

Estimation of Organic Carbon Blank Values and Error Structures of the Speciation Trends Network Data for Source Apportionment

Eugene Kim , Philip K. Hopke & Youjun Qin

To cite this article: Eugene Kim , Philip K. Hopke & Youjun Qin (2005) Estimation of Organic Carbon Blank Values and Error Structures of the Speciation Trends Network Data for Source Apportionment, Journal of the Air & Waste Management Association, 55:8, 1190-1199, DOI: [10.1080/10473289.2005.10464705](https://doi.org/10.1080/10473289.2005.10464705)

To link to this article: <http://dx.doi.org/10.1080/10473289.2005.10464705>



Published online: 01 Mar 2012.



Submit your article to this journal [↗](#)



Article views: 262



View related articles [↗](#)



Citing articles: 37 View citing articles [↗](#)

Estimation of Organic Carbon Blank Values and Error Structures of the Speciation Trends Network Data for Source Apportionment

Eugene Kim

Department of Civil and Environmental Engineering, Clarkson University, Potsdam, NY

Philip K. Hopke and Youjun Qin

Department of Chemical Engineering, Clarkson University, Potsdam, NY

ABSTRACT

Because the particulate organic carbon (OC) concentrations reported in U.S. Environment Protection Agency Speciation Trends Network (STN) data were not blank corrected, the OC blank concentrations were estimated using the intercept in particulate matter ≤ 2.5 μm in aerodynamic diameter ($\text{PM}_{2.5}$) regression against OC concentrations. The estimated OC blank concentrations ranged from 1 to 2.4 $\mu\text{g}/\text{m}^3$ showing higher values in urban areas for the 13 monitoring sites in the northeastern United States. In the STN data, several different samplers and analyzers are used, and various instruments show different method detection limit (MDL) values, as well as errors. A comprehensive set of error structures that would be used for numerous source apportionment studies of STN data was estimated by comparing a limited set of measured concentrations and their associated uncertainties. To examine the estimated error structures and investigate the appropriate MDL values, $\text{PM}_{2.5}$ samples collected at a STN site in Burlington, VT, were analyzed through the application of the positive matrix factorization. A total of 323 samples that were collected between December 2000 and December 2003 and 49 species based on several variable selection criteria were used, and eight sources were successfully identified in this study with the estimated error structures and min values among different MDL values from the five instruments: secondary sulfate aerosol (41%), secondary nitrate aerosol (20%), airborne soil (15%), gasoline vehicle emissions (7%), diesel emissions

(7%), aged sea salt (4%), copper smelting (3%), and ferrous smelting (2%). Time series plots of contributions from airborne soil indicate that the highly elevated impacts from this source were likely caused primarily by dust storms.

INTRODUCTION

U.S. Environmental Protection Agency (EPA) established the Speciation Trends Network (STN) to characterize particulate matter ≤ 2.5 μm in aerodynamic diameter ($\text{PM}_{2.5}$) composition in urban areas and to assist identifying areas out of attainment of the promulgated new national ambient air quality standards for airborne PM. Advanced source apportionment studies for the STN $\text{PM}_{2.5}$ measurements are needed for developing effective control strategies for $\text{PM}_{2.5}$, as well as for the source-specific community epidemiology to relate adverse health effects to apportioned source contributions. Positive matrix factorization (PMF)¹ has been successfully used to apportion $\text{PM}_{2.5}$ sources in several Interagency Monitoring of Protected Visual Environments (IMPROVE)² monitoring sites.^{3–7}

The reported particulate organic carbon (OC) concentrations in STN data were not blank corrected. There are four different types of samplers used in STN. Particulate carbons were analyzed by three analyzers, anions were analyzed by two instruments, and cations were analyzed by three instruments. Particulate metals were analyzed by five instruments located in three laboratories. Because various instruments show different method detection limit (MDL) values and uncertainties, a comprehensive set of error structures and MDL values that would be used for PMF studies were required.

The objectives of this study were to estimate the OC blank concentrations, error structures, and MDL values for the STN data. To examine the use of those estimations for the analyses of STN data, PMF was applied to ambient

IMPLICATIONS

Although there are major challenges in using STN data for source apportionment, it is possible to estimate OC blank values and errors to permit quantitative receptor modeling to be applied to these data.

PM_{2.5} speciation data collected at the STN monitoring site in Burlington, VT, in the present study. The PMF-derived PM_{2.5} sources and their seasonal trends are discussed.

EXPERIMENTAL WORK

Sample Collection and Chemical Analysis

Integrated 24-hr PM_{2.5} samples were collected at 54 core STN monitoring sites and >170 other monitoring sites. Four types of samplers were used to collect ambient PM_{2.5} samples in the STN: Spiral Aerosol Speciation Samplers (Met One Instruments, Grants Pass, OR), R&P Speciation Sampler 2300 (Rupprecht & Patashnick, East Greenbush, NY), Mass Aerosol Speciation Sampler (URG, Chapel Hill, NC), and Reference Ambient Air Sampler (Andersen Instruments, Smyrna, GA). PM_{2.5} samples were taken on a 1-in-3 or 1-in-6 day schedule. Field blanks were collected for 1-in-10 routine samples, and trip blanks were obtained at 1-in-30 routine samplers.⁸ Unlike IMPROVE network, the STN used multiple types of samples and multiple analytical laboratories to produce the data. There were also differences in the nature of the collected blanks and the treatment of the resulting data. PM_{2.5} samples were collected on Teflon, Nylon, and quartz filters. The Teflon filter was used for mass concentrations and analyzed via five energy dispersive X-ray fluorescence (XRF) spectrometers for the elemental analysis located in three laboratories: Chester LabNet, Cooper Environmental Services, and Research Triangle Institute (RTI). The Nylon filter was analyzed for sulfate (SO₄²⁻), nitrate (NO₃⁻), ammonium (NH₄⁺), sodium (Na⁺), and potassium (K⁺) via ion chromatography (IC). To minimize the sampling artifacts for NO₃⁻, a magnesium oxide (MgO) or sodium carbonate denuder was included at the upstream of the Nylon filter.^{9,10} Two instruments for anions and three instruments for cations in RTI were used for the Nylon filter analyses. The quartz filter was analyzed by three instruments at RTI via National Institute for Occupational Safety and Health/Thermal Optical Transmittance protocol¹¹ for OC and elemental carbon (EC). Carbon denuders that minimize positive sampling artifact caused by adsorption of gaseous organic materials were not included at the upstream of quartz filter in the STN samplers.^{12,13} None of the reported STN data were blank corrected.¹⁴

OC Blank Corrections

Tolocka et al.¹⁵ in a comparison study among STN samplers (i.e., Reference Ambient Air Sampler, Federal Reference Method Sampler, and Versatile Air Pollution Sampler) observed that the OC concentration measured by both STN and Federal Reference Method samplers that did not include carbon denuder at the upstream of quartz filter were consistently higher than the values sampled by Versatile Air Pollution Sampler that had a carbon denuder

preceding the quartz filter. Because the reported particulate OC concentrations were not blank corrected and there appears to be a positive artifact in the OC concentrations measured by STN samplers,¹⁵ approaches to obtaining an integrated estimate of the OC blank concentrations including trip and field blank, as well as OC-positive artifact on quartz filter, were tested. One of the ways for the reasonable estimation was using the intercept of the regression of OC concentrations against PM_{2.5}.¹⁵

Figure 1 shows the location of the 13 northeastern United States STN monitoring sites in which integrated OC blank concentrations were estimated. To examine the method for OC blank estimation, samples for which PM_{2.5} or OC mass concentrations were not available were excluded. The sample that showed extreme OC value on July 7, 2002, caused by a Canadian wildfire was excluded. Comparing colocated PM_{2.5} mass concentrations measured by STN and Federal Reference Method, outliers were censored before the regression analyses between STN PM_{2.5} and OC concentrations. To illustrate this approach, PM_{2.5} mass concentrations were compared with OC concentrations in Figure 2 for the four STN sites that have the highest or lowest intercepts of the 13 studied sites. The intercepts in PM_{2.5} regression against OC concentrations were then considered to be integrated OC blank concentrations that includes trip blank concentrations as well as positive sampling artifacts by adsorption of gaseous organic matter. The results for the 13 STN monitoring sites are summarized in Table 1. The estimated OC blank concentrations ranged from 1 to 2.4 μg/m³. The lower OC blank concentrations were found at rural sites (Whiteface,

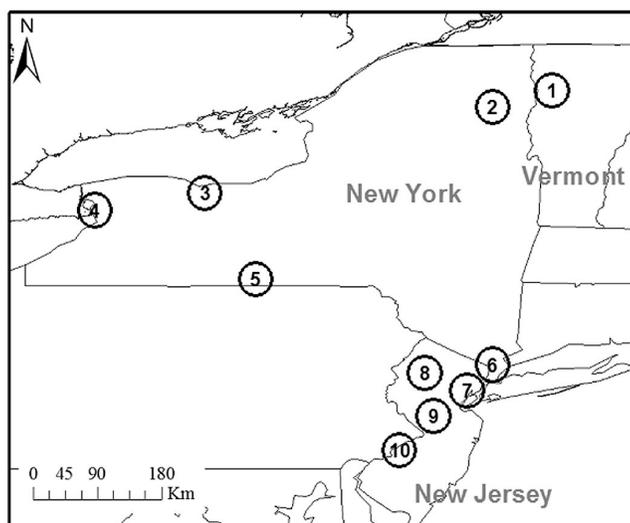


Figure 1. Location of EPA STN sites in NY, NJ, and VT. (1) Burlington; (2) Whiteface; (3) Rochester; (4) Buffalo; (5) Pinnacle State Park; (6) NY Botanical Garden, IS52, Canal St., and Queens College; (7) Elizabeth Lab.; (8) Chester; (9) New Brunswick; and (10) Camden.

NY, Pinnacle State Park, NY, and Chester, NJ), whereas the higher concentrations were observed in urban areas (Elizabeth Lab., NJ, Canal St., NY, and IS52, NY). The results suggest the influence of the positive artifact from the adsorption of gaseous organic matter by quartz filters.¹³ For the source apportionment study, the reported STN OC concentrations were blank corrected by subtracting the estimated OC blank concentrations.

MDL Values and Error Estimates

The application of PMF depends on the estimated uncertainties for each of the measured data. The uncertainty estimation based on the analytical uncertainties and laboratory MDL values provides a useful tool to decrease the weight of missing and below detection limit (BDL) data in these methods. Polissar et al.³ suggested a procedure for estimating uncertainties for the PMF study from the analyses of seven IMPROVE PM_{2.5} speciation datasets, in which data uncertainties and MDL values were well defined. In STN data, various instruments were used to analyze samples, and they produced different MDL values

Table 1. Summary of EPA STN sites in NY, NJ, and VT and estimated PM_{2.5} OC blank concentrations.

Monitoring Site	Sampler	Sampling Period	OC Blank, μg/m ³
Burlington, VT	SASS	12/2000-12/2003	1.83
Whiteface, NY	RPSPEC	5/2001-12/2003	1.02
Rochester, NY	RPSPEC	4/2001-12/2003	1.67
Buffalo, NY	RPSPEC	1/2002-12/2003	1.44
Pinnacle State Park, NY	RPSPEC	4/2001-12/2003	1.25
NY Botanical Garden, NY	SASS	4/2001-12/2003	1.41
IS52, NY	RPSPEC	12/2000-12/2003	1.85
Canal St., NY	SASS	8/2002-12/2003	2.44
Queens College, NY	RPSPEC	4/2001-12/2003	1.57
Chester, NJ	SASS	1/2002-12/2002	1.35
New Brunswick, NJ	SASS	1/2002-12/2002	1.49
Camden, NJ	SASS	1/2002-12/2002	1.65
Elizabeth Lab., NJ	SASS	1/2002-12/2002	2.19

Notes: SASS = Spiral Aerosol Speciation Sampler; RPSPEC = Rupprecht & Patashnick Speciation Sampler.

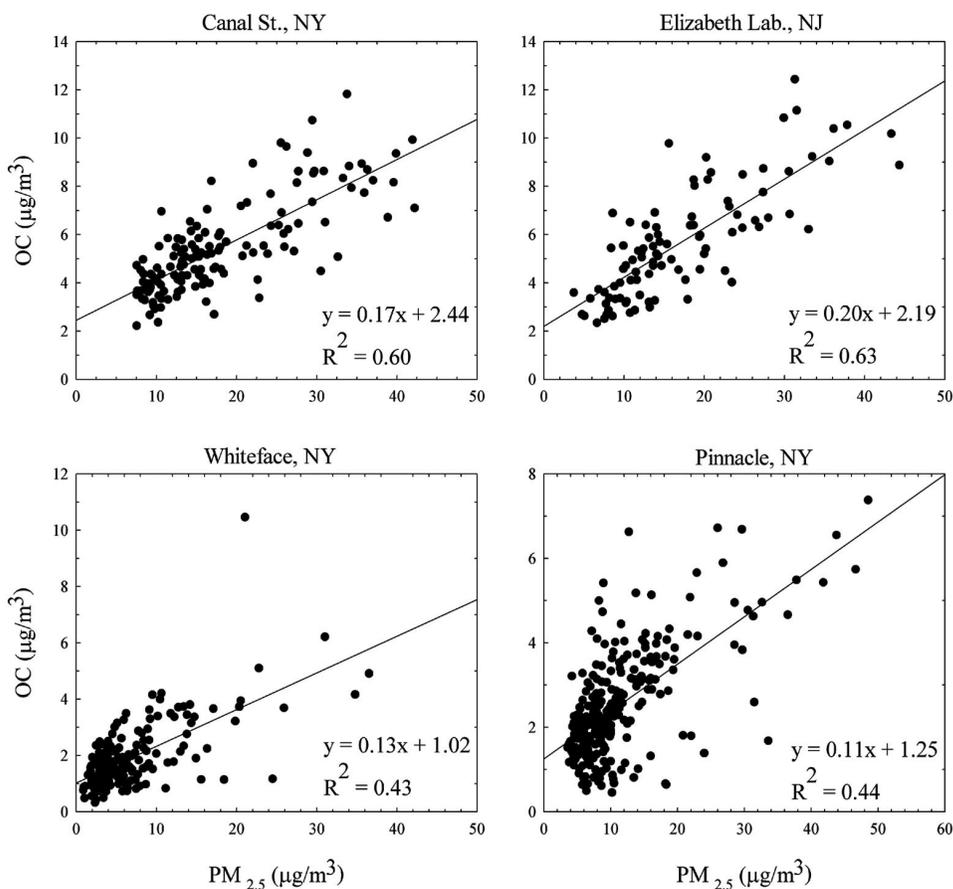


Figure 2. PM_{2.5} mass concentrations vs. OC concentrations for four STN sites.

and analytical uncertainties.¹⁴ A comprehensive set of MDL values and error structures that can be used for source apportionment studies need to be estimated, because before July 2003, the STN data were not accompanied by MDL values and uncertainties. It is not possible to identify which instrument was used for the analysis of any particular sample so that it is not possible to assign its particular MDL values and uncertainties for that sample.

The filter mass loading MDL values for OC, EC, anions, and cations (Table 2) are the same for the multiple instruments at RTI.¹⁴ Table 3 presents the filter mass loading MDL values for the 48 elements analyzed by the five XRF spectrometers (Chester770, Chester771, Cooper,

Table 2. Laboratory MDLs for carbons, anions, and cations (RTI, 2004).

	MDLs (μg/filter)
OC, EC	2.352
SO ₄ ²⁻	0.120
NH ₄ ⁺	0.160
NO ₃ ⁻	0.084
K ⁺	0.134
Na ⁺	0.290

Table 3. Laboratory MDLs for the elements analyzed by X-ray fluorescent methods (RTI, 2004).

	MDLs ($\mu\text{g}/\text{filter}$)						
	Chester 770	Chester 771	Cooper	RTI XRF1	RTI XRF2	Min. ($\mu\text{g}/\text{filter}$)	Ave. ($\mu\text{g}/\text{filter}$)
Al	0.157	0.163	0.035	0.219	0.219	0.035	0.159
Sb	0.213	0.158	0.164	0.267	0.267	0.158	0.214
As	0.036	0.037	0.005	0.014	0.014	0.005	0.021
Ba	0.850	0.601	0.022	0.099	0.099	0.022	0.334
Br	0.029	0.031	0.005	0.011	0.011	0.005	0.017
Cd	0.152	0.095	0.057	0.100	0.100	0.057	0.101
Ca	0.050	0.044	0.014	0.071	0.071	0.014	0.050
Ce	1.242	1.113	0.020	0.080	0.080	0.020	0.507
Cs	0.533	0.436	0.021	0.097	0.097	0.021	0.237
Cl	0.084	0.132	0.030	0.078	0.078	0.030	0.080
Cr	0.023	0.019	0.007	0.019	0.019	0.007	0.017
Co	0.020	0.018	0.005	0.014	0.014	0.005	0.014
Cu	0.019	0.019	0.006	0.024	0.024	0.006	0.018
Eu	0.162	0.084	0.017	0.036	0.036	0.017	0.067
Ga	0.048	0.071	0.009	0.026	0.026	0.009	0.036
Au	0.072	0.127	0.011	0.037	0.037	0.011	0.057
Hf	0.378	0.155	0.016	0.257	0.257	0.016	0.213
In	0.163	0.108	0.081	0.135	0.135	0.081	0.124
Ir	0.086	0.165	0.020	0.040	0.040	0.020	0.070
Fe	0.028	0.024	0.005	0.020	0.020	0.005	0.019
La	1.004	0.795	0.017	0.083	0.083	0.017	0.396
Pb	0.079	0.085	0.010	0.027	0.027	0.010	0.046
Mg	0.266	0.427	0.077	0.175	0.175	0.077	0.224
Mn	0.033	0.026	0.009	0.015	0.015	0.009	0.020
Hg	0.063	0.065	0.010	0.033	0.033	0.010	0.041
Mo	0.069	0.085	0.019	0.082	0.082	0.019	0.067
Ni	0.018	0.015	0.005	0.016	0.016	0.005	0.014
Nb	0.060	0.067	0.015	0.033	0.033	0.015	0.042
P	0.090	0.056	0.016	0.102	0.102	0.016	0.073
K	0.049	0.062	0.018	0.106	0.106	0.018	0.068
Rb	0.031	0.031	0.008	0.014	0.014	0.008	0.020
Sm	0.089	0.068	0.020	0.041	0.041	0.020	0.052
Sc	0.035	0.031	0.009	0.000	0.000	0.009	0.015
Se	0.031	0.033	0.005	0.025	0.025	0.005	0.024
Si	0.109	0.098	0.039	0.178	0.178	0.039	0.120
Ag	0.151	0.091	0.045	0.082	0.082	0.045	0.090
Na	0.738	1.580	0.419	0.526	0.526	0.419	0.758
Sr	0.036	0.036	0.010	0.017	0.017	0.010	0.023
S	0.095	0.204	0.014	0.085	0.085	0.014	0.097
Ta	0.282	0.255	0.016	0.074	0.074	0.016	0.140
Tb	0.109	0.073	0.014	0.042	0.042	0.014	0.056
Sn	0.258	0.131	0.114	0.192	0.192	0.114	0.177
Ti	0.030	0.034	0.010	0.048	0.048	0.010	0.034
V	0.022	0.021	0.008	0.031	0.031	0.008	0.023
W	0.199	0.208	0.017	0.059	0.059	0.017	0.108
Y	0.044	0.044	0.011	0.021	0.021	0.011	0.028
Zn	0.021	0.017	0.008	0.025	0.025	0.008	0.019
Zr	0.052	0.054	0.014	0.027	0.027	0.014	0.035

RTI1, and RTI2) operated by three participating laboratories.¹⁴ To investigate appropriate MDL values for PMF analyses, the min and average MDL values among five MDL values were tested using STN data measured in the Burlington, VT, site.

A limited set of the XRF analytical uncertainties for 13 eastern STN sites for samples collected between March 2001 and November 2003 were acquired from EPA. The reported analytical uncertainties for sulfur, silicon (Si), potassium (K), and iron (Fe) from the five instruments in three laboratories were compared in Figure 3. Various species, instruments, and laboratories show different analytical uncertainty structures.

As can be seen from Figure 3, the uncertainties are given as fractions of measured mass concentrations. To develop a comprehensive set of errors that could be used for PMF studies across the STN, a general fractional error was estimated by comparing the available measured concentrations and their associated uncertainties. To generate the error structures, the fractional errors that are estimated as a fraction of the measured concentrations are chosen to encompass most of the reported uncertainties as shown by the lines in Figure 3 and to provide the most reasonable PMF solution. For the species for which analytical uncertainties were not available (i.e. OC, EC, SO_4^{2-} , NH_4^+ , NO_3^- , sodium [Na^+], and potassium [K^+]), the fractional errors between 5% and 30% were tested, and the fractions that started to provide the interpretable PMF solution were chosen. The specific values for each element are shown in Table 4. Thus, based on the studies of Polissar et al.,³ the error structures (s_{ij}) were calculated using the following equation:

$$s_{ij} = [\text{MDL}]/3 + k \times x_{ij} \quad (1)$$

where x_{ij} is the j th species concentration measured in the i th sample, and the values of k are given in Table 4. The minimum or mean MDL values could be used, and the specific choice of MDL values need to be determined by the studies described below.

Multivariate Receptor Modeling

An ambient $\text{PM}_{2.5}$ compositional data set of 24-hr integrated samples collected at a STN site in Burlington, VT, were analyzed through the application of PMF to examine the estimated error structures and to investigate the appropriate MDL values. The receptor modeling problem can be expressed in terms of the contribution from p independent sources to all of the chemical species in a given sample as follows:^{16,17}

$$x_{ij} = \sum_{s=1}^p g_{is}f_{sj} + e_{ij} \quad (2)$$

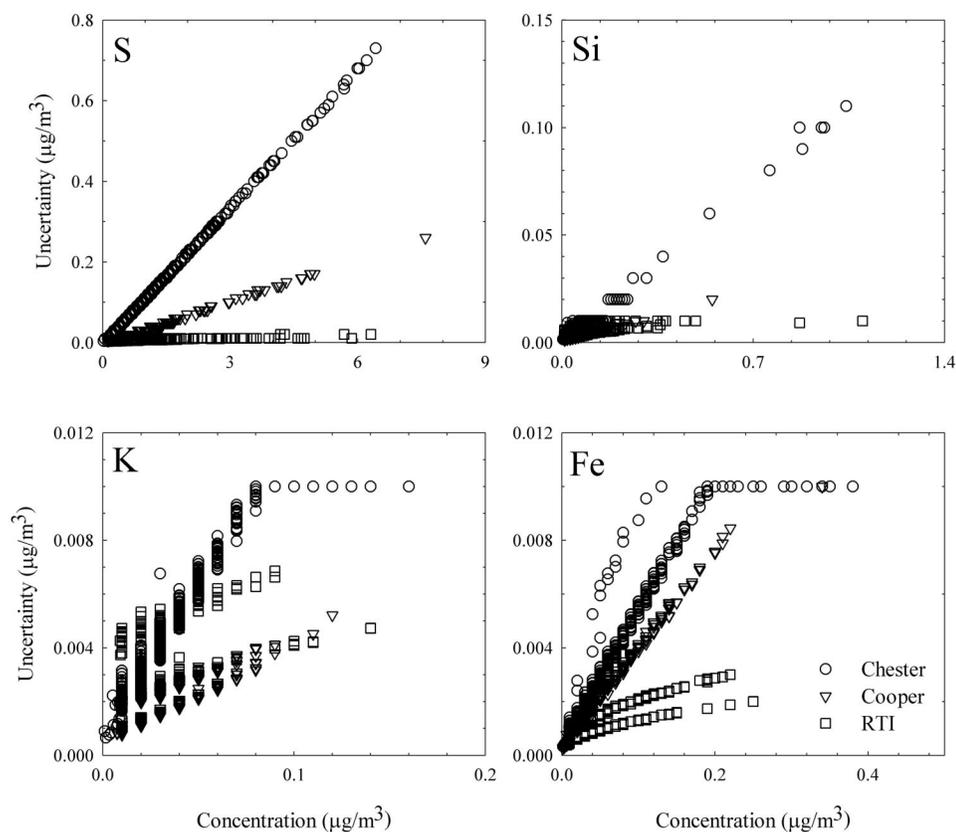


Figure 3. The comparison between measured concentrations and associated analytical uncertainties.

where g_{is} is the particulate mass concentration from the s th source contributing to the i th sample, f_{sj} is the j th species mass fraction from the s th source, e_{ij} is residual associated with the j th species concentration measured in the i th sample, and p is the total number of independent sources. PMF provides a solution that minimizes an object function, $Q(E)$, based on uncertainties for each observation.^{3,18}

$$Q(E) = \sum_{i=1}^n \sum_{j=1}^m \left[\frac{x_{ij} - \sum_{s=1}^p g_{is} f_{sj}}{u_{ij}} \right]^2 \quad (3)$$

where u_{ij} is an uncertainty estimate in the j th constituent measured in the i th sample.

There are an infinite number of possible combinations of source contribution and profile matrices to the multivariate receptor modeling problem because of the free rotation of matrices.¹⁹ PMF uses non-negativity constraints on the factors to decrease rotational ambiguity. Also, the parameter FPEAK and the matrix FKEY are used to control the rotations.^{20,21} By setting a non-zero value of FPEAK, the routine is forced to add one source contribution vector to another and subtract the corresponding source profile factors from each other and thereby yield

more physically realistic solutions. PMF was run with different FPEAK values to determine the range within which the scaled residuals remain relatively constant.^{21,23} The optimal solution should lie in this FPEAK range. In this way, subjective bias was reduced to a large extent. External information can be imposed on the solution to control the rotation. If specific species in the source profiles are known to be zero, then it is possible to pull down those values toward lower concentration through appropriate settings of FKEY resulting in the most interpretable source profiles. Each element of the FKEY matrix controls the pulling-down of the corresponding element in the source profile matrix by setting a nonzero integer value in FKEY matrix.²⁰

Based on the studies of Polissar et al.,³ the measured concentrations below the MDL values were replaced by half of the MDL values, and their uncertainties were set at 5/6 of the MDL values. Missing concentrations were replaced by the geometric mean of the concentrations, and their accompanying uncertainties were set at four times this geometric mean concentration.

For the application of PMF to STN data, only samples for which $PM_{2.5}$ or OC mass concentrations were not available were excluded from data set measured in Burlington, VT. A summary of original $PM_{2.5}$ speciation data, MDL values, and signal-to-noise ratios are provided in Table 5. To obtain reasonable model fit, the Canadian wildfire sample on July 7, 2002, in which $PM_{2.5}$ and OC mass concentrations were unusually high (i.e., 7 times and 25 times geometric mean, respectively) was excluded in the source apportionment study. Overall, 7% of the original data was excluded from this study. XRF S and IC SO_4^{2-} showed excellent correlations ($slope = 3.1$, $r^2 = 0.98$), so it is reasonable to exclude XRF S from the analysis to prevent double counting of mass concentrations. Also, IC Na^+ and IC K^+ were chosen because of the higher analytical precision compared with XRF Na and XRF K. Chemical species that have below MDL values >90% (i.e., antimony, indium, samarium, and terbium) were excluded. Thus, a total of 323 samples collected between

Table 4. Estimated fractional uncertainties for EPA STN data.

Species	Uncertainty (%)	Species	Uncertainty (%)
OC	7	Pb	5
EC	7	Mg	5
SO ₄ ²⁻	7	Mn	5
NH ₄ ⁺	7	Hg	5
NO ₃ ⁻	7	Mo	5
K ⁺	7	Ni	5
Na ⁺	7	Nb	5
Al	10	P	10
Sb	5	K	10
As	20	Rb	5
Ba	5	Sm	5
Br	5	Sc	5
Cd	5	Se	5
Ca	11	Si	10
Ce	5	Ag	5
Cs	5	Na	10
Cl	10	Sr	5
Cr	5	S	11
Co	10	Ta	5
Cu	5	Tb	5
Eu	5	Sn	5
Ga	5	Ti	5
Au	5	V	5
Hf	5	W	5
In	5	Y	5
Ir	5	Zn	5
Fe	5	Zr	5
La	5	—	—

December 2000 and December 2003 and 49 species including PM_{2.5} mass concentrations were used. Species that have signal-to-noise ratios between 0.2 and 2 were considered weak variables (i.e., EC, cadmium, europium, gallium, cobalt, iridium, molybdenum, niobium, rubidium, scandium, silver, strontium, tin, yttrium, and zirconium), and their estimated uncertainties were increased by a factor of five to reduce their weight in the solution as recommended by Paatero and Hopke.²² The contaminated Na⁺ collected between October 2001 and January 2002 were reported,⁸ and they were treated as missing data.

In this study, the measured PM_{2.5} mass concentration was included as an independent variable in the PMF modeling to directly obtain the mass apportionment without the usual multiple regression. The utilization of PM_{2.5} mass concentration as a variable is specified in detail in Kim et al.²³

Finally, to obtain physically reasonable PMF solution, it was necessary to test different (min/average) MDL values, different numbers (5–11) of sources, and different FPEAK values (–1–1) with the final choice based on the

evaluation of the resulting source profiles, as well as the quality of the species fits. The global optimum of the PMF solutions were tested by using multiple random starts for the initial values used in the iterative fitting process.

RESULTS AND DISCUSSION

In a variety of source number solutions, FPEAK values, and FKEY matrices, the eight-source model, a value of FPEAK = 0, and min MDL values provided the most physically reasonable source profiles. In the seven-source model, the copper smelting was not deduced. In the nine-source model, the additional source was not interpretable. The source profiles of the best solution are presented in Figure 4, and the corresponding source contributions are shown in Figure 5. These results will be discussed below, but first, the results of the study of the choice of MDL values and PMF parameters will be assessed.

It should be noted that because actual point-by-point errors are not being used, it is not possible to use the comparison of the $Q(E)$ values with the theoretical $Q(E)$ values as a means of estimating the likely number of factors.²¹ The relative changes in the $Q(E)$ values may still be useful, but the absolute values may no longer be relevant, because the choice of estimated MDL values and errors precludes accurate calculation of the actual $Q(E)$ values for the problem.

In the results of the 7–10-source models, several trace elements (i.e., barium (Ba), tantalum (Ta), lanthanum (La), and zinc (Zn)) had unrealistically high concentrations. These elements were down-weighted by increasing their estimated uncertainties five times. To pull-down concentrations in the source profiles that appear high toward zero, a FKEY matrix was set. In the FKEY matrix, values of all elements were set to zero, except the following: a value of 6 for SO₄²⁻ in gasoline vehicle and diesel emissions. The use of FKEY is indicated by arrowheads above the source profile values that were adjusted through the use of FKEY.

In the study of the MDL values, the diesel emission profile compared with the results in Figure 6 showed high SO₄²⁻, NO₃⁻, Ba, Cerium, and La concentrations when the average MDL values were used. The average MDL values increased the estimated uncertainties in the PMF analysis and resulted in several unreasonably high species in source profiles. Because this profile of the diesel emissions was not interpretable, it seems appropriate to report the eight-source model results with min MDL values.

The average source contributions of each source to the PM_{2.5} mass concentrations are summarized in Table 6. A comparison of the measured PM_{2.5} concentrations with the summations of species included in this study showed a slope 0.67 ± 0.02 ($r^2 = 0.82$). In Figure 7, a comparison of the reconstructed PM_{2.5} contributions from all of the

Table 5. Summary of PM_{2.5} and 55 species mass concentrations in Burlington, VT.

Species	Arithmetic Mean, ng/m ³	Number of BDL Values, ^a %	Number of Missing Values, %	Minimum MDL, ^b ng/m ³	Average MDL, ng/m ³	Signal-to-Noise ^c
PM _{2.5}	11,043	0	0	746	NA	19
NH ₄ ⁺	1,106	1.85	0.62	17	NA	96
NO ₃ ⁻	1,388	0.31	0.62	8.7	NA	247
K ⁺	70	58	0.62	14	NA	3.9
Na ⁺	218	15	1.5	30	NA	12
SO ₄ ²⁻	2,591	0	0.62	12	NA	308
OC	2,193	15	0	244	NA	12
EC	392	17	0	244	NA	1.8
Al	24	47	0	3.6	16	9.3
Sb	9.2	93	0	16	22	0.6
As	1.5	53	0	0.49	2.2	2.9
Ba	34	26	0	2.3	35	16
Br	2.2	28	0	0.57	1.8	4.5
Cd	4.5	87	0	5.9	10	0.7
Ca	34	1.9	0	1.4	5.2	30
Ce	24	55	0	2.1	53	12
Cs	11	57	0	2.2	25	4.8
Cl	18	60	0	3.1	8.3	11
Cr	2.5	33	0	0.78	1.8	8.3
Co	0.59	83	0	0.49	1.5	0.9
Cu	3.1	23	0	0.67	1.9	6.0
Eu	3.1	88	0	1.8	7.0	1.2
Ga	1.2	84	0	0.94	3.7	1.1
Au	2.9	49	0	1.2	5.9	2.5
Hf	8.9	75	0	1.6	22	4.2
In	4.7	94	0	8.4	13	0.5
Ir	3.2	67	0	2.1	7.3	1.4
Fe	40	0.93	0	0.47	2.0	114
La	20	56	0	1.7	41	11
Pb	3.7	32	0	1.0	4.7	4.0
Mg	23	79	0	8.0	23	2.6
Mn	2.2	36	0	1.0	2.0	2.3
Hg	2.2	57	0	1.0	4.2	2.0
Mo	2.5	77	0	1.9	7.0	1.2
Ni	2.0	27	0	0.52	1.5	5.4
Nb	1.7	84	0	1.5	4.3	0.8
P	6.8	74	0	1.6	7.6	4.6
K	43	1.2	0	1.9	7.1	29
Rb	0.88	79	0	0.83	2.0	0.9
Sm	2.0	94	0	2.1	5.4	0.6
Sc	1.1	86	0	0.89	1.5	0.9
Se	1.2	55	0	0.48	2.5	2.5
Si	65	1.5	0	4.1	12	22
Ag	4.8	73	0	4.7	9.4	1.0
Na	133	55	0	43	79	2.8
Sr	1.5	72	0	1.0	2.4	1.2
S	825	0.31	0	1.5	10	823
Ta	15	46	0	1.6	15	10
Tb	1.5	93	0	1.5	5.8	0.7
Sn	14	56	0	12	18	1.2
Ti	4.4	19	0	1.0	3.5	5.2

Table 5. Continued.

Species	Arithmetic Mean, ng/m ³	Number of BDL Values, ^a %	Number of Missing Values, %	Minimum MDL, ^b ng/m ³	Average MDL, ng/m ³	Signal-to-Noise ^c
V	1.9	43	0	0.80	2.3	2.5
W	5.8	56	0	1.8	11	3.2
Y	1.1	85	0	1.2	2.9	0.8
Zn	8.6	11	0	0.80	2.0	14
Zr	1.5	82	0	1.4	3.6	0.9

Notes: NA = not available; ^aBelow laboratory MDL values based on minimum MDL values; ^bLaboratory MDL; ^cRatio of signal-to-noise based on minimum MDL values.

sources with measured STN PM_{2.5} concentrations shows that the PMF-resolved sources using estimated error structure effectively reproduce the measured values and account for most of the variation in the PM_{2.5} concentrations (*slope* = 0.78 ± 0.02 and *r*² = 0.81).

Secondary sulfate aerosol has a high concentration of SO₄²⁻ and NH₄⁺ and account for 41% of the PM_{2.5} mass concentration. The averaged seasonal contributions from each source are presented in Figure 8 (summer: April–September; winter: October–March). As shown in Figures 5 and 8, the secondary sulfate aerosol shows strong seasonal variation with higher concentrations in summer, when the photochemical activity is highest. The earlier results^{24,25} indicated that the sulfate was the result of emissions from coal-fired power plants.

When compared with the studies based on IMPROVE data,²⁴ the selenium data was inadequate to permit the winter-high secondary sulfate aerosol to be extracted in this study. Secondary sulfate aerosol profiles typically

include OC that becomes associated with the secondary sulfate aerosol,²⁶ and this association is consistent with previous studies that observed similar profiles.^{4,24,25} The secondary sulfate aerosol had the highest source contribution to PM_{2.5} mass concentrations (4 μg/m³). In the previous study using 7-year measurements between 1988 and 1995 from Underhill, VT, summer and winter-secondary sulfate aerosols contributed 4.9 μg/m³ to the PM_{2.5} concentrations.²⁴

The secondary nitrate aerosol is represented by its high concentration of NO₃⁻ and NH₄⁺. It accounts for 20% of the PM_{2.5} mass concentration. The average contribution of this source to the PM_{2.5} mass concentration was 2 μg/m³. This value is similar to the secondary nitrate aerosol contributions identified in Washington, DC (1.6 μg/m³).⁶ This source was not identified in previous northern Vermont aerosol study, because NO₃⁻ was not measured in the study.^{4,24} The secondary nitrate aerosol has seasonal variation with maxima in winter as shown in

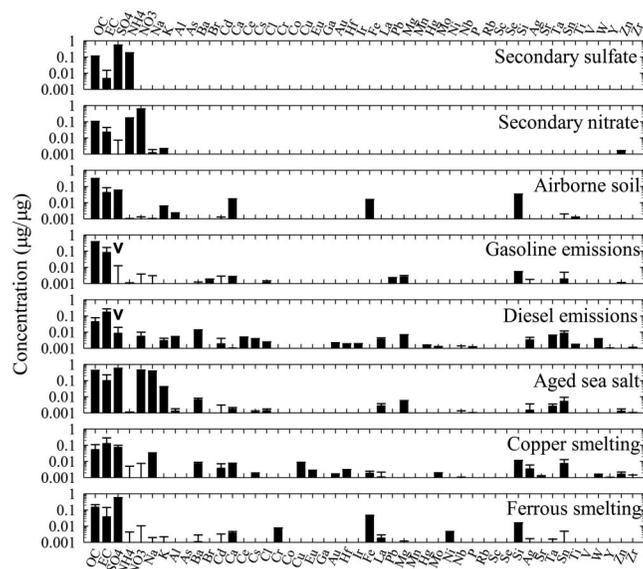


Figure 4. Source profiles deduced from PM_{2.5} samples using min method detection limit values (prediction ± standard deviation). The species that were pulled down by FKEY matrix is indicated by arrowheads.

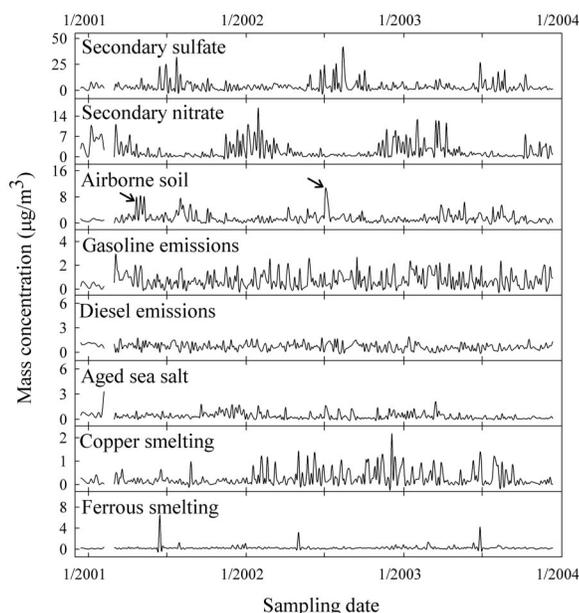


Figure 5. Time series plot of source contributions deduced using min method detection limit values. The dust storm impacts are indicated by arrows.

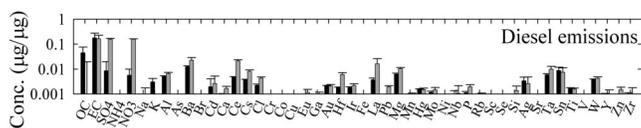


Figure 6. The comparison of diesel emissions profile deduced from Burlington, VT, data. The black and gray bars are derived using minimum and averaged MDL values, respectively.

Figures 5 and 8. These peaks in winter indicate that low temperature and high relative humidity help the formation of secondary nitrate aerosols. This seasonal variation is consistent with previous studies for Atlanta, GA,⁴ and northeastern United States.^{6,7}

Airborne soil is represented by Si, Fe, aluminum, and calcium^{27,28} contributing 15% (1.5 µg/m³) to the PM_{2.5} mass concentration. Crustal particles could be contributed by wind-blown soil dust and resuspended by road traffic. The airborne soil shows seasonal variation with higher concentrations in summer. The elevated airborne soil contributions on April 22, 2001, and July 4, 2002, as indicated in Figure 5, were likely to be caused by Asian dust storm and Saharan dust storm, respectively, based on the previous backward trajectory analyses.^{6,7,29}

Gasoline vehicle and diesel emissions are represented by high OC and EC, of which the abundances differ between these sources.²⁷ Gasoline vehicle emissions have high concentration of the OC. In contrast, diesel emissions were tentatively identified on the basis of the high concentration of EC. The averaged PM_{2.5} mass contributions from gasoline and diesel emissions were 0.73 and 0.69 µg/m³ accounting for 7.5 and 7.1% of the PM_{2.5} mass concentration, respectively. The fraction of mass contributions from gasoline and diesel emissions (14.6%) is consistent with 16% from the fraction of gasoline and diesel emissions contributing PM_{2.5} in rural Brigantine, NJ.⁷ These sources do not have a clear seasonal trend. Gasoline vehicle and diesel emissions were not identified in previous northern Vermont aerosol studies²⁴ because of the lack of OC and EC data.

Table 6. Average source contributions (%) to PM_{2.5} mass concentrations in Burlington, VT.

	Average Source Contribution (Standard Error)
Secondary sulfate	40.9 (3.1)
Secondary nitrate	20.2 (1.5)
Airborne soil	15.1 (0.8)
Gasoline vehicle emissions	7.5 (0.3)
Diesel emissions	7.1 (0.2)
Aged sea salt	3.9 (0.2)
Copper smelter	2.7 (0.2)
Ferrous smelting	2.6 (0.3)

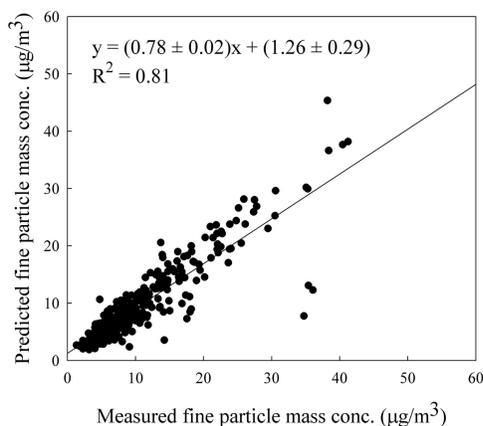


Figure 7. Measured vs. predicted PM_{2.5} mass concentration.

Aged sea salt is characterized by its high concentration of Na, SO₄²⁻, and NO₃⁻. The lack of chlorine in the profile is caused by chloride displacement by acidic gases. This source accounts for 4% (0.4 µg/m³) of the PM_{2.5} mass concentration. The averaged contribution of this source in Washington, DC,⁶ was 0.4 µg/m³, which is similar to the results of this study. In contrast, a previous northern Vermont aerosol study attributed 0.05 µg/m³ of the PM_{2.5} mass concentration to this source.²⁴ This source shows a winter-high seasonal pattern. The reported Na⁺ concentration has high peaks between October 2001 and January 2002 because of the filter contamination from the cleaning procedure for centrifuge tubes.⁸ Those Na⁺ values were down-weighted in this study to reduce their effects in the solution. However, there is the potential that this source contribution is still contaminated by some degree of artifact.

Two metal processing sources that are characterized by EC, Copper, Fe, and Si were resolved.^{29,30} These sources do not have a strong seasonal pattern. These sources accounts for 5% (0.5 µg/m³) of the PM_{2.5} mass concentration. The previous Vermont aerosol study attributed 0.4 µg/m³ of the PM_{2.5} mass concentration to Canadian Cu, Mn, and smelter sources.²⁴

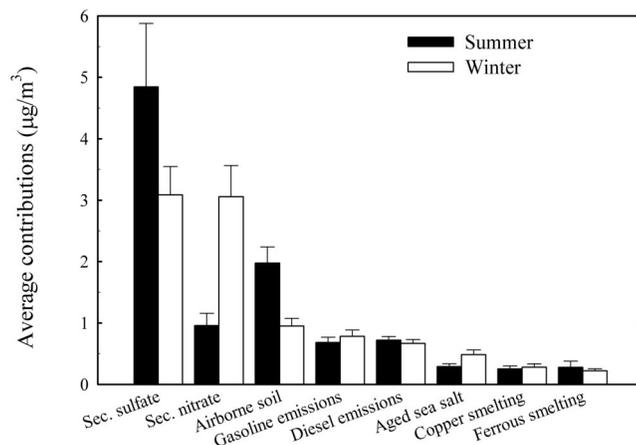


Figure 8. The seasonal comparison of source contributions to PM_{2.5} mass concentration (mean ± 95% confidence interval).

CONCLUSIONS

Integrated PM_{2.5} compositional data from samples collected at 13 STN monitoring site in the northeastern United States were used to estimate particulate OC blank concentrations using the intercept in PM_{2.5} regression against OC concentrations. The higher OC blank concentrations in urban areas suggest the influence of the positive artifact from the adsorption of gaseous organic materials by quartz filters. A comprehensive set of error structures for the STN samples was estimated, and appropriate MDL values were chosen. They were successfully examined through the application of PMF to STN data measured in Burlington, VT, and identified eight PM_{2.5} sources. Although this study is based on STN data collected in the northeastern United States, the results of this study will provide the appropriate OC blank value estimation, error structures, and MDL values for STN data source apportionment studies at a variety of sites in the network.

ACKNOWLEDGMENTS

This work was supported by the New York State Energy Research and Development Authority under agreement No. 7919, Health Effects Institute under agreement 4706-RFPA03-4/04-2, and the EPA Science to Achieve Results (STAR) program under grant R831078. Although the research described in this article has been funded in part by EPA, it has not been subjected to the agency's required peer and policy review and, therefore, does not necessarily reflect the views of the agency, and no official endorsement should be inferred.

REFERENCES

- Paatero, P. Least Square Formulation of Robust Non-Negative Factor Analysis; *Chemometrics Intelligent Lab. Systems* **1997**, *37*, 23-35.
- Malm, W.C.; Sisler, J.F.; Huffman, D.; Eldred, R.A.; Cahill, T.A. Spatial and Seasonal Trends in Particle Concentration and Optical Extinction in the United States; *J. Geophys. Res.* **1994**, *99*, 1347-1370.
- Polissar, A.V.; Hopke, P.K.; Paatero, P.; Malm, W.C.; Sisler, J.F. Atmospheric Aerosol over Alaska 2. Elemental Composition and Sources; *J. Geophys. Res.* **1998**, *103*, 19045-19057.
- Song, X.H.; Polissar, A.V.; Hopke, P.K. Source of Fine Particle Composition in the Northeastern U.S.; *Atmos. Environ.* **2001**, *35*, 5277-5286.
- Maykut, N.N.; Lewtas, J.; Kim, E.; Larson, T.V. Source Apportionment of PM_{2.5} at an Urban IMPROVE Site in Seattle, WA.; *Environ. Sci. Technol.* **2003**, *37*, 5135-5142.
- Kim, E.; Hopke, P.K. Source Apportionment of Fine Particles at Washington, DC Utilizing Temperature Resolved Carbon Fractions; *J. Air & Waste Manage. Assoc.* **2004**, *54*, 773-785.
- Kim, E.; Hopke, P.K. Improving Source Identification of Fine Particles in a Rural Northeastern US Area Utilizing Temperature Resolved Carbon Fractions; *J. Geophys. Res.* **2004**, *109*, D09204.
- Extended Semi-Annual Data Summary Report for the Chemical Speciation of PM_{2.5} Filter Samples Project. RTI/07565/06DS, Research Triangle Institute, Research Triangle Park, NC, 2004.
- Koutrakis, P.; Wolfson, J.M.; Slater, J.L.; Brauer, M.; Spengler, J.D. Evaluation of an Annular Denuder/Filter Pack System to Collect Acidic Aerosols and Gases; *Environ. Sci. Technol.* **1988**, *22*, 1463-1468.
- Hering, S.; Cass, G. The Magnitude of Bias in the Measurement of PM_{2.5} Regression Arising from Volatilization of Particulate Nitrate from Teflon Filters; *J. Air & Waste Manage. Assoc.* **1999**, *49*, 725-733.
- Birch, M.E.; Cary, R.A. Elemental Carbon-Based Method for Monitoring Occupational Exposures to Particulate Diesel Exhaust; *Aerosol Sci. Technol.* **1996**, *25*, 221-241.
- Gundel, L.A.; Lee, V.C.; Mahanama, K.R.R.; Stevens, R.K.; Rau, J.A. Direct Determination of the Phase Distributions of Semi-Volatile Polycyclic Aromatic Hydrocarbons Using Annular Denuders; *Atmos. Environ.* **1995**, *29*, 1719-1733.
- Pankow, J.F.; Mader, B.T. Gas/Solid Partitioning of Semivolatile Organic Compounds (SOCs) to Air Filters. 3. An Analysis of Gas Adsorption Artifacts in Measurements of Atmospheric SOCs and Organic Carbon (OC) When Using Teflon Membrane Filters and Quartz Fiber Filters; *Environ. Sci. Technol.* **2001**, *35*, 3422-3432.
- Quality Assurance Projects Plan Chemical Speciation of PM_{2.5} Filter Samples; Research Triangle Park, NC, 2004; available at <http://www.epa.gov/ttn/amtic/files/ambient/pm25/spec/rtiqap.pdf> (accessed June 1, 2005).
- Tolocka, M.P.; Solomon, P.A.; Mitchell, W.; Norris, G.A.; Gemmill, D.B.; Wiener, R.W.; Vanderpool, R.W.; Homolya, J.B.; Rice, J. East Versus West in the U.S.: Chemical Characteristics of PM_{2.5} During the Winter of 1999; *Aerosol Sci. Technol.* **2001**, *34*, 88-96.
- Miller, M.S.; Friedlander, S.K.; Hidy, G. M. A Chemical Element Balance for the Pasadena Aerosol; *J. Colloid Interface Sci.* **1972**, *39*, 165-176.
- Hopke, P.K. *Receptor Modeling in Environmental Chemistry*. John Wiley & Sons: New York, 1985.
- Paatero, P. Least Square Formulation of Robust Non-Negative Factor Analysis; *Chemometrics Intelligent Lab. Systems* **1997**, *37*, 23-35.
- Henry, R.C. Current Factor Analysis Models Are Ill-Posed; *Atmos. Environ.* **1987**, *21*, 1815-1820.
- Lee, E.; Chun, C.K.; Paatero, P. Application of Positive Matrix Factorization in Source Apportionment of Particulate Pollutants; *Atmos. Environ.* **1999**, *33*, 3201-3212.
- Paatero, P.; Hopke, P.K.; Song, X.H.; Ramadan, Z. Understanding and Controlling Rotations in Factor Analytic Models; *Chemometrics Intelligent Lab. Systems* **2002**, *60*, 253-264.
- Paatero, P.; Hopke, P.K. Discarding or Downweighting High-Noise Variables in Factor Analytic Models; *Anal. Chim. Acta* **2003**, *490*, 277-289.
- Kim, E.; Hopke, P.K.; Edgerton, E.S. Source Identification of Atlanta Aerosol by Positive Matrix Factorization; *J. Air & Waste Manage. Assoc.* **2003**, *53*, 731-739.
- Polissar, A.V.; Hopke, P.K.; Poirot, R.L. Atmospheric Aerosol Over Vermont: Chemical Composition and Sources; *Environ. Sci. Technol.* **2001**, *35*, 4604-4621.
- Poirot, R.L.; Wishinski, P.R.; Hopke, P.K.; Polissar, A.V. Comparative Application of Multiple Receptor Methods to Identify Aerosol Sources in Northern Vermont; *Environ. Sci. Technol.* **2001**, *35*, 4622-4636.
- Liu, D.; Wenzel, R.J.; Prather, K.A. Aerosol Time-of-Flight Mass Spectrometry during the Atlanta Supersite Experiment: 1. Measurements; *J. Geophys. Res.* **2003**, *108*, D7.
- Watson, J.G.; Chow, J.C.; Houck, J.E. PM_{2.5} Chemical Source Profiles for Vehicle Exhaust, Vegetative Burning, Geological Material, and Coal Burning in Northwestern Colorado During 1995; *Chemosphere* **2001**, *43*, 1141-1151.
- Watson, J.G.; Chow, J.C. Source Characterization of Major Emission Sources in the Imperial and Mexicali Valleys Along the US/Mexico Border; *Sci. Total Environ.* **2001**, *276*, 33-47.
- Small, M.; Germani, M.S.; Small, A.M.; Zoller, W.H.; Moyers, J.L. Airborne Plume Study of Emissions from the Processing of Copper Ores in Southeastern Arizona; *Environ. Sci. Technol.* **1981**, *15*, 293-299.
- SPECIATE version 3.2. U.S. Environmental Protection Agency: Research Triangle Park, NC. 2002.

About the Authors

Eugene Kim is a research assistant professor with the Department of Civil and Environmental Engineering, Clarkson University. Philip K. Hopke is the Baynard D. Clarkson Distinguished Professor, and Youjun Qin is a research associate with the Department of Chemical Engineering, Clarkson University. Address correspondence to: Philip K. Hopke, Department of Chemical Engineering, Clarkson University, Box 5708, Potsdam, NY 13699; phone: +1-315-268-3861; fax: +1-315-268-4410; e-mail: hopkepk@clarkson.edu.