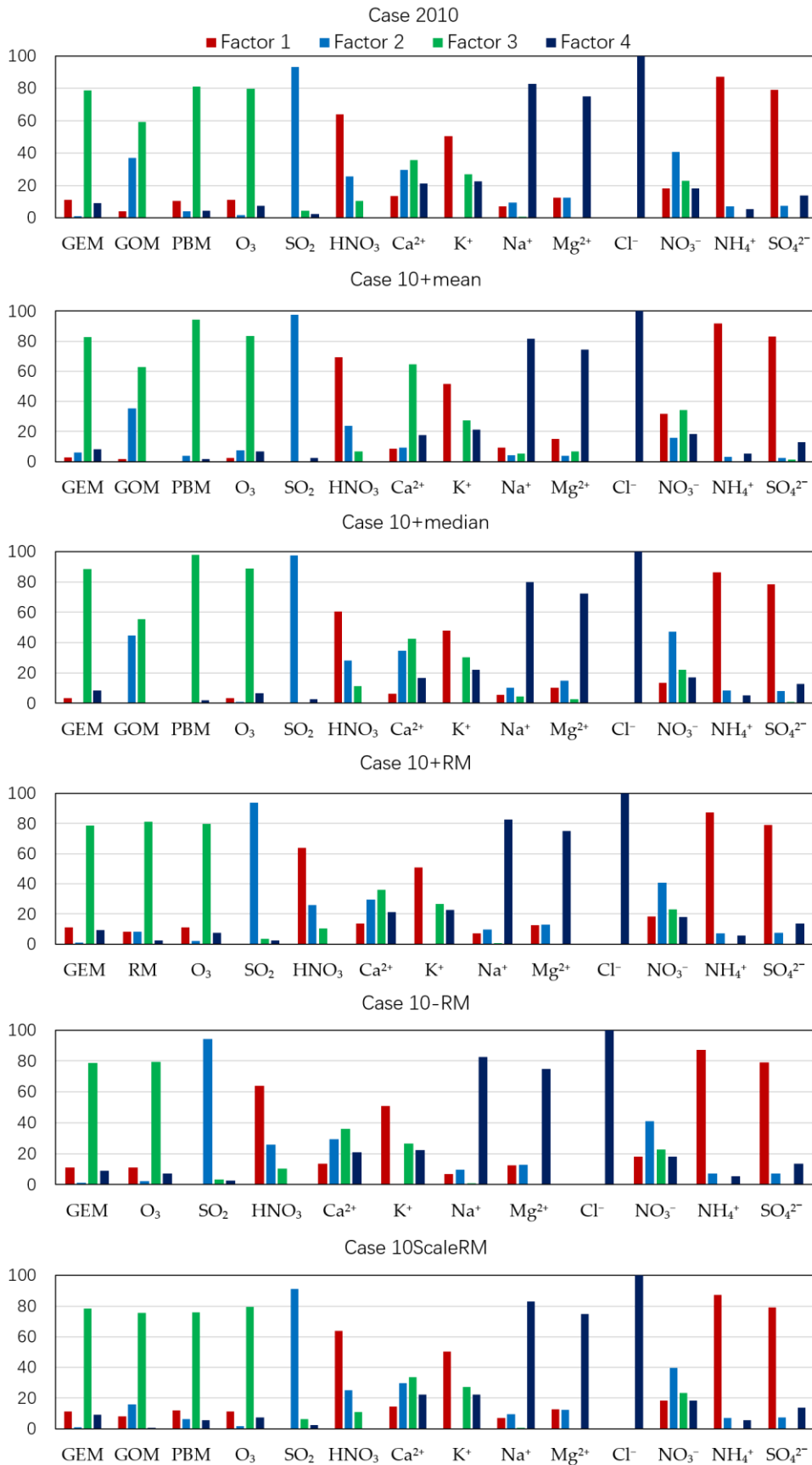


Figure 23. PMF source profiles in percent of concentration, 2010. (subscription/superscription updated)

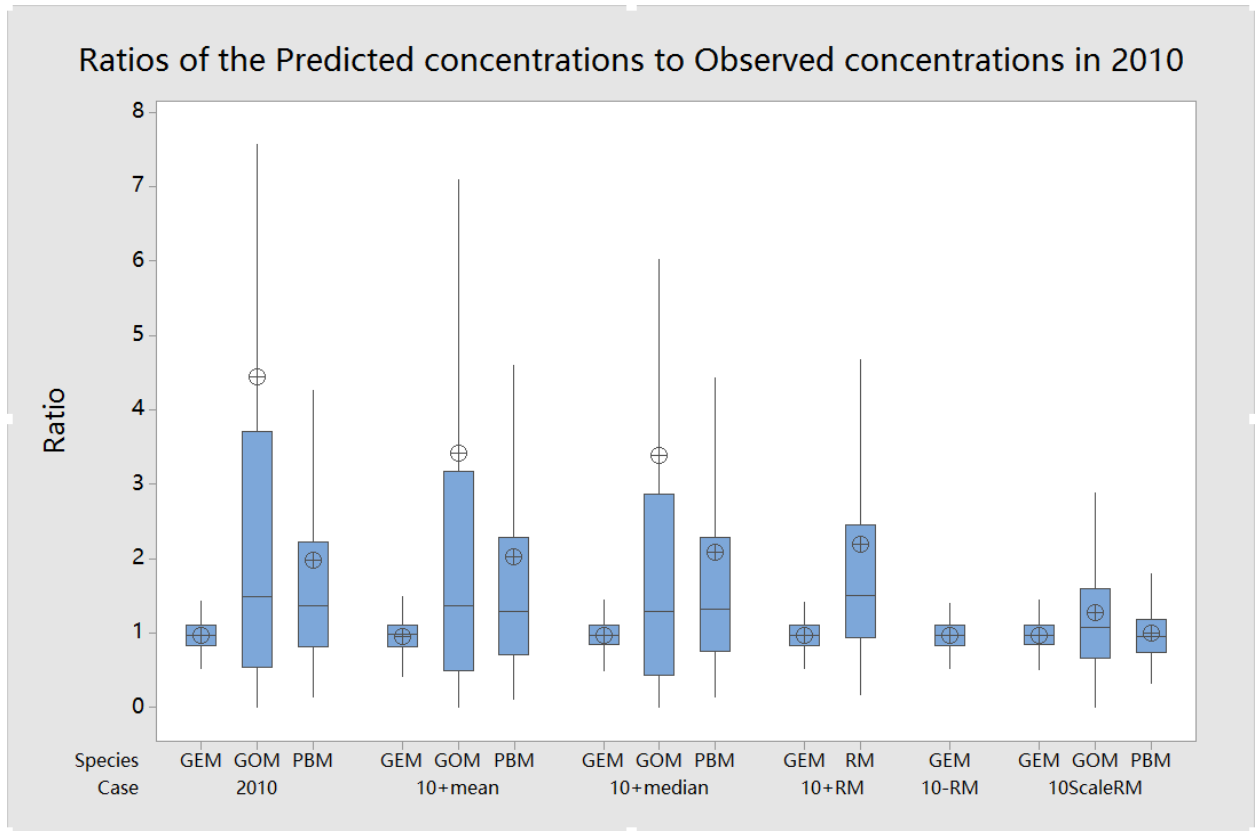


Figure 43. Box plot of ~~model-reproduced~~ predicted to observed concentration ratios (upper whisker-upper 25% of the distribution excluding outliers; interquartile range box - middle 50% of the data; horizontal line in the box - median; lower whisker- lower 25% of the distribution excluding outliers;⊕ - mean).

Supplementary Information

Potential sources and processes affecting speciated atmospheric mercury at Kejimikujik National Park, Canada: comparison of receptor models and data treatment methods

Xiaohong Xu, Yanyin Liao, Irene Cheng, Leiming Zhang

Contents

Section 1. Selection of the number of PMF factors

Section 2. Stability of PMF model runs

Table S1. Point Source eEmissions of Hg and other pollutants reported in NPRI in the province of Nova Scotia within 150 Km of the sampling site (Data source: Environmental Canada, 2016). Bold facilities are shown in Figure 1.

Table S2. Coefficients of cross-correlation among all variables in 2009 (bold numbers are significant at $p < 0.05$).

Table S3. Coefficients of cross-correlation among all variables in 2010 (bold numbers are significant at $p < 0.05$).

Table S4. PMF factor contributions to speciated Hg and ratios of predicted to observed annual Hg concentrations in 2009.

Table S5. PMF factor contributions to speciated Hg and ratios of predicted to observed annual Hg concentrations in 2010.

Table S6. Pearson correlation coefficients between Hg forms and other compounds in Case 2009, Case 09+mean, and Case 09+median (bold numbers are significant at < 0.05).

Table S7. Pearson correlation coefficients between Hg forms and other compounds in Case 2010, Case 10+mean, and Case 10+median (bold numbers are significant at < 0.05).

Figure S51. Obs/Pred scatter plot in 2009. a) Case 2009, b) Case 09+mean, c) Case 09+median, d) Case 09+RM, e) Case 09-RM, and f) Case 09ScaleRM, observed GOM and PBM have been scaled.

Figure S62. Obs/Pred scatter plot in 2010. a) Case 2010, b) Case 10+mean, c) Case 10+median, d) Case 10+RM, e) Case 10-RM, and f) Case 10ScaleRM, observed GOM and PBM have been scaled.

Figure S73. Obs/Pred time series in 2009. a) Case 2009, b) Case 09+mean, c) Case

09+median, d) Case 09+RM, e) Case 09-RM, and f) Case 09ScaleRM, observed GOM and PBM have been scaled.

Figure S84. Obs/Pred time series in 2010. a) Case 2010, b) Case 10+mean, c) Case 10+median, d) Case 10+RM, e) Case 10-RM, and f) Case 10ScaleRM, observed GOM and PBM have been scaled.

Section 1. Selection of the member of PMF factors

The number of PMF factors needs to be chosen according to the understanding of the sources impacting the samples utilized. When the background information is not enough to determine the number of 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 can be used to identify the range of the number of factors. IM and IS will show a drastic drop when the number of factors increase up to a critical value (Lee et al., 1999). The optimal number of factors should be no less than the critical value. The trend of dQ also provides useful information on deciding the number of factors. 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 should be conducted. The model performance and the interpretability of factors in each run should be evaluated. The optimal PMF solution should be a compromise of those indexes and the interpretability of the factor profiles and their comparability with those from the literature (Belis et al, 2015a, 2015b; Cesari et al., 2016).

In this study, 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 S1). In the line plot of $Q(\text{Robust})$ and $Q(\text{true})$ vs. the number of factors (Figure S2), no significant decreases were found when the number of factors is larger than 5 in 2009. Therefore, the PMF was run using the number of factors from 3 to 5 in 2009. In 2010, the decrease of IS value was gradual while the IM value experienced a drastic drop when the number of factors increased to 3 (Figure S3). The trend of the $Q(\text{Robust})$ and $Q(\text{True})$ in 2010 is similar to 2009 (Figure S4). Therefore, the PMF runs with the number of factors from 3 to 5 were also conducted in 2010. The number of the factors selected (4) is a compromise of the trends of these indexes and the physical meanings of the factors obtained following Cesari et al. (2016). A detailed comparison of the physical meanings of solutions with different number of factors can be found in Liao (2016).

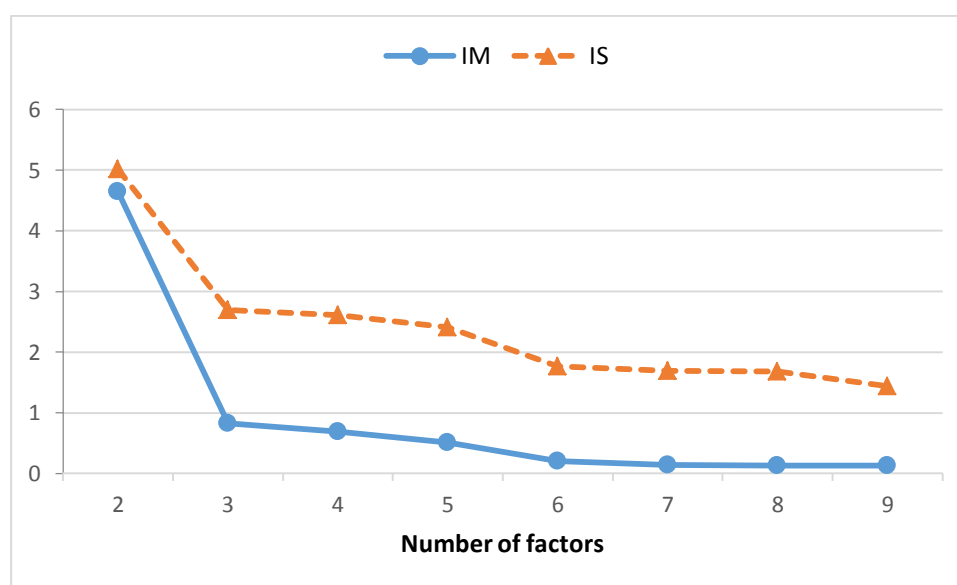


Figure S1. IM and IS vs number of PMF factors in 2009.

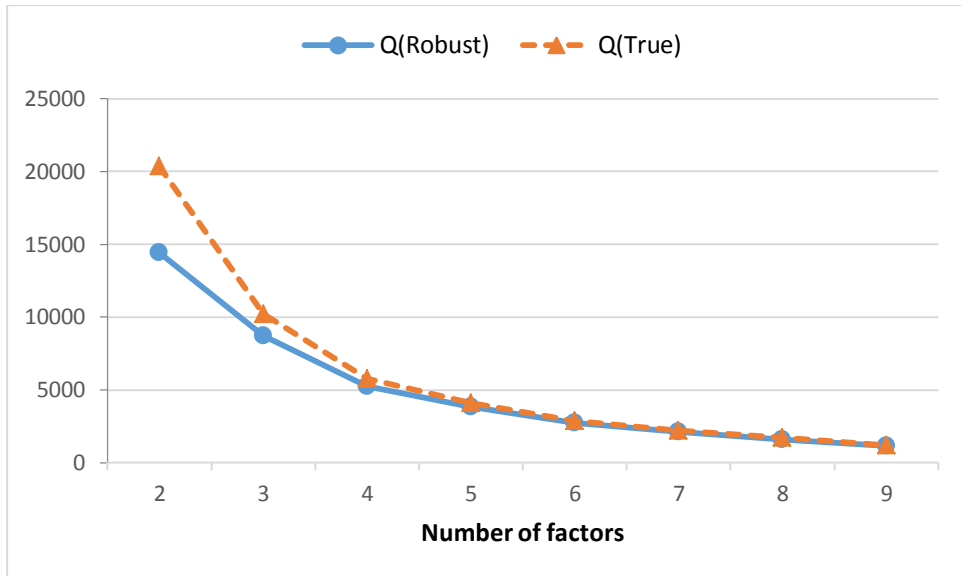


Figure S2. Q(Robust) and Q(true) vs number of PMF factors in 2009.

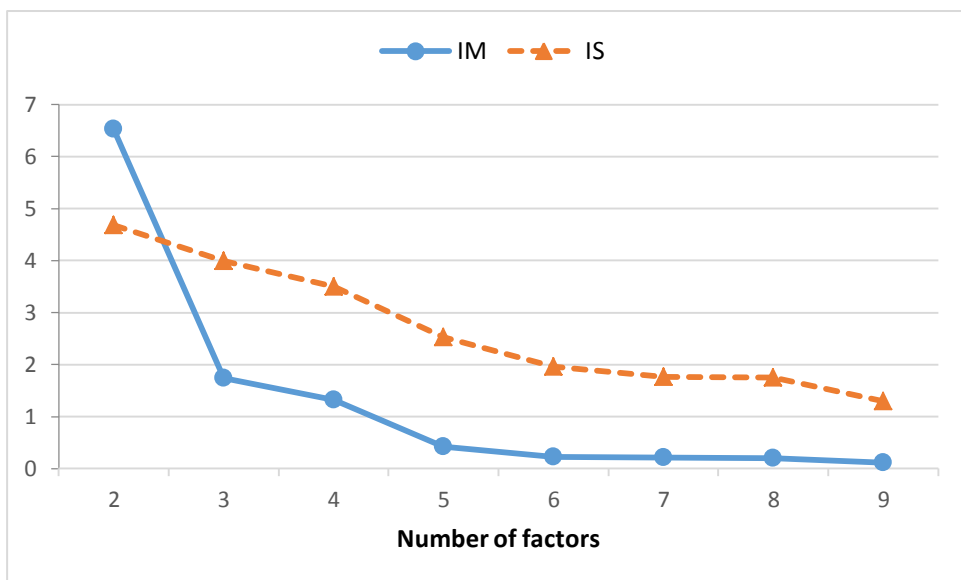


Figure S3. IM and IS vs number of PMF factors in 2010.

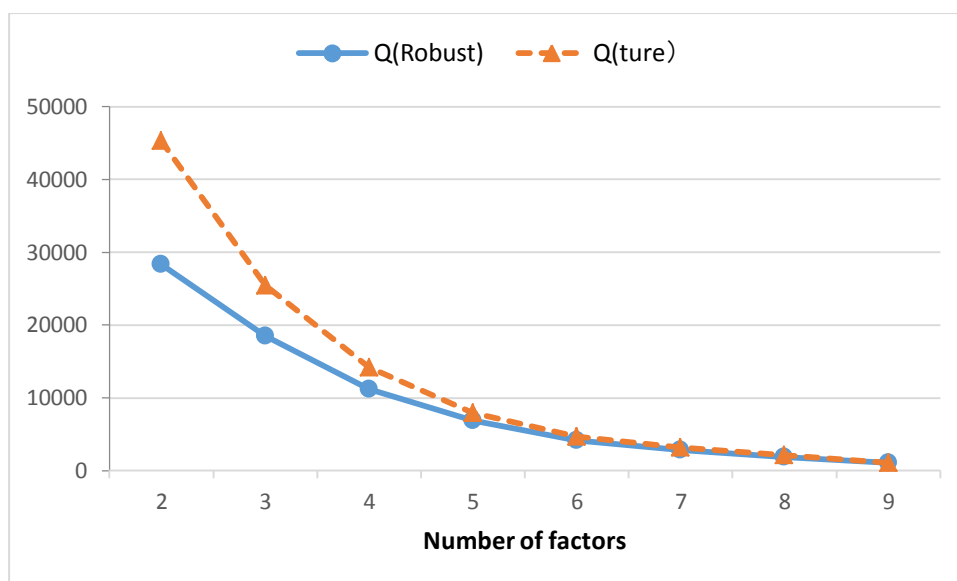


Figure S4. Q(Robust) and Q(true) vs number of PMF factors in 2010.

References:

- Belis, C. A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B. R., Gerboles, M., and Hopke, P. K.: A new methodology to assess the performance and uncertainty of source apportionment models in intercomparison exercises, *Atmos. Environ.*, 119, 35-44, 10.1016/j.atmosenv.2015.08.002, 2015a.
- Belis, C. A., Karagulian, F., Amato, F., Almeida, M., Artaxo, P., Beddows, D. C. S., Bernardoni, V., Bove, M. C., Carbone, S., Cesari, D., Contini, D., Cuccia, E., Diapouli, E., Eleftheriadis, K., Favez, O., El-Haddad, I., Harrison, R. M., Hellebust, S., Hovorka, J., Jang, E., Jorquera, H., Kammermeier, T., Karl, M., Lucarelli, F., Mooibroek, D., Nava, S., Nøjgaard, J. K., Paatero, P., Pandolfi, M., Perrone, M. G., Petit, J. E., Pietrodangelo, A., Pokorná, P., Prati, P., Prevot, A. S. H., Quass, U., Querol, X., Saraga, D., Sciare, J., Sfetsos, A., Valli, G., Vecchi, R., Vestenius, M., Yubero, E., and Hopke, P. K.: A New Methodology to Assess the Performance and Uncertainty of Source Apportionment Models II: The Results of Two European Intercomparison Exercises, *Atmos. Environ.*, 123, 240-250, 10.1016/j.atmosenv.2015.10.068, 2015b.
- Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, *Sci. Total Environ.*, 518-519, 626-635, dx.doi.org/10.1016/j.scitotenv.2015.01.022, 2015.
- Cesari, D., Amato, F., Pandolfi, M., Alastuey, A., Querol, X., and Contini, D.: An inter-comparison of PM10 source apportionment using PCA and PMF receptor models in three European sites, *Environment Science and Pollution Research*, 23, 15133-15148, 10.1007/s11356-016-6599-z, 2016.
- Hopke, P. K.: A guide to positive matrix factorization, available at: <https://www3.epa.gov/ttnamti1/files/ambient/pm25/workshop/laymen.pdf> (last access: May 30, 2016), 2000.

- Lee, E., Chan, C. K., & Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmos. Environ.*, 33(19), 3201-3212, 10.1016/S1352-2310(99)00113-2, 1999.
- Liao, Y.: Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimikujik National Park, Nova Scotia, Master's thesis, University of Windsor, Windsor, Ontario, Canada, 2016.

Section 2. Stability of PMF model runs

As pointed out in the PMF User's Guide (US EPA, 2014), when the Q (robust) values over several runs are highly variable, the stability of the result is poor. In this study, the differences of the Q (Robust) value between different runs were all smaller than 5 indicating that the results were quite stable in 2009 and 2010. More information can be found in Liao (2016).

References:

- Liao, Y.: Analysis of potential sources and processes affecting ambient speciated mercury concentrations at Kejimikujik National Park, Nova Scotia, Master's thesis, University of Windsor, Windsor, Ontario, Canada, 2016.
- US Environmental Protection Agency (US EPA): EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide, available at: https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf (last access: May 30, 2016), 2014.

Table S1. Point source emissions of Hg and other pollutants reported in NPRI within Nova Scotia (EC, 2016). Bold facilities are shown in Figure 1. (updated)

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO ₂ (Tonnes)		NO ₂ (Tonnes)		NH ₃ (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Brooklyn Power	Brooklyn (44.05°N, 64.70°W)	50 Km southeast	0	0	9.9	26	309	259	0	0
Michelin North America (Canada)- Bridgewater Plant	Bridgewater (44.39°N, 64.54°W)	53 Km east	0	0	195	184	68	63	0	0
High Liner Foods Inc.	Lunenburg (44.37°N, 64.30°W)	72 Km east	0	0	27	27	0	0	0	0
Department of National Defence – 14 Wing Greenwood	Greenwood (44.98°N, 64.91°W)	75 Km north	0	0	55	68	19	18	0	0
Louisiana Pacific Canada Ltd.	East River (44.58°N, 64.16°W)	88 Km northeast	0	0	122	102	100	99	0	0
Maple Leaf Foods – Larsen Packers Limited	Berwick (45.05°N, 64.75°W)	89 Km northeast	0	0	51	38	0	0	0	0
Michelin North America (Canada) - Waterville Plant	Waterville (45.05°N, 64.65°W)	92 Km northeast	0	0	162	182	57	62	0	0
Acadia University – Acadia Campus	Wolfville (45.08°N, 64.37°W)	108 Km northeast	0	0	77	73	27	26	0	0
CKF. Inc.	Hantsport (45.06°N, 64.17°W)	116 Km northeast	0	0	66	57	21	72	0	0
Minas Basin Pulp and Power	Hantsport (45.07°N, 64.17°W)	116 Km northeast	0	0	225	260	66	76	0	0
Mount Saint Vincent University	Halifax (44.67°N, 63.65°W)	129 Km northeast	0	0	27	13	7.2	3.9	0	0
Department of National Defence – Canadian Forces Ammunition Depot	Bedford (44.71°N, 63.63°W)	131 Km northeast	0	0	56	50	0	0	0	0

Table S1 – Continued 1

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO ₂ (Tonnes)		NO ₂ (Tonnes)		NH ₃ (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Department of National Defence - Windsor Park	Halifax (44.66°N, 63.61°W)	132 Km northeast	0	0	59	44	36	30	0	0
Department of National Defence – Stadacona/Dockyard	Halifax (44.66°N, 63.58°W)	133 Km northeast	0	0	211	177	58	51	0	0
Capital Health – Camp Hill Site Heating Plant	Halifax (44.64°N, 63.59°W)	133 Km northeast	0	0	15	12	14	20	0	0
Dalhousie University	Halifax (44.64°N, 63.59°W)	133 Km northeast	0.18	0.15	253	260	89	72	0	0
Saint Mary’s University	Halifax (44.63°N, 63.58°W)	133 Km northeast	0	0	1.2	0	3	0	0	0
Oland Brewery	Halifax (44.66°N, 63.60°W)	133 Km Northeast	0	0	31	0	0	0	0	0
Nova Scotia Power – Tufts Cove Generating Station	Dartmouth (44.67°N, 63.60°W)	134 Km northeast	0	0	2,205	2,205	3,054	3,054	0	0
Capital Health-Victoria General Hospital Central Heating Plant	Halifax (44.64°N, 63.58°W)	134 Km northeast	0	0	215	7.6	60	19	0	0
Maritime Paper Products Ltd.	Dartmouth (44.70°N, 63.60°W)	134 Km northeast	0	0	7.2	0.868	3.1	2.1	0	0
Nova Scotia Power –Burnside Combustion Turbines	Dartmouth (44.71°N, 63.61°W)	134 Km northeast	0	0	0	0	60	40	0	0
Capital Health – Nova Scotia Hospital Central Heating Plant	Dartmouth (44.65°N, 63.55°W)	136 Km northeast	0	0	3.3	1.1	9.3	8.7	0	0
Imperial Oil – Dartmouth Refinery	Dartmouth (44.64°N, 63.54°W)	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 (44.63°N, 63.51°W)	138 Km northeast	0	0	150	127	43	38	0	0
Martells Contracting	Elmsdale (44.96°N, 63.48°W)	154 Km northeast	0	0	28	17	4.5	2.8	0	0

Table S1 – Continued 2

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO ₂ (Tonnes)		NO ₂ (Tonnes)		NH ₃ (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
The Shaw Group Ltd.	Hardwoodlands (45.07°N, 63.52°W)	160 Km northeast	0	0	0	0	27	19	0	0
Lafarge Canada Inc. – Brookfield Cement Plant	Brookfield (45.24°N, 63.33°W)	180 Km northeast	5	5.9	562	667	498	591	0	0
Polycello	Amherst (45.82°N, 64.23°W)	183 Km northeast	0	0	0.003	0.002	0.462	0.335	0	0
Enligna Canada Inc.	Middle Musquodoboit (45.13°N, 62.95°W)	188 Km northeast	0	0	2.8	2.9	25	26	0	0
Oxford Frozen Foods	Oxford (45.73°N, 63.85°W)	188 Km northeast	0	0	66	59	0	0	0.9	0
Municipality of the county of Colchester – Wastewater Treatment Facility	Truro (45.37°N, 63.34°W)	188 Km northeast	0	0	0	0	0	0	2	0.08
Crossley Carpet Mills Limited	Truro (45.35°N, 63.29°W)	189 Km northeast	0	0	40	32	12	11	0	0
Rothsay	Truro (45.36°N, 63.31°W)	189 Km northeast	0	0	77	60	0	0	0	0
Stanfield's Ltd.	Truro (45.37°N, 63.28°W)	191 Km northeast	0	0	21	21	0	0	0	0
Stella-Jones Inc.	Truro (45.38°N, 63.27°W)	192 Km northeast	0	0	12	19	2.9	4.2	0	0
The Canadian Salt Company Limited – Pugwash Mine and Refinery	Pugwash (45.84°N, 63.66°W)	209 Km northeast	0	0	168	153	32	31	0	0
Michelin North America (Canada) – Pictou County Plant	New Glasgow (45.62°N, 62.74°W)	245 Km northeast	0	0	209	229	72	78	0	0
Maritime Steel and Foundries Limited	New Glasgow (45.58°N, 62.64°W)	245 Km northeast	0	0	0.25	0	0.875	0	0	0

Table S1 – Continued 3

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO ₂ (Tonnes)		NO ₂ (Tonnes)		NH ₃ (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Nova Scotia Power – Trenton Generating Station	Trenton (45.62°N, 62.64°W)	248 Km northeast	33	19	30,429	19,257	5,126	5,577	0	0
Nova Forge Corporation	Trenton (45.62°N, 62.64°W)	248 Km northeast	0	0	3.1	0	0	0	0	0
Northern Pulp Nova Scotia Corporation	New Glasgow (45.65°N, 62.72°W)	266 Km northeast	0	0	246	89	688	676	42	46
St. Francis Xavier University	Antigonish (45.62°N, 61.99°W)	291 Km northwest	0	0	41	36	25	17	0	0
Exxonmobil Canada Properties – Goldboro Gas Plant	Goldboro (45.17°N, 61.61°W)	300 Km northeast	0	0	0	0	521	415	0	0
Nova Scotia Power – Point Tupper Generating Station	Port Hawkesbury (45.58°N, 61.35°W)	335 Km northeast	12	9.5	9,394	5,721	1,952	1,952	0	0
Newpage Port Hawkesbury Corp.	Port Hawkesbury (45.60°N, 61.36°W)	355 Km northeast	0	0	294	85	404	306	0.23	0.23
Exxonmobil Canada Properties – Point Tupper Fractionation Plant	Port Hawkesbury (45.58°N, 61.34°W)	335 Km northeast	0	0	0	0	48	23	0	0
Exxonmobil Canada Properties – Thebaud Platform	Offshore (43.01°N, 59.98°W)	402 Km east	0	0	0	0	135	126	0	0
Exxonmobil Canada Properties – North Triumph Platform	North Triumph Platform (43.01°N, 58.98°W)	433 Km east	0	0	0	0	26	29	0	0
Nova Scotia Power – Point Aconi Generating Station	Point Aconi (46.32°N, 60.30°W)	442 Km northeast	2.7	2.8	3,627	3,365	1,759	1,747	0	0
Exxonmobil Canada Properties – Venture Platform	Venture Platform (44.06°N, 59.58°W)	450 Km east	0	0	18	0	54	51	0	0

Table S1 – Continued 4

Facility	Location (lat, long)	Distance to KEJ/direction	Hg (Kg)		SO ₂ (Tonnes)		NO ₂ (Tonnes)		NH ₃ (Tonnes)	
			2009	2010	2009	2010	2009	2010	2009	2010
Cape Breton University	Sydney (46.17°N, 60.09°W)	450 Km northeast	0	0	60	57	0	0	0	0
Nova Scotia Power – Lingan Generating Station	Lingan (46.23°N, 60.04°W)	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

Table S2. Coefficients of cross-correlation among all variables in 2009 (bold numbers are significant at $p < 0.05$).

	GOM	PBM	PM	O ₃	SO ₂	HNO ₃	Ca ²⁺	K ⁺	Na ⁺	Mg ²⁺	Cl ⁻	NO ₃ ⁻	NH ₄ ⁺	SO ₄ ²⁻	Temperature	Relative humidity	Wind speed	Precipitation
GEM	0.37	0.28	0.15	0.48	0.11	0.18	0.13	0.01	0.06	0.07	-0.01	0.14	0.18	0.24	0.03	0.06	0.24	0.17
GOM		0.10	0.31	0.27	0.21	0.45	0.39	0.17	-0.09	-0.02	-0.18	0.17	0.28	0.27	0.46	-0.38	-0.01	-0.09
PBM			0.47	0.56	0.63	0.42	0.28	0.11	0.20	0.23	0.06	0.50	0.53	0.54	-0.43	-0.38	-0.09	-0.15
PM				0.52	0.64	0.70	0.68	0.43	0.37	0.44	0.19	0.66	0.80	0.79	0.17	-0.36	-0.05	-0.18
O ₃					0.49	0.52	0.51	0.05	0.21	0.26	0.09	0.36	0.50	0.58	-0.17	-0.49	0.20	-0.09
SO ₂						0.80	0.52	0.29	0.29	0.34	0.09	0.71	0.74	0.70	-0.21	-0.31	-0.10	-0.19
HNO ₃							0.67	0.33	0.16	0.23	-0.05	0.61	0.80	0.77	0.16	-0.37	-0.13	-0.16
Ca ²⁺								0.47	0.44	0.53	0.31	0.59	0.57	0.58	0.23	-0.37	0.05	-0.09
K ⁺									0.64	0.66	0.57	0.53	0.34	0.41	0.37	-0.04	0.14	0.03
Na ⁺										0.99	0.96	0.61	0.24	0.39	-0.10	0.00	0.25	-0.10
Mg ²⁺											0.93	0.64	0.30	0.45	-0.06	-0.06	0.25	-0.11
Cl ⁻												0.38	0.02	0.18	-0.12	0.07	0.29	-0.07
NO ₃ ⁻													0.72	0.68	-0.02	-0.17	-0.01	-0.13
NH ₄ ⁺														0.94	0.12	-0.31	-0.09	-0.13
SO ₄ ²⁻															0.08	-0.31	0.01	-0.13
Temperature																-0.11	0.11	0.09
Relative humidity																	0.26	0.39
Wind speed																		0.39

Table S3. Coefficients of cross-correlation among all variables in 2010 (bold numbers are significant at $p < 0.05$).

	GOM	PBM	O ₃	SO ₂	HNO ₃	Ca ²⁺	K ⁺	Na ⁺	Mg ²⁺	Cl ⁻	NO ₃ ⁻	NH ₄ ⁺	SO ₄ ²⁻	Temperature	Relative humidity	Wind speed	Precipitation	
GEM	0.31	0.11	0.70	0.02	-0.11	-0.01	-0.13	0.09	0.08	0.10	-0.01	-0.10	-0.11	-0.48	-0.02	0.38	0.18	
GOM		0.29	0.55	0.30	0.24	0.07	0.07	-0.03	-0.01	-0.10	0.18	0.06	0.06	-0.04	-0.66	-0.06	-0.18	
PBM			0.32	0.06	0.04	-0.02	-0.03	-0.07	-0.07	-0.08	0.03	0.01	0.00	-0.16	-0.39	-0.15	-0.22	
O ₃				0.18	0.11	0.02	-0.07	0.05	0.05	0.01	0.06	0.00	0.02	-0.29	-0.39	0.31	0.01	
SO ₂					0.63	0.13	0.16	0.05	0.09	-0.04	0.25	0.26	0.31	-0.03	-0.31	-0.10	-0.13	
HNO ₃						0.25	0.34	-0.11	0.00	-0.24	0.28	0.48	0.53	0.33	-0.25	-0.23	-0.13	
Ca ²⁺							0.57	0.01	0.38	0.00	0.70	0.78	0.71	0.19	-0.10	-0.05	-0.02	
K ⁺								0.09	0.33	0.06	0.55	0.68	0.67	0.32	-0.16	-0.08	-0.07	
Na ⁺									0.92	0.96	0.21	-0.09	0.02	-0.16	0.08	0.28	-0.01	
Mg ²⁺										0.89	0.45	0.21	0.28	-0.08	0.04	0.24	-0.01	
Cl ⁻											0.08	-0.13	-0.06	-0.20	0.15	0.32	0.03	
NO ₃ ⁻												0.68	0.64	0.10	-0.23	-0.05	-0.11	
NH ₄ ⁺													0.97	0.28	-0.15	-0.14	-0.09	
SO ₄ ²⁻														0.29	-0.15	-0.13	-0.10	
Temperature															-0.10	-0.19	0.03	
Relative humidity																	0.24	0.41
Wind speed																		0.47

Table S4. PMF factor contributions to speciated Hg and ratios of predicted to observed annual Hg concentrations in 2009.

a) Case 2009

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	20	56	97	97
	Average	4	6	77	14
	Median	2	3	83	9
Ratio of predicted to observed annual mean:					0.97
GOM (%)	Min	0	0	0	0
	Max	100	0	100	0
	Average	25	0	70	0
	Median	19	0	78	0
Ratio of predicted to observed annual mean:					0.86
PBM (%)	Min	0	0	0	0
	Max	0	93	100	92
	Average	0	21	69	9
	Median	0	14	74	5
Ratio of predicted to observed annual mean:					1.03

b) Case 09+mean

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	79	34	98	99
	Average	5	4	78	13
	Median	3	2	83	7
Ratio of predicted to observed annual mean:					0.94
GOM (%)	Min	0	0	0	0
	Max	97	0	100	0
	Average	17	0	83	0
	Median	12	0	88	0
Ratio of predicted to observed annual mean:					1.19
PBM (%)	Min	0	0	0	0
	Max	0	87	100	94
	Average	0	23	67	10
	Median	0	19	71	5
Ratio of predicted to observed annual mean:					1.19

c) Case 09+median

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	89	28	98	99
	Average	6	3	79	13
	Median	3	1	83	7

Ratio of predicted to observed annual mean:					0.93
GOM (%)	Min	0	0	0	0
	Max	95	100	100	0
	Average	14	1	85	0
	Median	10	0	90	0
Ratio of predicted to observed annual mean:					1.20
PBM (%)	Min	0	0	0	0
	Max	0	86	100	96
	Average	0	20	70	10
	Median	0	15	75	6
Ratio of predicted to observed annual mean:					1.14

d) Case 09+RM

Factor	Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt	
GEM (%)	Min	0	0	0	
	Max	16	54	98	
	Average	3	5	78	
	Median	2	2	83	
Ratio of predicted to observed annual mean:					0.97
RM (%)	Min	0	0	0	
	Max	37	83	100	
	Average	10	16	73	
	Median	7	11	78	
Ratio of predicted to observed annual mean:					1.04

e) Case 09-RM

Factor	Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt	
GEM (%)	Min	0	0	0	
	Max	34	15	98	
	Average	7	1	79	
	Median	4	1	84	
Ratio of predicted to observed annual mean:					0.97

f) Case 09ScaleRM

Factor name	Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt	
GEM (%)	Min	0	0	0	
	Max	0	65	100	
	Average	0	7	78	
	Median	0	4	83	
Ratio of predicted to observed annual mean:					0.97
GOM (%)	Min	0	0	0	
	Max	97	0	100	
	Average	23	0	75	
	Median	18	0	81	
Ratio of predicted to observed annual mean:					0.75

	Min	0	0	0	0
PBM	Max	0	88	100	96
(%)	Average	0	16	74	10
	Median	0	10	80	6
Ratio of predicted to observed annual mean:					0.94

Table S5. PMF factor contributions to speciated Hg and ratios of predicted to observed annual Hg concentrations in 2010.

a) Case 2010

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	100	9	99	100
	Average	11	1	79	9
	Median	7	1	85	4
Ratio of predicted to observed annual mean:					0.98
GOM (%)	Min	0	0	0	0
	Max	100	100	100	0
	Average	5	29	67	0
	Median	2	28	68	0
Ratio of predicted to observed annual mean:					1.34
PBM (%)	Min	0	0	0	0
	Max	100	28	98	99
	Average	11	4	80	5
	Median	6	3	86	2
Ratio of predicted to observed annual mean:					1.00

b) Case 10+mean

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	28	70	98	100
	Average	3	6	83	8
	Median	2	3	88	4
Ratio of predicted to observed annual mean:					0.96
GOM (%)	Min	0	0	0	0
	Max	22	100	100	0
	Average	2	28	70	0
	Median	1	22	76	0
Ratio of predicted to observed annual mean:					1.35
PBM (%)	Min	0	0	0	0
	Max	0	85	100	99
	Average	0	4	93	2
	Median	0	2	97	1
Ratio of predicted to observed annual mean:					0.87

c) Case 10+median

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	39	0	100	100
	Average	3	0	88	8
	Median	2	0	93	4

Ratio of predicted to observed annual mean:					0.97	
GOM (%)	Factor	Min	0	0	0	0
	Max	0	100	100	100	0
	Average	0	36	64	64	0
	Median	0	38	62	62	0
Ratio of predicted to observed annual mean:					1.32	
PBM (%)	Factor	Min	0	0	0	0
	Max	0	0	100	100	100
	Average	0	0	97	97	3
	Median	0	0	99	99	1
Ratio of predicted to observed annual mean:					0.88	

d) Case 10+RM

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt	
GEM (%)	Min	0	0	0	0	
	Max	100	9	99	100	
	Average	11	1	79	9	
	Median	7	1	85	4	
Ratio of predicted to observed annual mean:					0.98	
RM (%)	Factor	Min	0	0	0	0
	Max	100	47	98	96	
	Average	9	8	80	3	
	Median	5	6	86	1	
Ratio of predicted to observed annual mean:					1.16	

e) Case 10-RM

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt
GEM (%)	Min	0	0	0	0
	Max	100	10	99	100
	Average	11	1	78	9
	Median	7	1	85	4
Ratio of predicted to observed annual mean:					0.98

f) Case 10ScaleRM

Factor		Combustion emission	Industrial sulfur	Photochemistry & re-emission	Sea salt	
GEM (%)	Min	0	0	0	0	
	Max	100	10	99	100	
	Average	11	1	78	9	
	Median	7	1	85	4	
Ratio of predicted to observed annual mean:					0.98	
GOM (%)	Factor	Min	0	0	0	0
	Max	100	69	99	75	
	Average	8	14	77	1	
	Median	4	11	80	0	
Ratio of predicted to observed annual mean:					1.23	

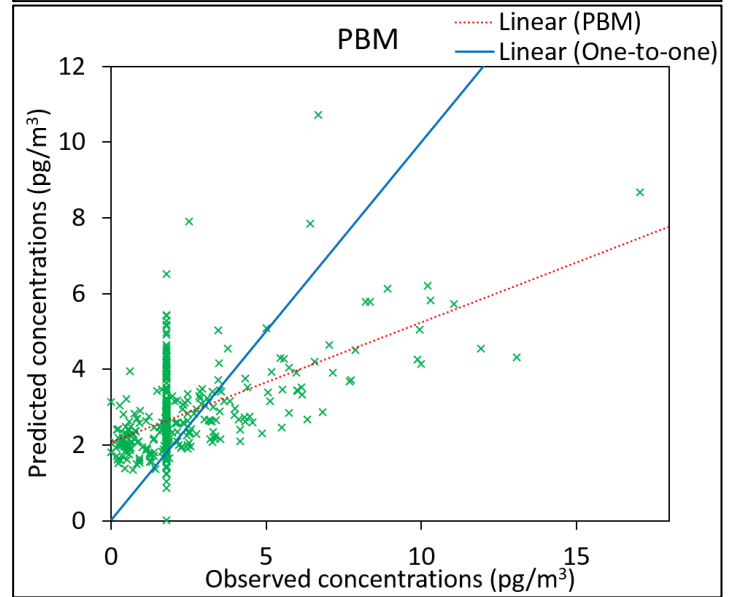
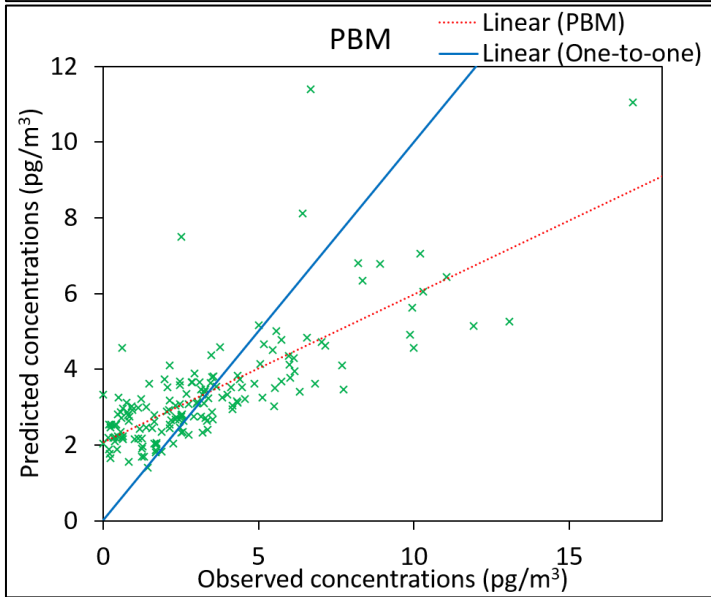
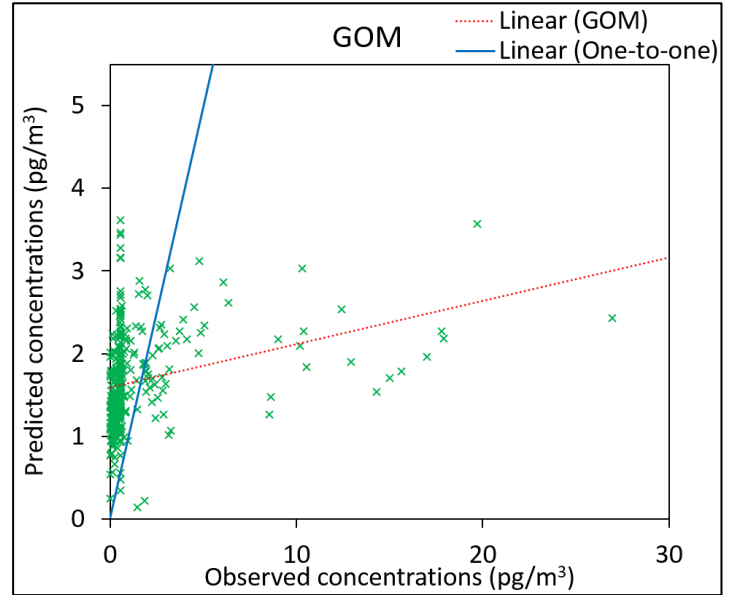
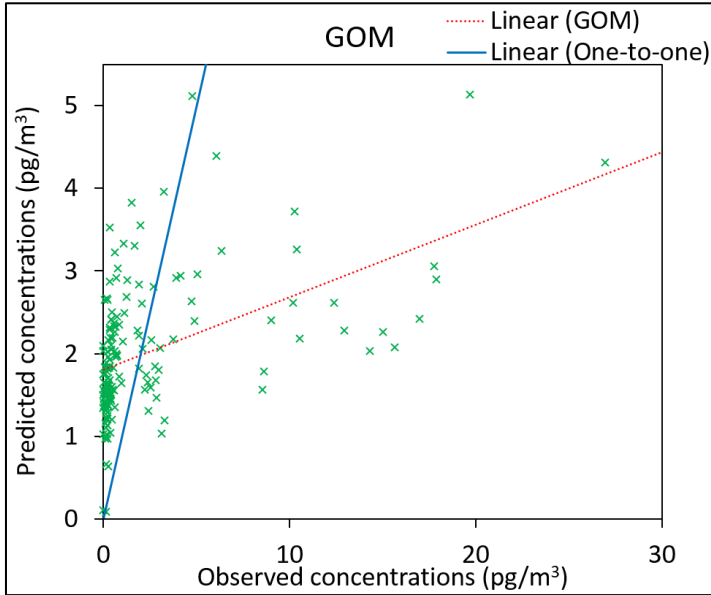
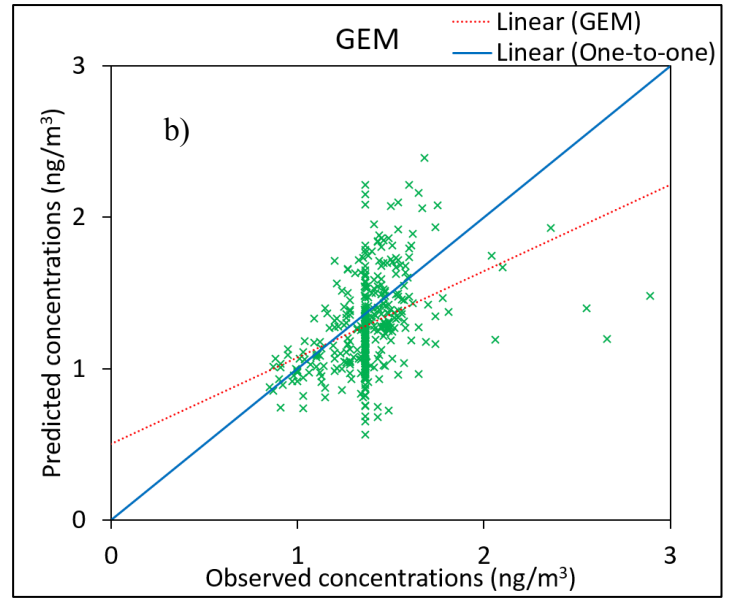
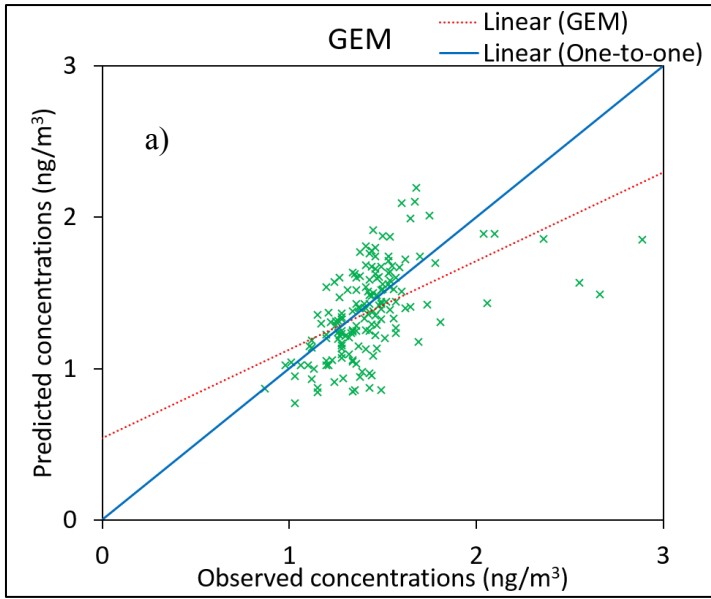
	Min	0	0	0	0
PBM	Max	100	40	97	98
(%)	Average	12	6	76	6
	Median	7	4	82	2
Ratio of predicted to observed annual mean:					0.88

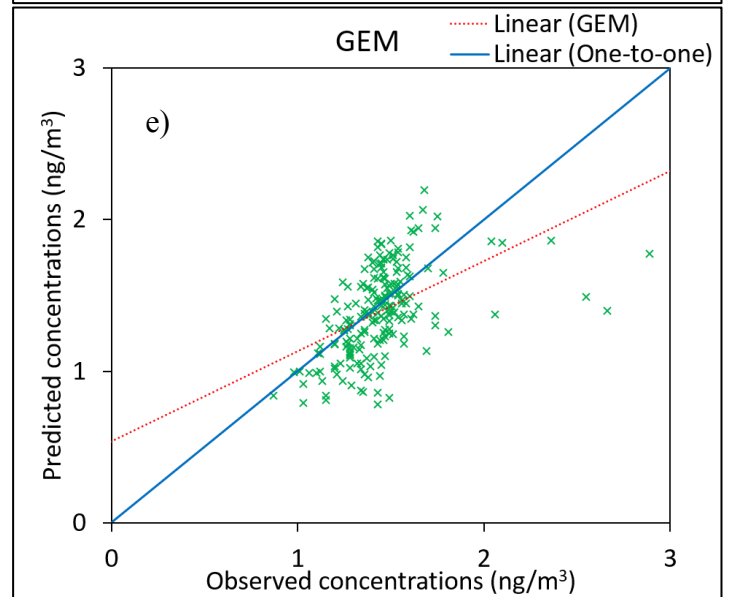
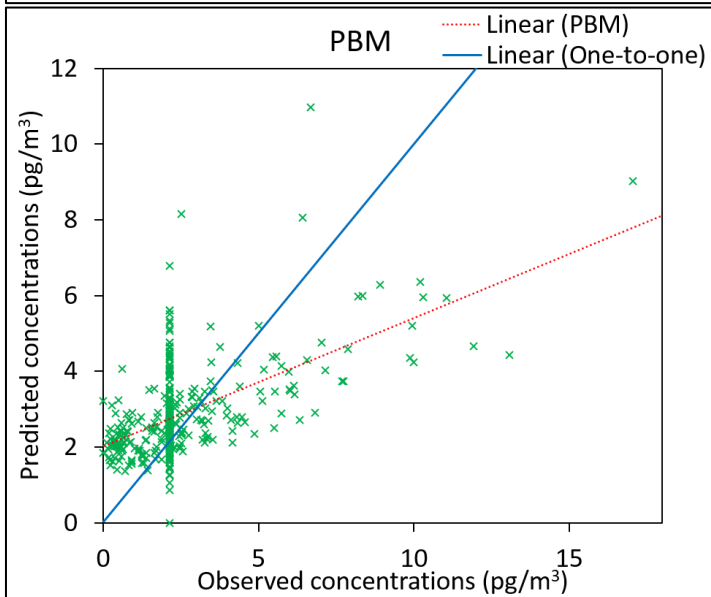
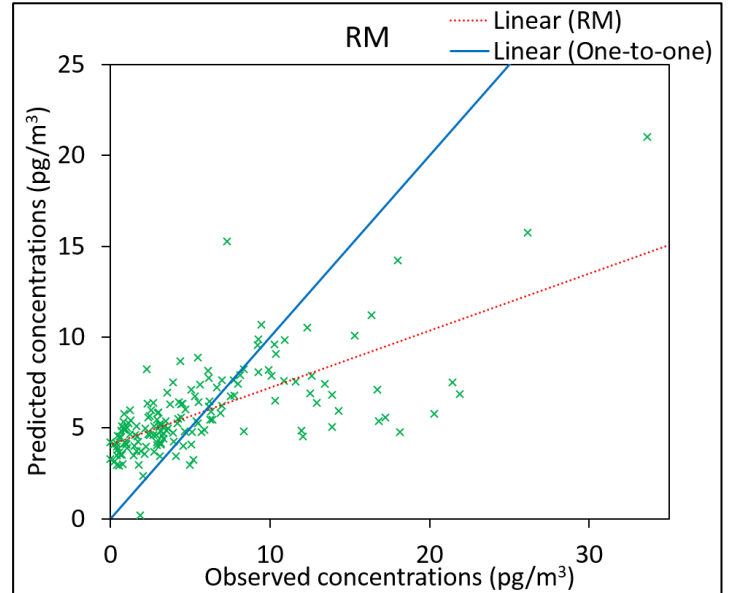
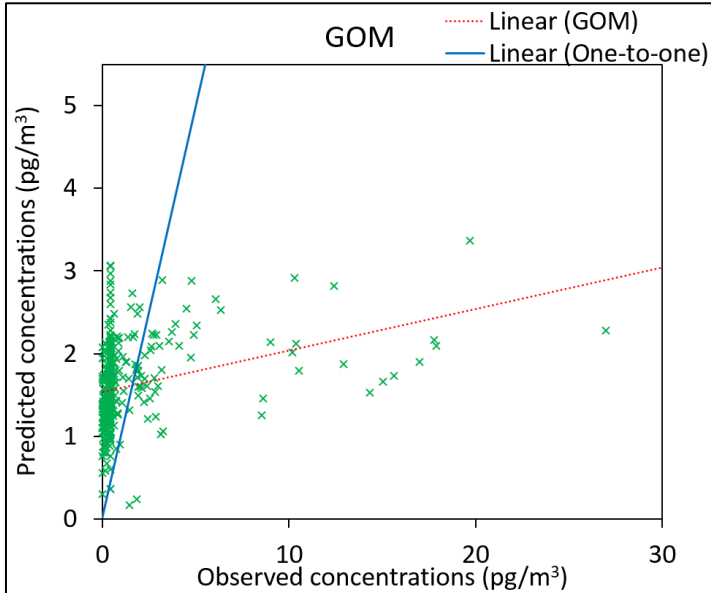
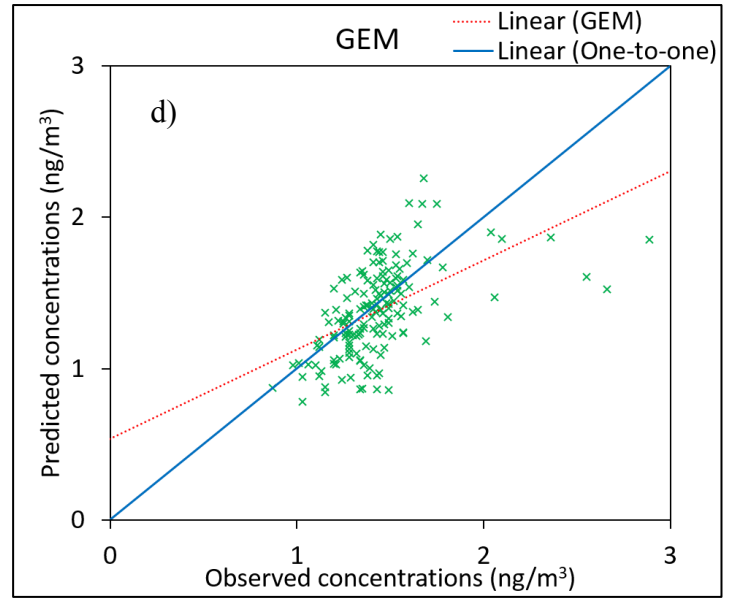
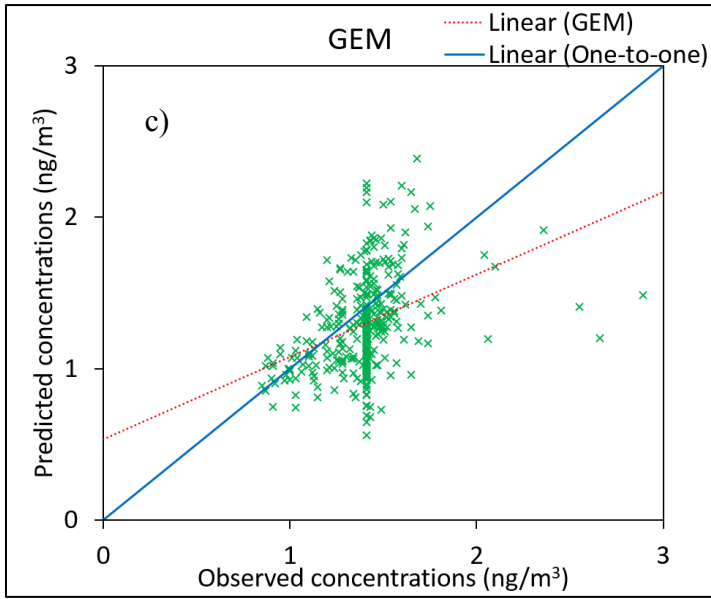
Table S6. Pearson correlation coefficients between Hg forms and other compounds in Case 2009, Case 09+mean, and Case 09+median (bold numbers are significant at <0.05).

	GEM200 9	GEM09+ mean	GEM09+ median	GOM200 9	GOM09+ mean	GOM09+ median	PBM200 9	PBM09+ mean	PBM09+ median
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	0.12	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	0.51	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	-0.03	-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	0.10	0.09	0.53	0.22	0.24
SO ₄ ²⁻	0.24	0.13	0.15	0.27	0.06	0.05	0.53	0.17	0.19

Table S7. Pearson correlation coefficients between Hg forms and other compounds in Case 2010, Case 10+mean, and Case 10+median (bold numbers are significant at <0.05).

	GEM201 0	GEM10+ mean	GEM10+ median	GOM201 0	GOM10+ mean	GOM10+ median	PBM201 0	PBM10+ mean	PBM10+ median
GEM	1.00	1.00	1.00	0.32	0.29	0.29	0.11	0.19	0.19
GOM	0.32	0.29	0.29	1.00	1.00	1.00	0.29	0.15	0.15
PBM	0.11	0.19	0.19	0.29	0.15	0.15	1.00	1.00	1.00
O ₃	0.70	0.68	0.68	0.56	0.51	0.51	0.32	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⁺	-0.01	-0.02	-0.02	0.07	0.08	0.09	-0.02	-0.05	-0.05
K ⁺	-0.13	-0.12	-0.13	0.07	0.09	0.10	-0.03	-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 ⁻	-0.02	-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	-0.11	0.06	0.08	0.08	0.00	-0.05	-0.05





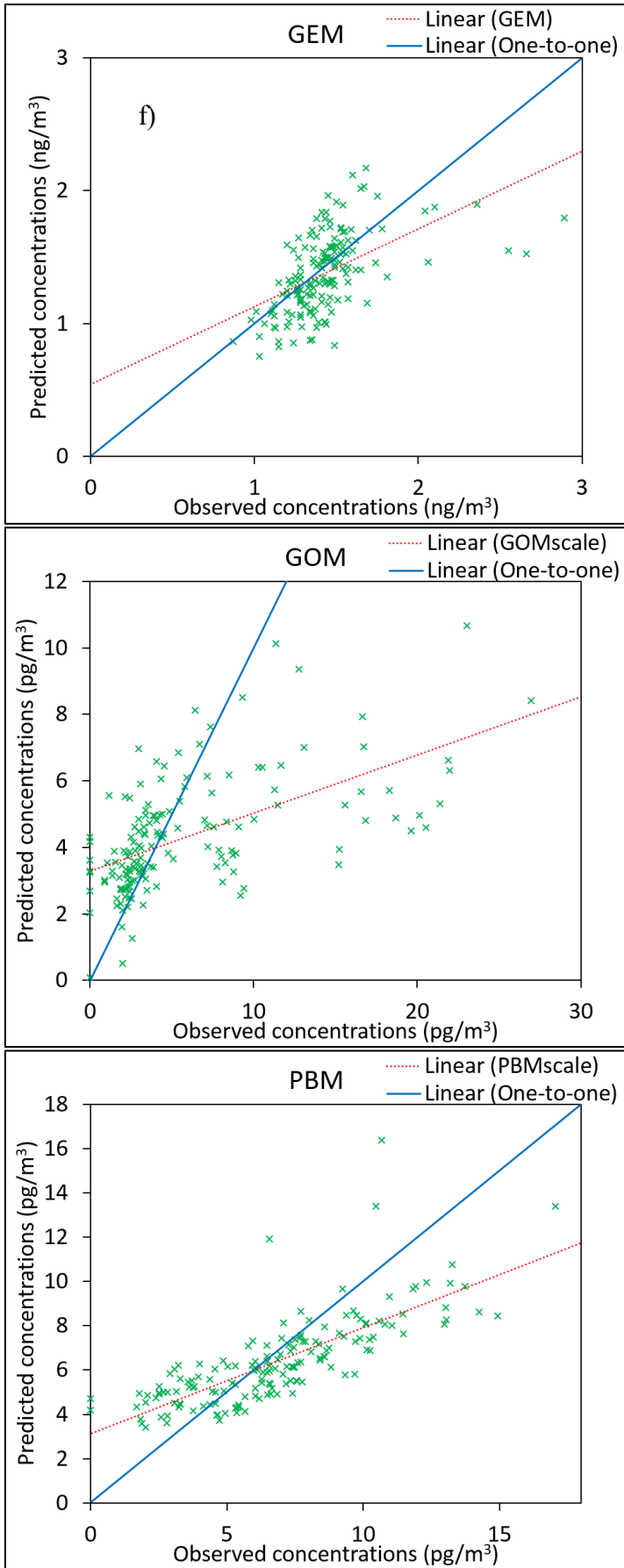
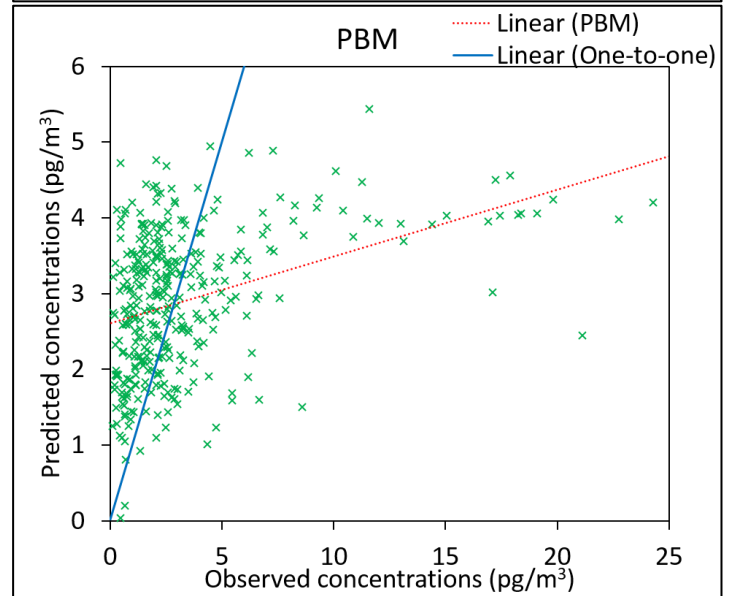
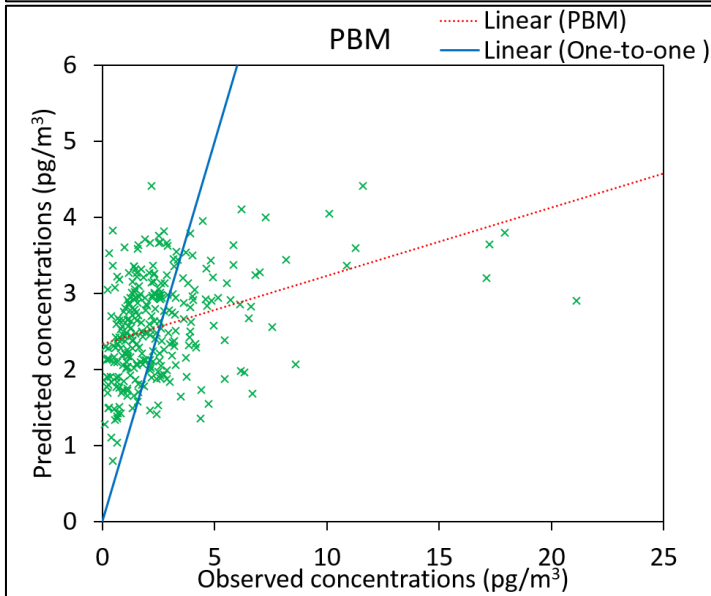
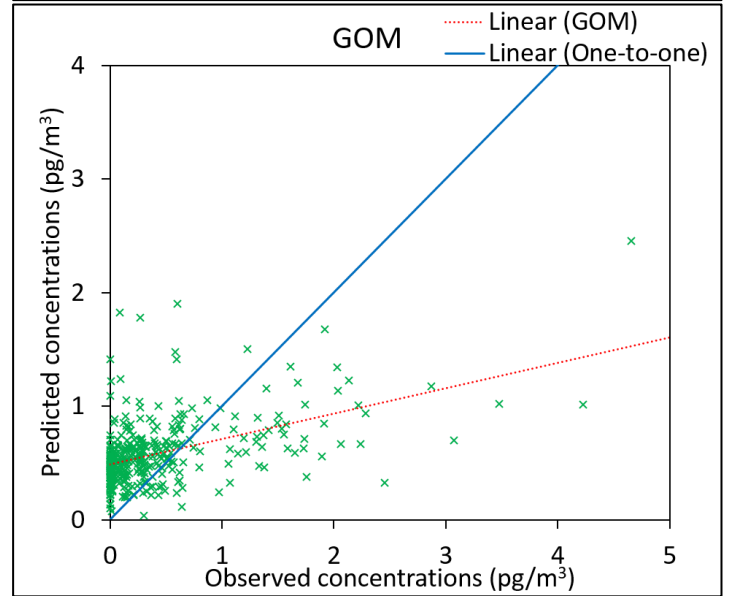
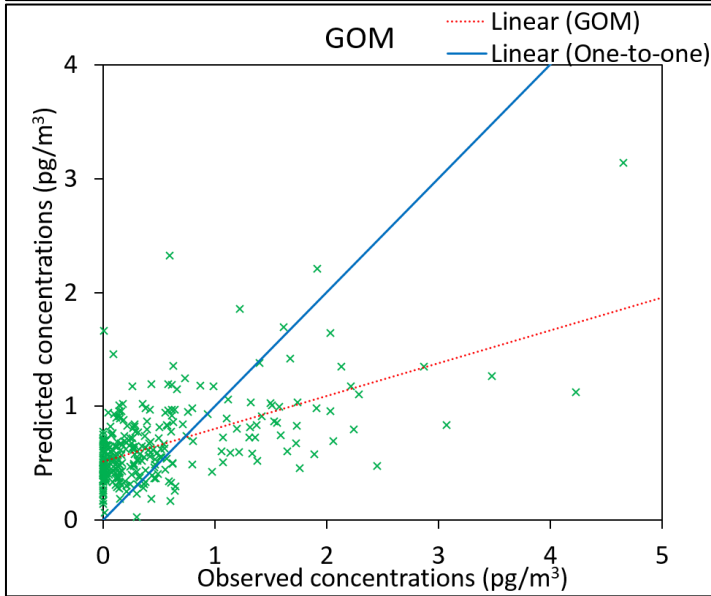
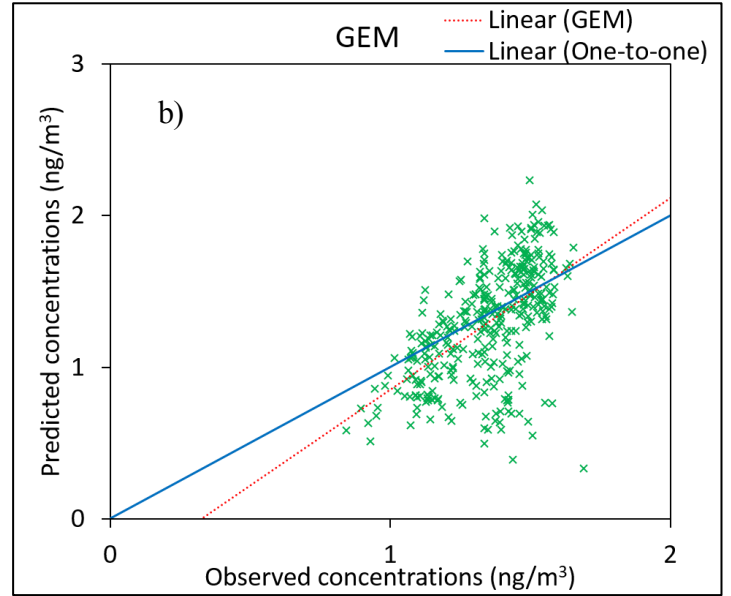
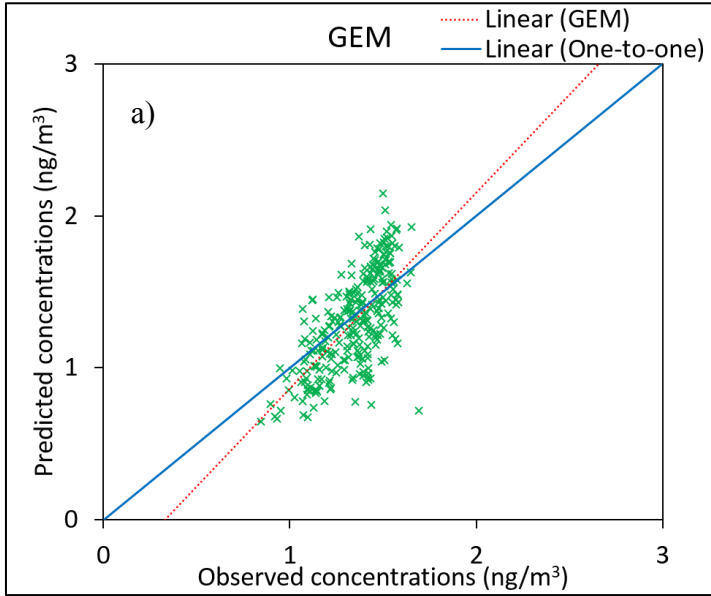
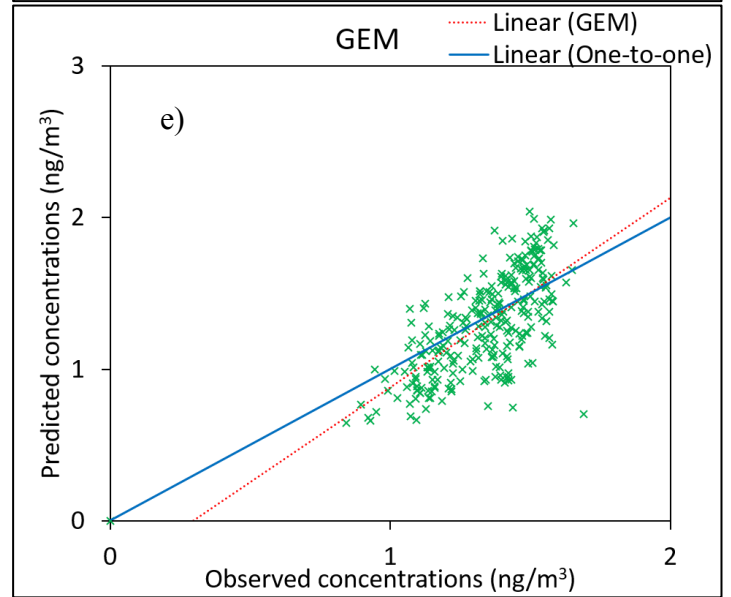
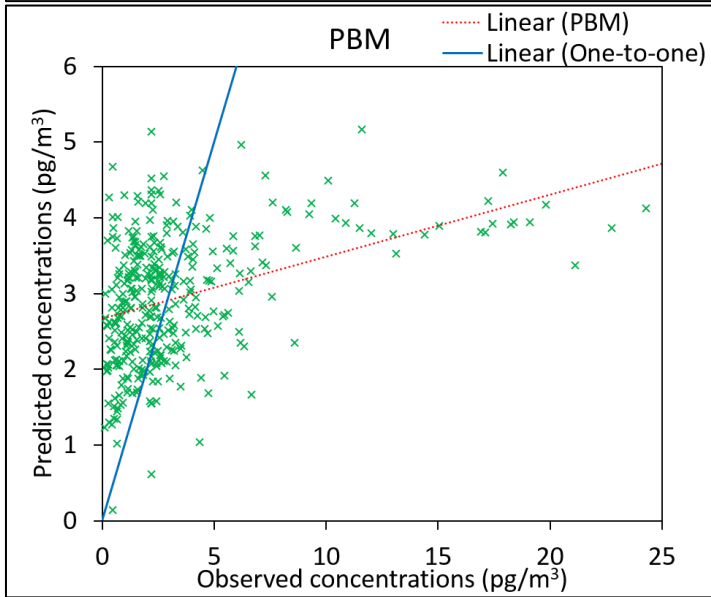
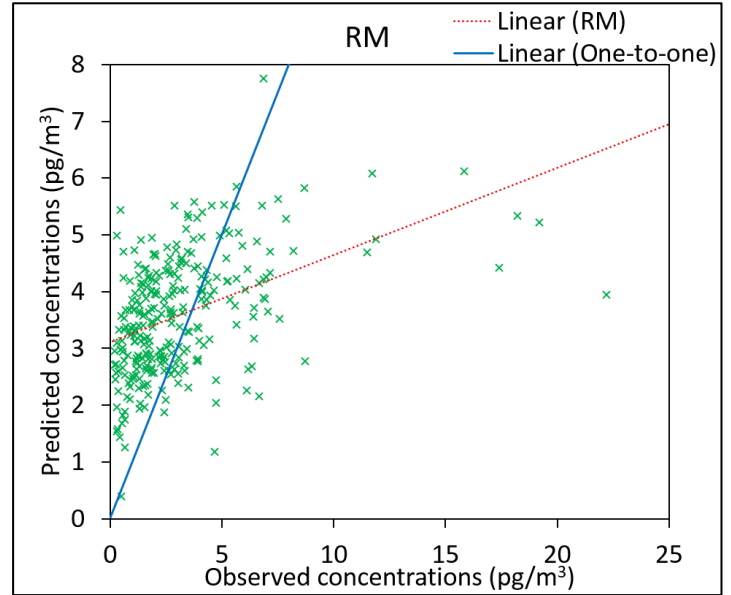
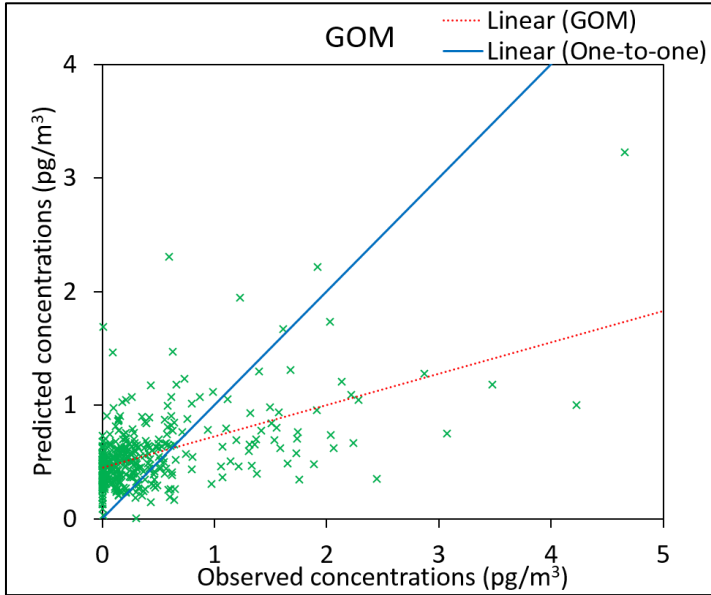
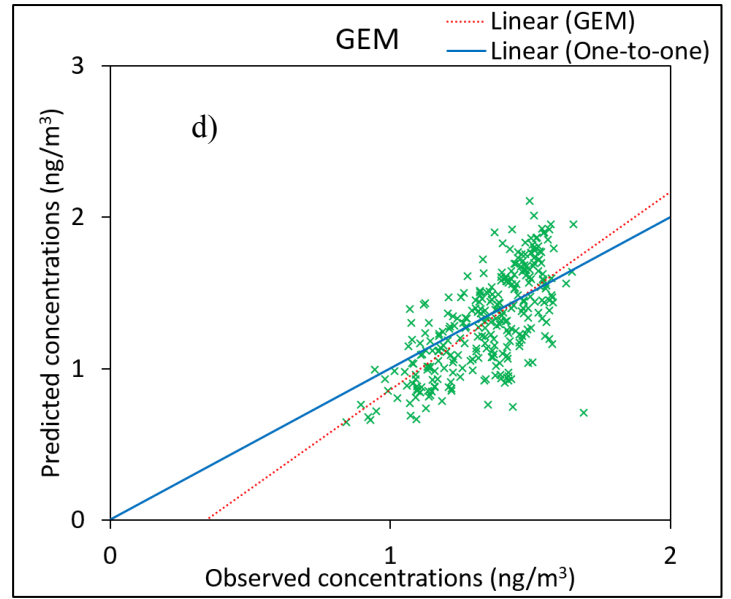
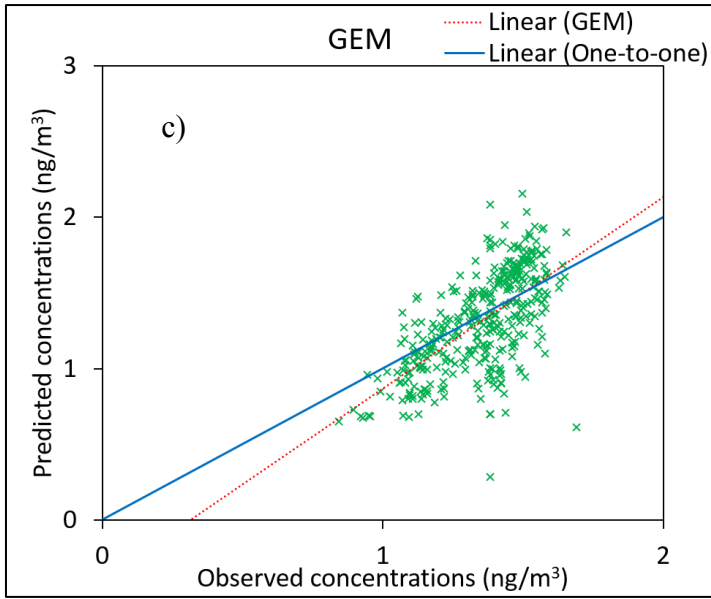


Figure S51. Obs/Pred scatter plot in 2009. a) Case 2009, b) Case 09+mean, c) Case 09+median, d) Case 09+RM, e) Case 09-RM, and f) Case 09ScaleRM, observed GOM and PBM have been

scaled.





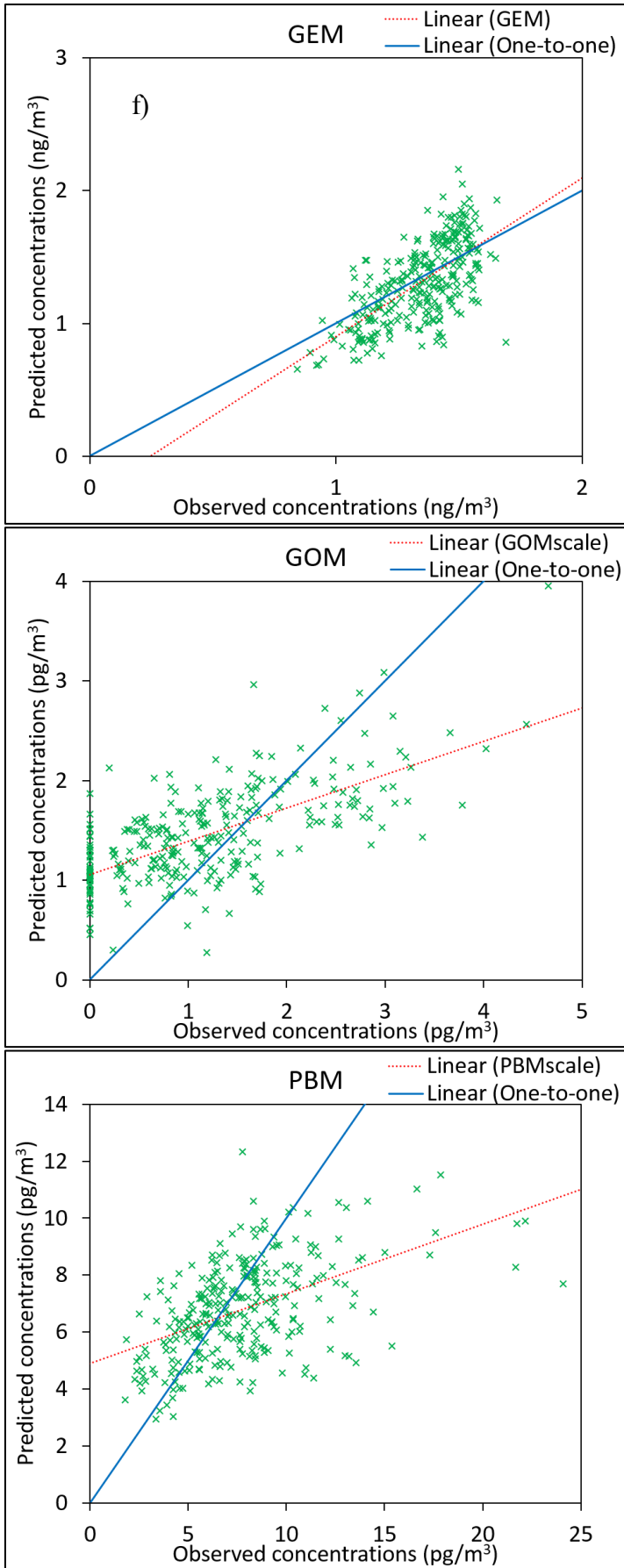
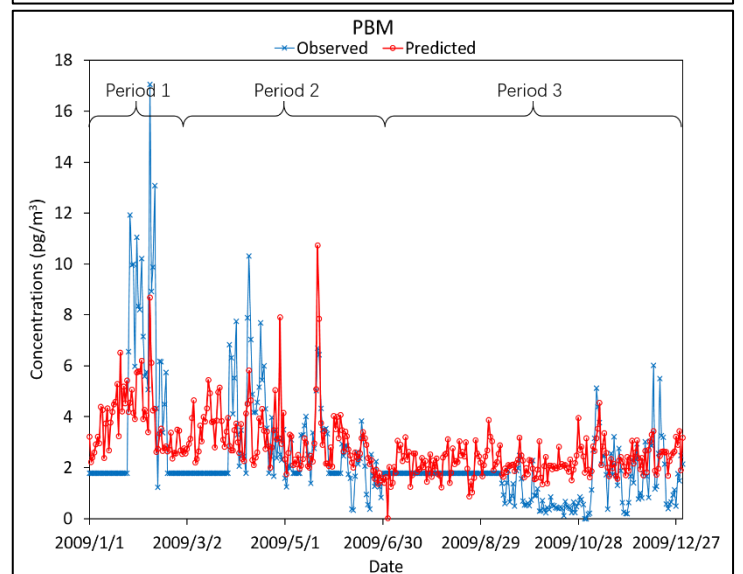
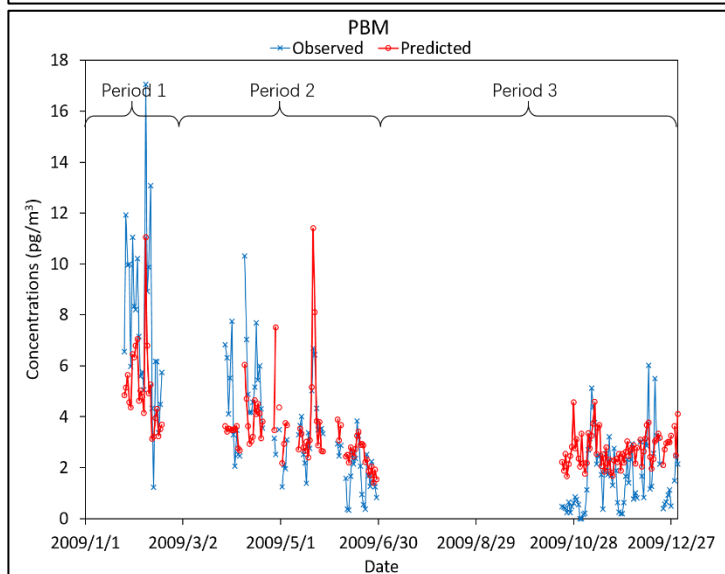
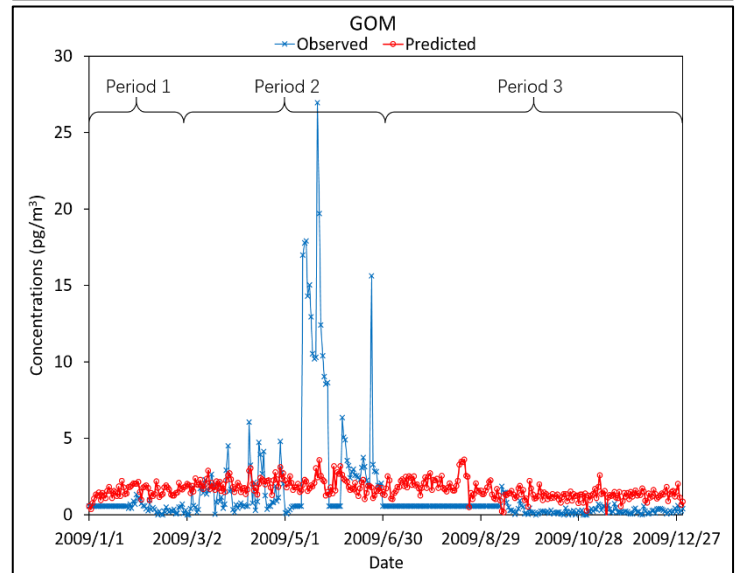
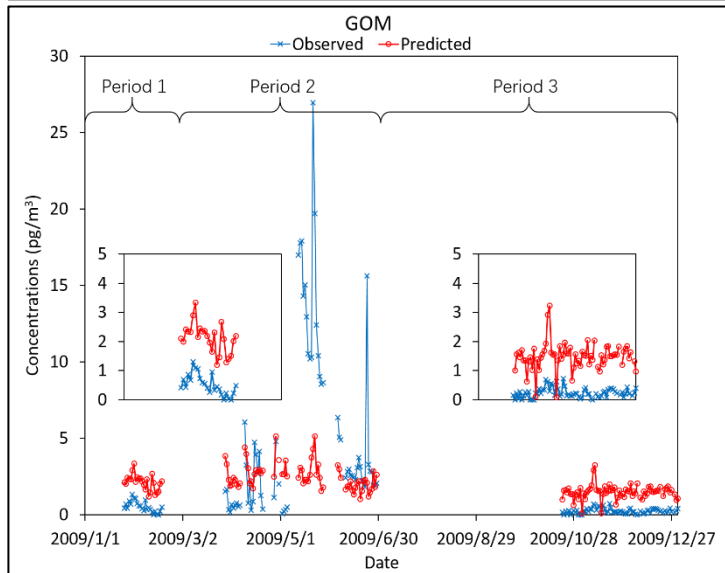
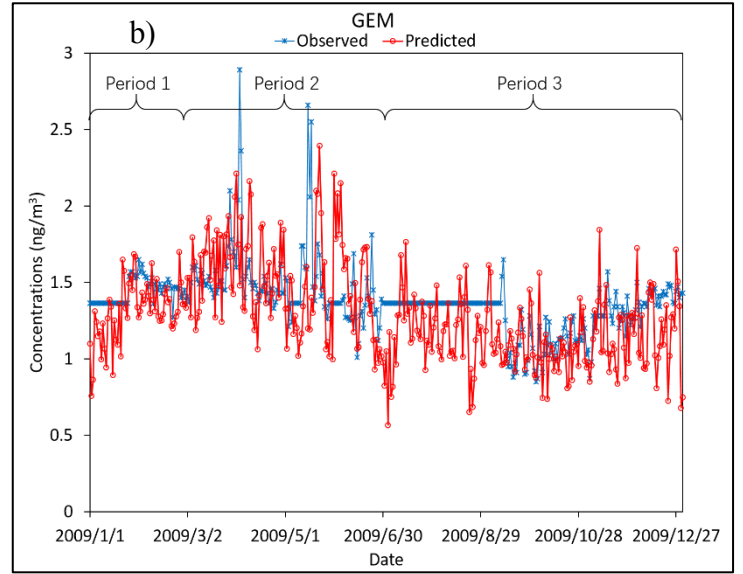
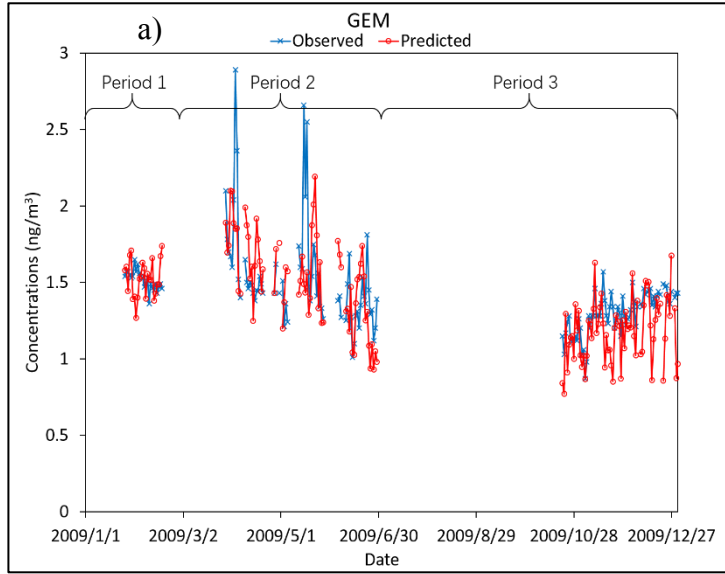
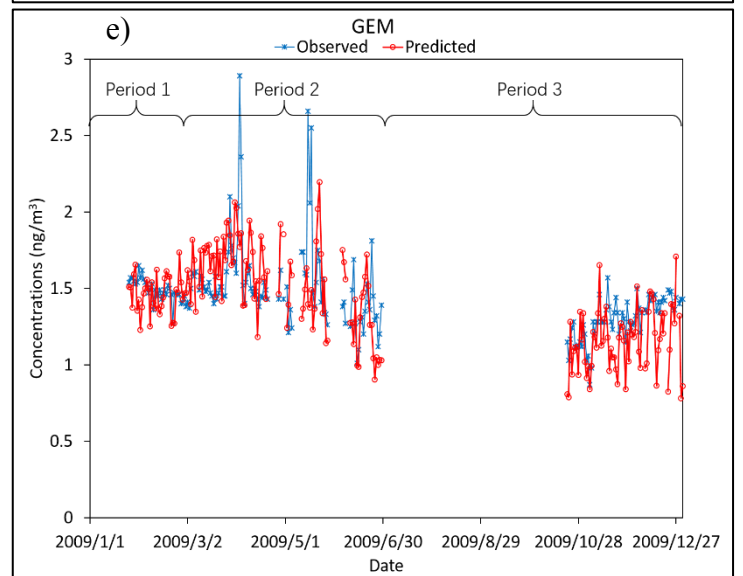
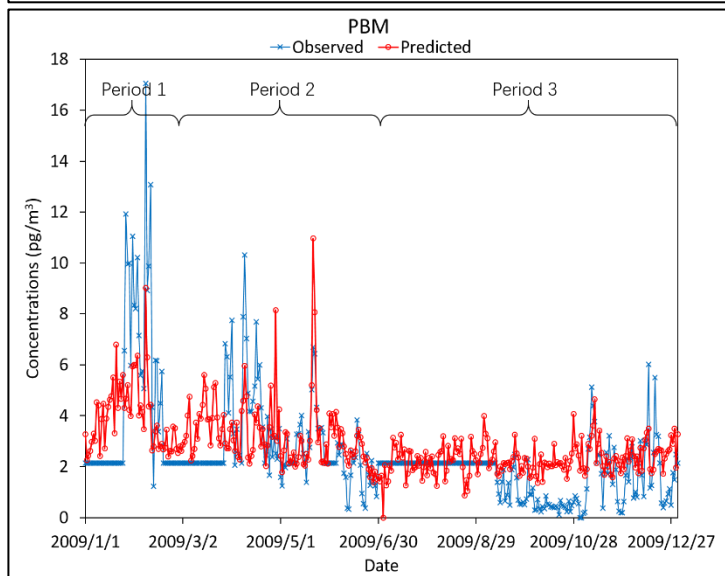
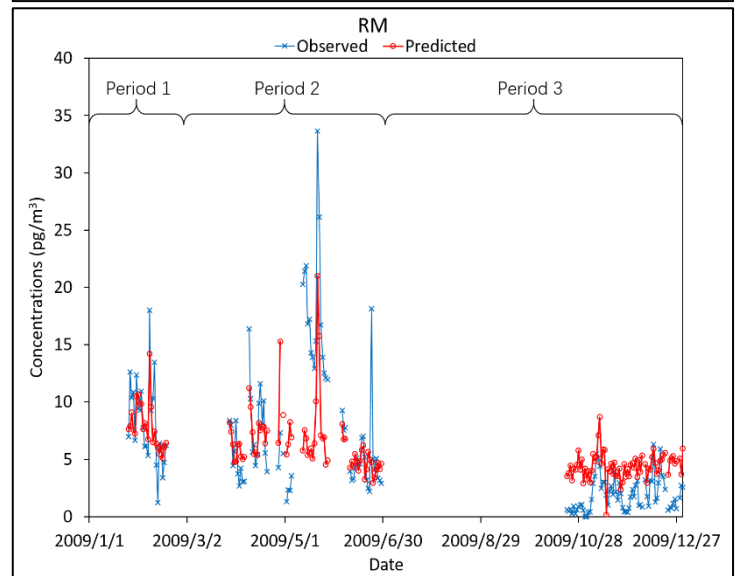
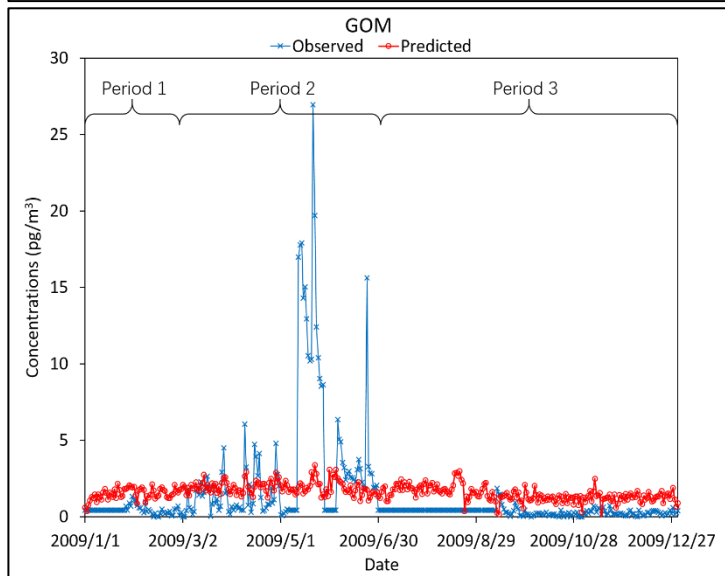
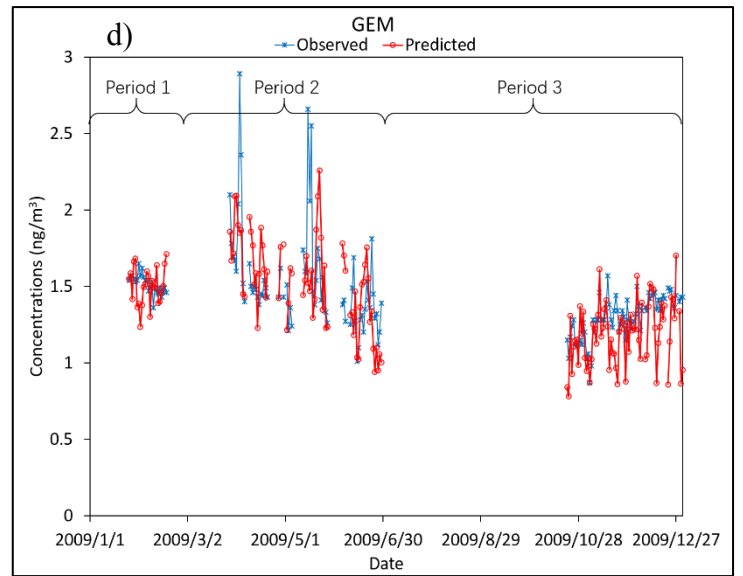
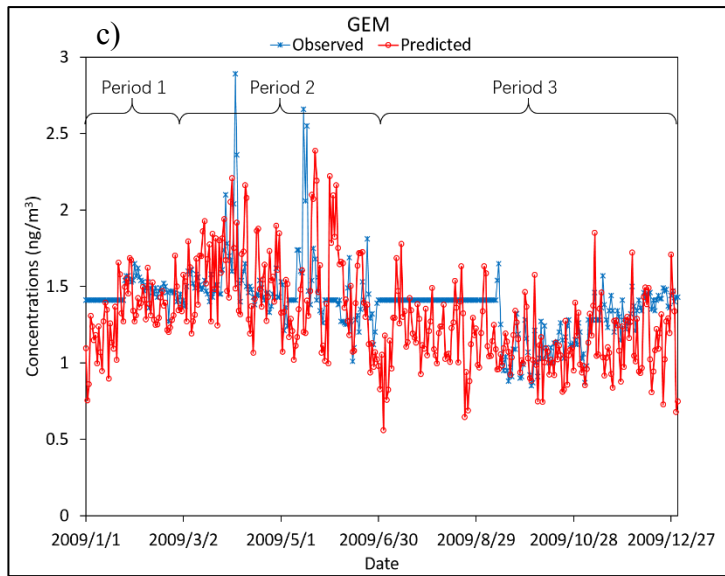


Figure S62. Obs/Pred scatter plot in 2010. a) Case 2010, b) Case 10+mean, c) Case 10+median, d) Case 10+RM, e) Case 10-RM, and f) Case 10ScaleRM, observed GOM and PBM have been

scaled.





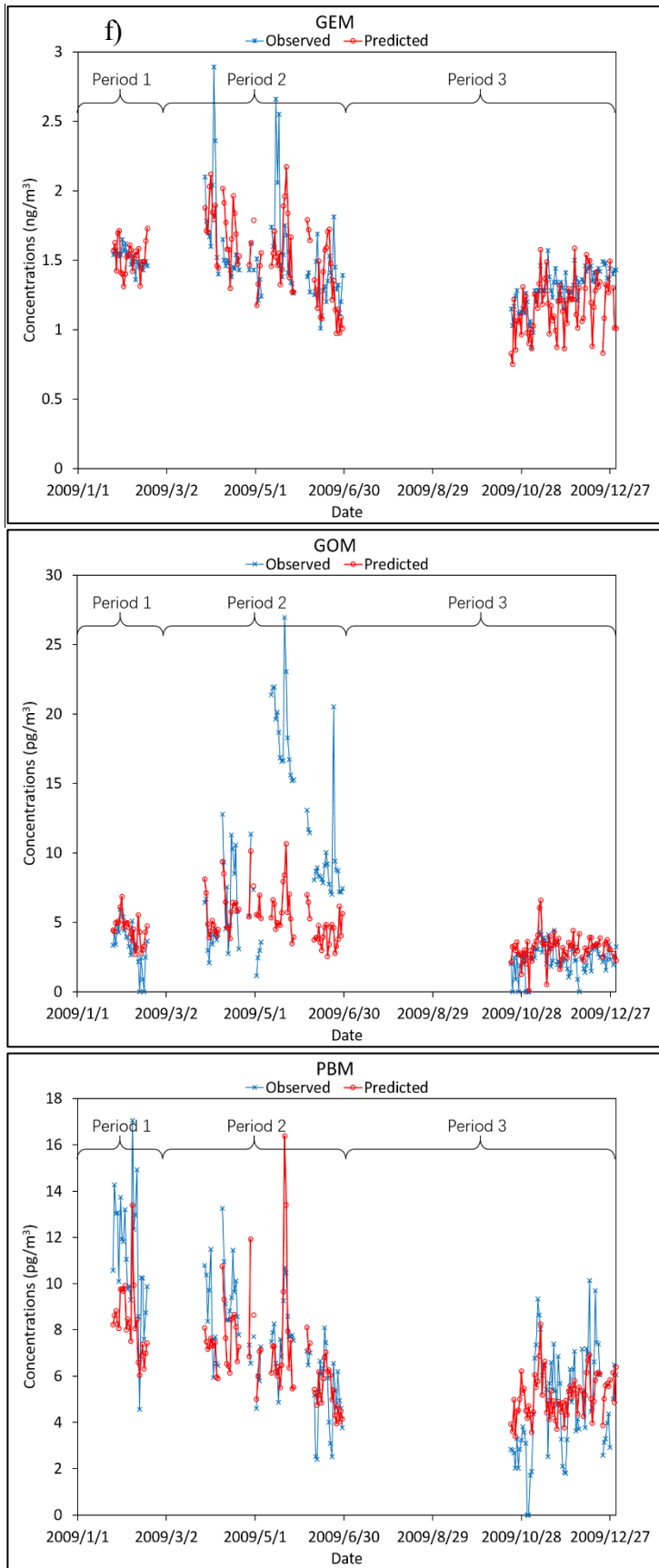
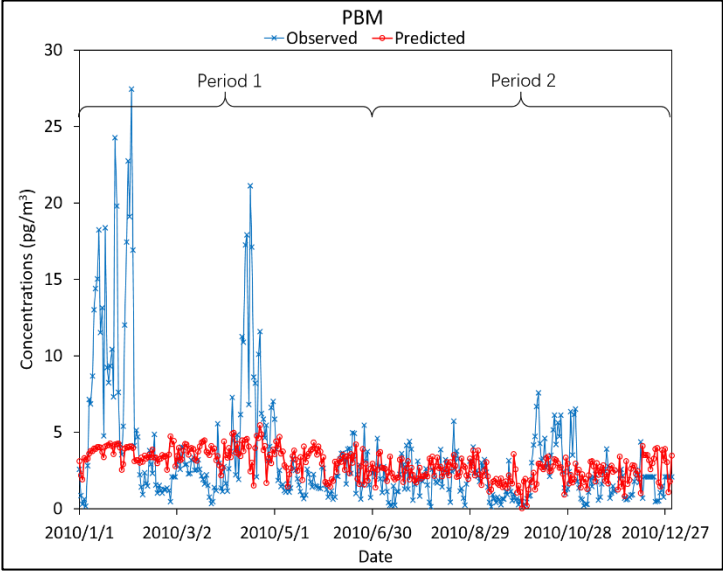
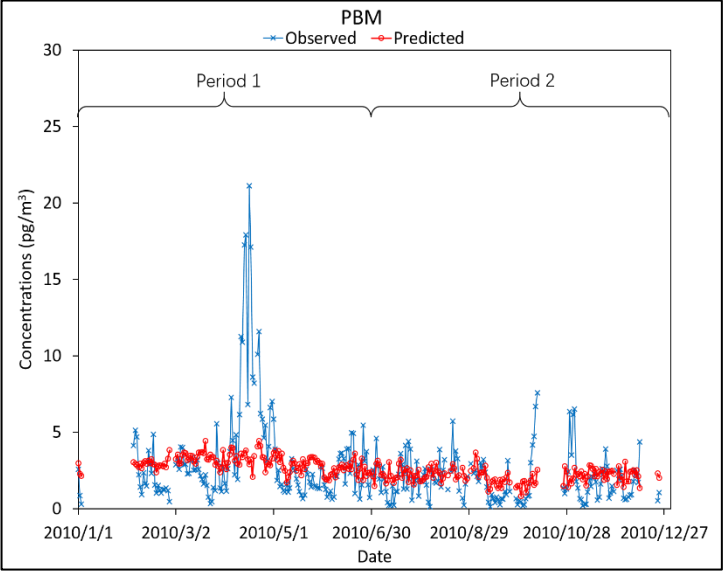
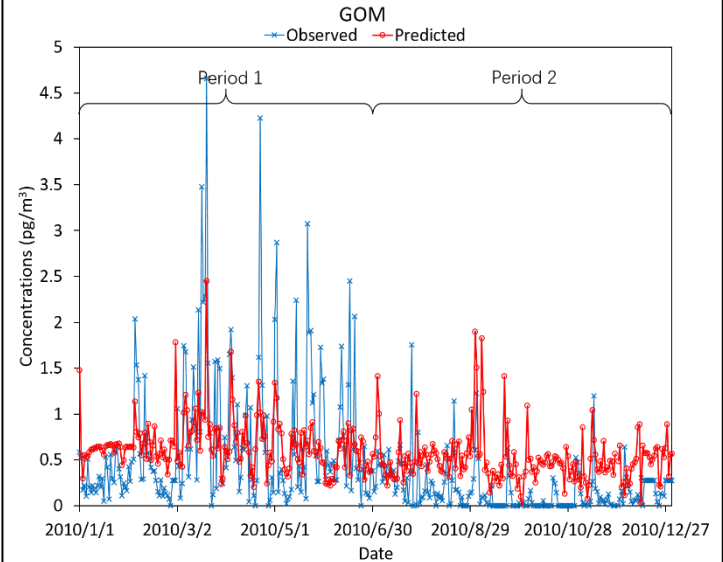
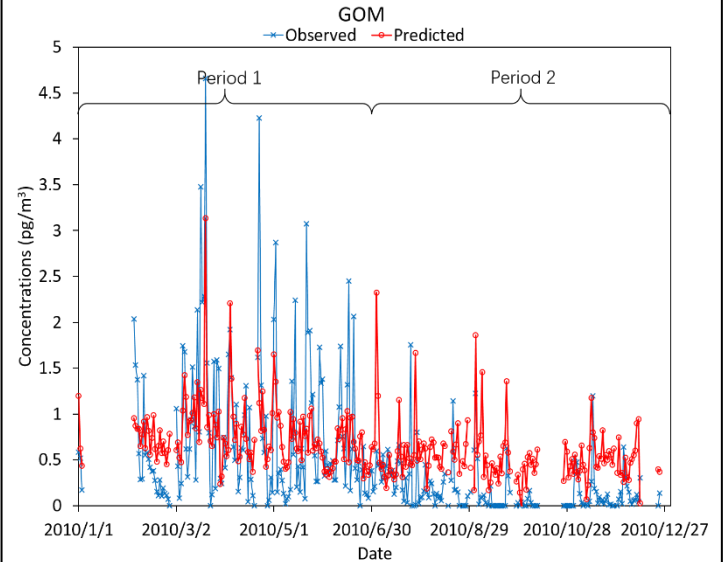
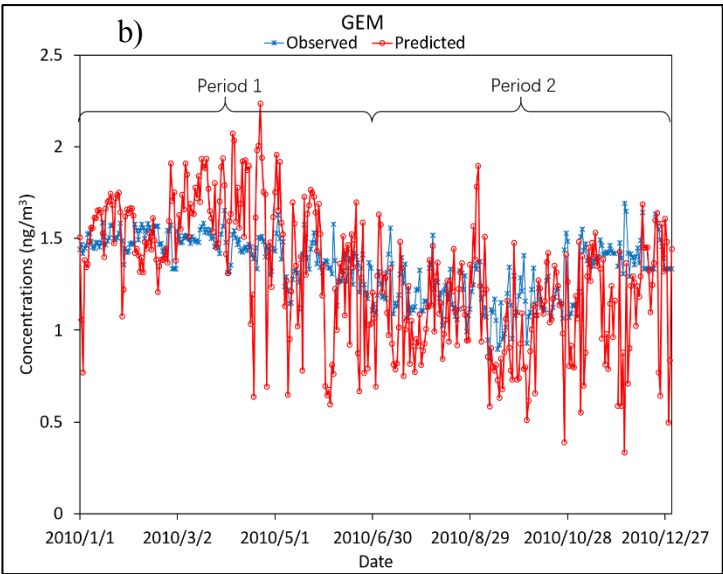
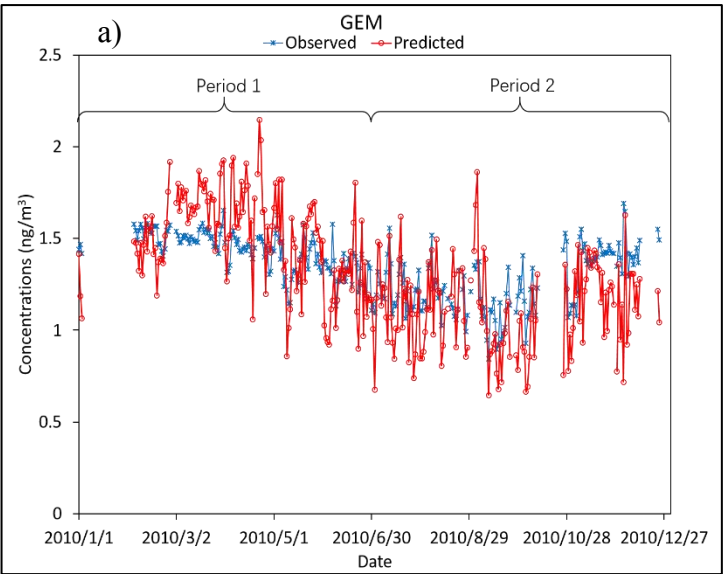
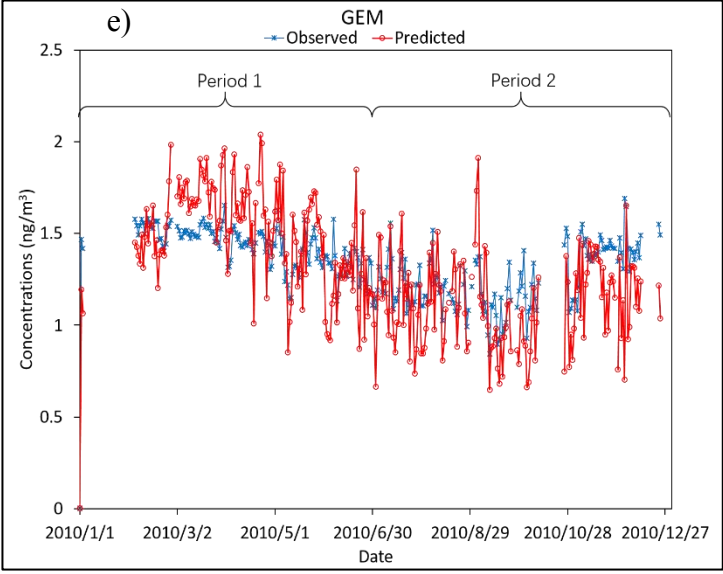
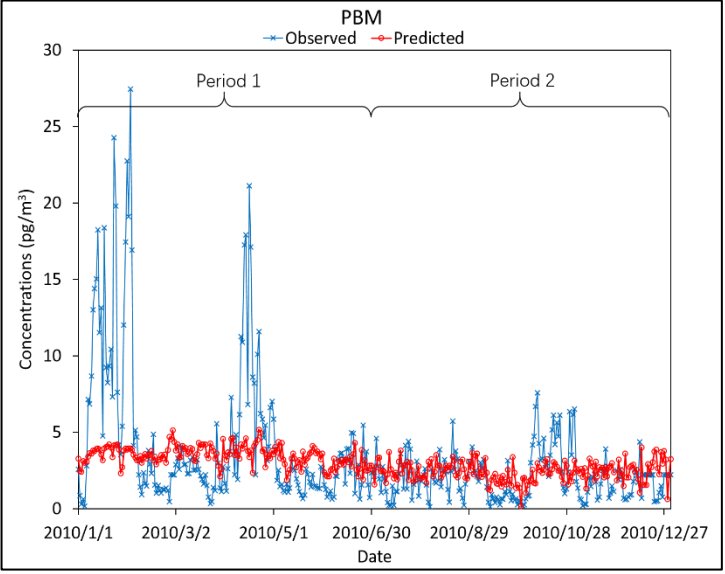
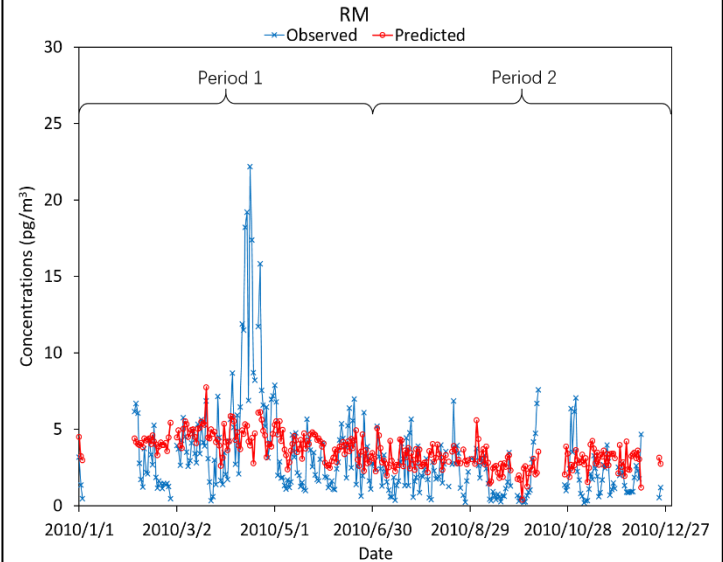
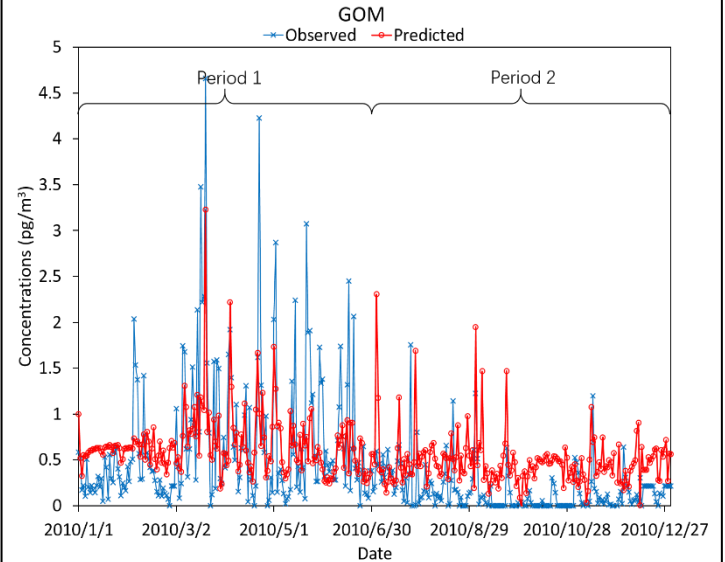
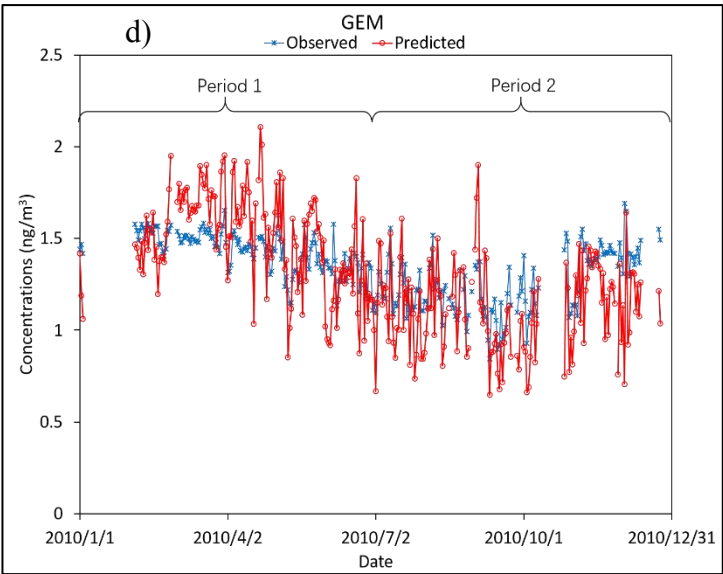
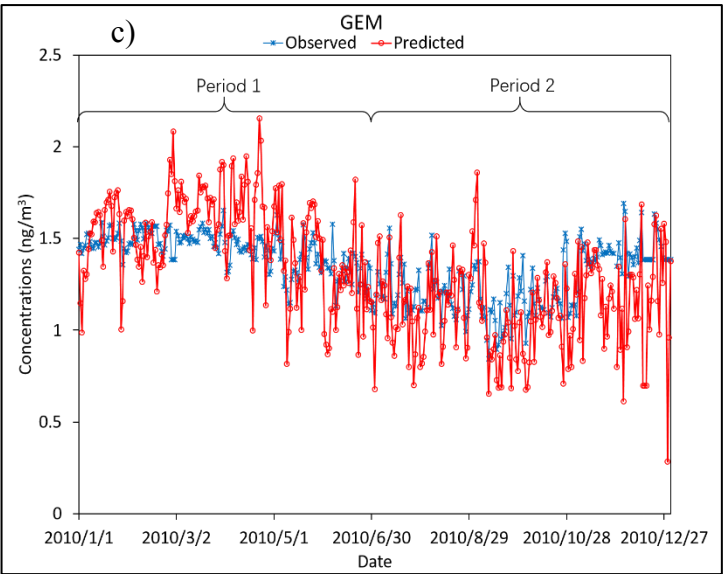


Figure S73. Obs/Pred time series in 2009. a) Case 09, b) Case 09+mean, c) Case 09+median, d) Case 09+RM, e) Case 09-RM, and f) Case 09ScaleRM, observed GOM and PBM have been scaled.





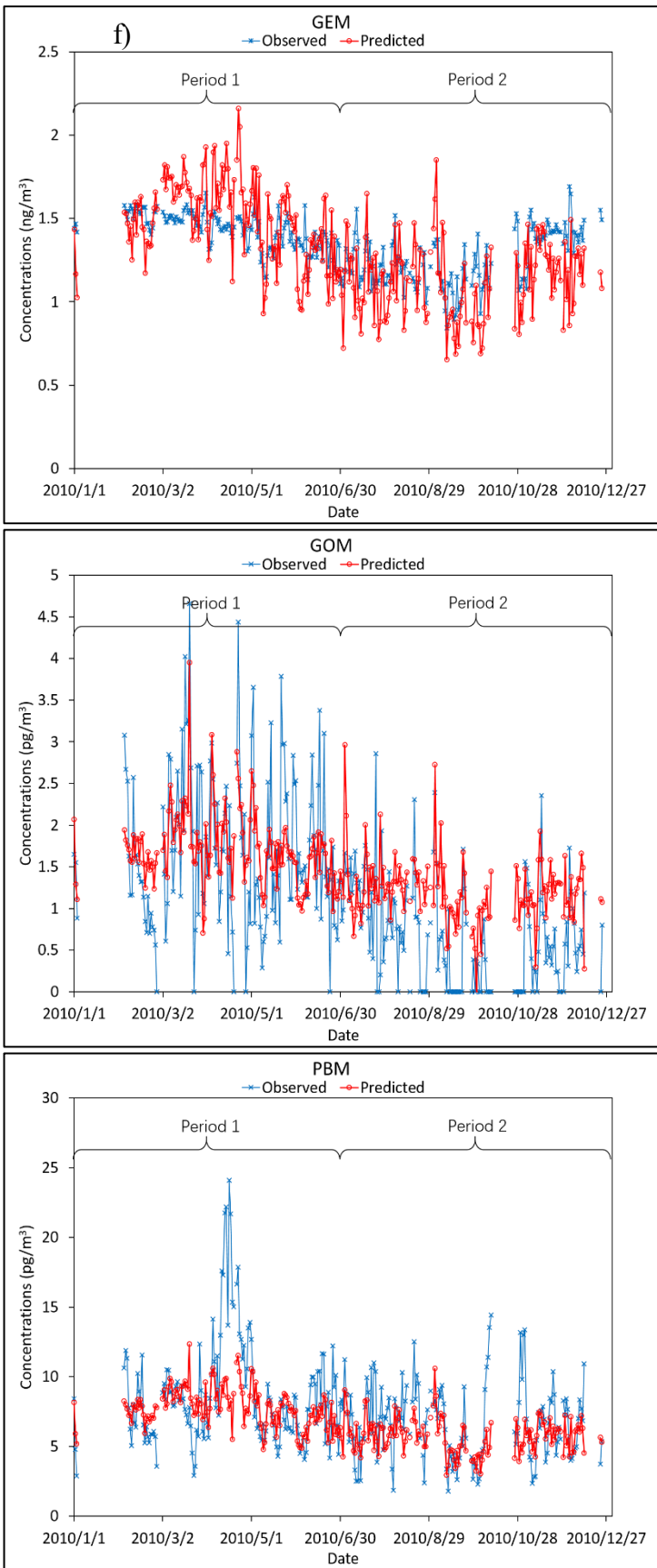


Figure S84. Obs/Pred time series in 2010. a) Case 2010, b) Case 10+mean, c) Case 10+median, d) Case 10+RM, e) Case 10-RM, and f) Case 10ScaleRM, observed GOM and PBM have been scaled.