

## Comparison of Atmospheric Aerosols between Two Sites over Golden Triangle of Texas

Chiou, P.<sup>1\*</sup>, Tang, W.<sup>2</sup>, Lin, C. J.<sup>3,4</sup>, Chu, H. W.<sup>5</sup> and Ho, T. C.<sup>2</sup>

<sup>1</sup>Department of Mathematics, Lamar University, Box 10047, Beaumont, TX 77710, USA

<sup>2</sup>Department of Chemical Engineering, Lamar University, Box 10053, Beaumont, TX 77710, USA

<sup>3</sup>Department of Civil Engineering, Lamar University, Box 10024, Beaumont, TX 77710, USA

<sup>4</sup>School of Environ. Sci. and Eng., South China University of Technology, Guangzhou City, China

<sup>5</sup>Department of Mechanical Engineering, Lamar University, Box 10028, Beaumont, TX 77710, USA

Received 12 July 2008;

Revised 15 Jan. 2009;

Accepted 25 Jan. 2009

**ABSTRACT:** Speciated samples of PM<sub>2.5</sub> were collected at the Hamshire and Orange sites in Golden Triangle of Texas by US EPA (Environmental Protection Agency) from July of 2003 to August of 2005. A total of 269 samples for the Hamshire site and 293 samples for the Orange site with 52 species were measured; however, 22 species were excluded because of too many below-detection-limit data. Among the 22 species excluded, 20 species are common to both sites. Due to the laboratory change about November 1<sup>st</sup> of 2004 and possible analytical artifacts, phosphorous was excluded as well. The two data sets were analyzed by positive matrix factorization (PMF) to infer the sources of PM observed at the two sites. The analysis identified nine common source-related factors: sulfate-rich secondary aerosol, cement/carbon-rich, wood smoke, motor vehicle/road dust, metal processing, nitrate-rich secondary aerosol, soil, sea salt, and chloride depleted marine aerosol. Sulfate and nitrate mainly exist as ammonium salts. The sulfate-rich secondary aerosol accounts for 42% and 43% of the PM mass concentrations at the Hamshire and the Orange sites, respectively. The factor containing highest concentrations of Cl and Na was attributed to sea salt due to the proximity of the monitoring sites to the Gulf of Mexico. The chloride depleted marine aerosol was related to the sea salt aerosol but was identified separately due to the chlorine replacement reactions. Basically, the factors of sulfate, motor vehicle/road dust, nitrate, soil, sea salt, and chloride depleted marine aerosol at the two sites showed similar chemical composition profiles and seasonal variation reflecting these six factors were likely to be Golden Triangle regionally related sources. Cement/carbon-rich, wood smoke, and metal processing factor were likely to be the local sources.

**Key words:** Airborne particulate matter, Positive matrix factorization, Receptor modeling, Factor contribution, Factor profiles

### INTRODUCTION

Beaumont, Texas (30°05'N, 94°06'W) located on the west bank of the Neches River, 130 km east of Houston and 45 km north of the Gulf of Mexico, is a medium size urban area in Southeast Texas. With two smaller neighboring cities, Port Arthur and Orange, it constitutes the so-called

\*Corresponding author E-mail:paul.chiou@lamar.edu

Golden Triangle in Texas, a major industrial area on the Texas Gulf Coast. The Golden Triangle metropolitan area has a population of over 380 thousand. Shipbuilding, livestock raising, and rice farming spread in the surrounding area. Several major chemical, petrochemical, and paper plants, refineries, rice mills, and waste management sites are located in Beaumont and Port Arthur. The

Hamshire and Orange monitoring sites operated by US EPA are located in the region of Golden Triangle. In an effort to better characterize the ambient air quality in Golden Triangle of Texas, it is important to identify the possible sources of  $PM_{2.5}$  in the region. To understand the source/receptor relationship, multivariate receptor models have been applied to the observed speciated PM over the years. The multivariate approach is based on the fundamental principle that mass conservation can be assumed, and a mass balance analysis can be used to identify and apportion sources of airborne particulate matter in the atmosphere (Hopke, 1985; 1991). Among the multivariate receptor models, positive matrix factorization (PMF) is a relatively new technique developed by Paatero (1997) and Paatero and Tapper (1993; 1994). It has been successfully applied to several source attribution studies (Anttila *et al.*, 1995; Chueinta *et al.*, 2000; Junnto and Paatero, 1994; Kim *et al.*, 2003; 2004; Lee *et al.*, 1999; Paterson *et al.*, 1999; Polissar *et al.*, 1996; 1998; 1999; 2001; Xie *et al.*, 1999).

The objectives of this study are to (1) identify the sources of particulate pollutants at the two sites, (2) estimate the source contributions as well as source composition of each possible source (Chiou *et al.*, 2007; Hansen *et al.*, 2003; Liu *et al.*, 2005; Ramadan *et al.*, 2000; 2003; Song *et al.*, 2001; Zheng *et al.*, 2002), and (3) investigate the regional-local source contrast using estimated source contributions of each common factor for the two sites. Such comparison between the two sites in this region of Texas has not been reported in earlier literature.

## MATERIALS & METHODS

The  $PM_{2.5}$  composition sample data analyzed in this study was downloaded from the US EPA website at <http://www.epa.gov/ttn/airs/airsaqs/>, and processed to conform to the PMF data format. The original 24-h integrated samples were collected at the Hamshire monitoring site (29°51'50"N, 94°19'04"W) and Orange site (30°11'39"N, 93°52'01"W) using a fine particle sequential sampler. The Hamshire site, 3 km off the interstate highway I-10, is located 35 km southwest of Beaumont, and the Orange site, 6 km north of I-10, is located 30 km northeast of

Beaumont (Fig.1). The sequential sampler used at the two sites was Partisol-Plus Model 2025 Sequential Air Sampler (Rupprecht/Patashnick Co. Inc.) with very sharp cut cyclone fractionators. Ambient air samples are drawn at a constant flow rate into a specially shaped inlet that removes particles with aerodynamic diameters greater than 2.5  $\mu m$ , and the remaining particles are passed through a filter. Integrated 24-h  $PM_{2.5}$  particle samples were collected on Teflon filters. Total mass was then determined gravimetrically from the filters. Most of PM samples were collected every third day and some were collected daily during the time period between July 2003 and August 2005. A total of 269 and 293 samples were separately obtained at the Hamshire and Orange sites. Both mass concentration and elemental chemical speciation were determined using an energy dispersive X-ray fluorescence (XRF). An ion chromatography (IC) was used to analyze sulfate ( $SO_4^{2-}$ ), ammonium ( $NH_4^+$ ), and nitrate ( $NO_3^-$ ) concentrations. The thermal optical transmission technique was used to measure both organic carbon (OC) and elemental carbon (EC). A total of 52 chemical elements was analyzed, including: Ag, Al, As, Au, Ba, Br, Ca, Cd, Ce, Cl, Co, Cr, Cs, Cu, Eu, Fe, Ga, Hf, Hg, In, Ir, K, La, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Rb, S, Sb, Se, Si, Sm, Sn, Sr, Ta, Tb, Ti, V, Y, Zn, Zr, W, OC, EC, SO,  $NH_4^+$ ,  $NO_3^-$ .

In the data, the concentration of XRF S and  $SO_4^{2-}$  were highly correlated (slope = 2.83,  $r^2 = 0.95$  for the Hamshire data; slope = 2.72,  $r = 0.94$  for the Orange data), thus it is reasonable to exclude XRF S from the analysis (Kim *et al.*, 2004). The XRF analysis of PM speciation filters at sites in Texas was conducted by the Research Triangle Institute. After October 31<sup>st</sup>, 2004, the XRF analysis of filters from all except three of Texas PM speciation sites was switched to the Desert Research Institute laboratory. The two sites used in this PMF analysis are among those switched to. Due to an analytical artifact in the data resulting from a change in analytical laboratories after October 31<sup>st</sup>, 2004, 94% and 96% of phosphorous concentrations at the Hamshire

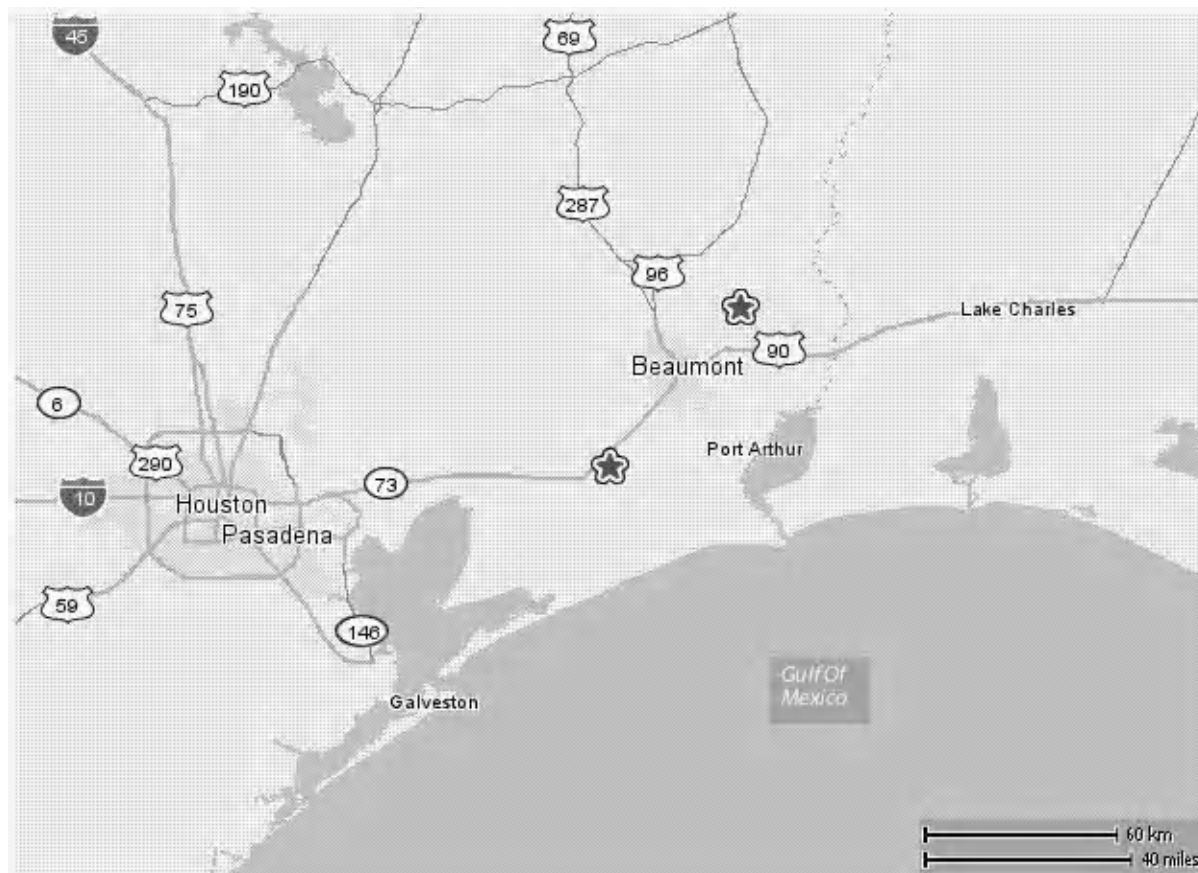


Fig. 1. Location of Hamshire and Orange monitoring sites in star

and Orange sites, respectively, were below detection limits prior to November 1<sup>st</sup> of 2004, while the phosphorous levels were suddenly above the detection limits after the laboratory change. Furthermore, there were a high correlations of = 0.97 and 0.87 between phosphorous and sulfur at the Hamshire and Orange sites, respectively, posterior to October 31<sup>st</sup>, 2004. On the other hand, there are 22 chemical species with a lower signal-to-noise ratio because of too many below-detection-limit measurements. As a result, these 22 species and phosphorous were excluded in the PMF analysis. Among the 22 species excluded, 20 are common species to the two sites. The analysis of the compositional data, however, still revealed a mass closure violation after excluding these species. The comparison of measured PM mass to the sum of PM compositional data indicates that 7.1% for the Hamshire and 7.5% for the Orange site of the measured PM mass concentrations were less than the sum of species concentrations. In the data matrices, there were missing and below-detection-limit values. The analytical uncertainty estimates associated with

each measured concentration and the detection limits for instruments were also reported. The possible source of measurement errors which can be considered as the uncertainty include the collection of over- and under- sized particles, the mass gain or loss of the collected PM mass, and the analytical sensitivity. (Table 1 & Table 2) summarize the PM speciation data used in this study.

In this study, PMF was used with the data collected at the Hamshire and Orange site as discussed previously. PMF is an approach of factor analysis, and it is described in detail by Paatero (1997). Only a brief description of this approach is provided here. PMF uses the method of weighted least-squares to solve a general receptor modeling problem. The general model assumes there are  $p$  sources, source types or source regions (termed factors) impacting a receptor and the observed concentrations of various species at the receptor are linear combinations of the impacts from the  $p$  factors. The factor analysis model (PMF) can be written as:

**Table 1. Summary of PM<sub>2.5</sub> and 28 species mass concentrations at Hamshire used for PMF analysis**

Species	Concentration (µg/m <sup>3</sup> )				Number of BDL <sup>b</sup> values (%)	Number of missing values (%)
	Geometric mean <sup>a</sup>	Arithmetic mean	Minimum	Maximum		
PM <sub>2.5</sub>	9865	10986	2324	40400	0 (0.0)	5 (1.8)
Al	16	61	4.6	1499	128 (47.6)	5 (1.8)
As	0.56	0.90	0.1	4.7	142 (52.8)	5 (1.8)
Ba	6.3	10.5	1.7	354	194 (72.1)	5 (1.8)
Br	2.1	2.6	0.2	13	79 (29.4)	5 (1.8)
Ca	46	58	1.5	273	1 (0.37)	5 (1.8)
Cl	9.9	89	0.75	1400	97 (36.1)	5 (1.8)
Cr	0.66	0.95	0.35	21	196 (72.9)	5 (1.8)
Cs	4.5	7.0	2	317	229 (85.1)	5 (1.8)
Cu	0.88	1.33	0.5	38	183 (68.0)	5 (1.8)
Fe	34	64	0.43	815	3 (1.1)	5 (1.8)
Hg	0.77	0.90	0.35	5.3	229 (85.1)	5 (1.8)
K	65	82	9.8	768	0 (0.0)	5 (1.8)
Mg	12	27	3.6	373	170 (63.2)	5 (1.8)
Mn	1.0	1.4	0.31	12	151 (56.1)	5 (1.8)
Na	66	151	11	1130	105 (39.0)	5 (1.8)
Ni	0.75	0.92	0.34	4.7	133 (49.4)	5 (1.8)
Pb	1.7	2.4	0.55	39	155 (57.6)	5 (1.8)
Se	0.46	0.58	0.1	3.4	228 (84.8)	5 (1.8)
Si	84	192	3.1	2607	4 (1.5)	5 (1.8)
Sr	1.1	1.6	0.3	7.6	147 (54.6)	5 (1.8)

<sup>a</sup>Data below the limit of detection were replaced by half of the reported detection limit values for the geometric mean calculations

<sup>b</sup>Below detection limit

$$x_{ij} = y_{ij} + e_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij}, i=1, \dots, n; j=1, \dots, m, \quad (1)$$

where  $x_{ij}$  is the  $j$ th species concentration measured in the  $i$ th sample,  $g_{ik}$  is the particulate mass concentration from the  $k$ th source (or factor) contributing to the  $i$ th sample,  $f_{kj}$  is the  $j$ th species mass fraction from the  $k$ th source (or factor),  $e_{ij}$  is the residual associated with the  $j$ th species concentration measured in the  $i$ th sample, and  $p$  is the total number of sources (or factors). The objective of PMF is to estimate the mass contributions and the mass fractions (profiles) in Eq. (1) by the weighted least-squares. The task

of PMF is thus to minimize the sum of the squares of the residuals weighted inversely with error estimates (estimated uncertainties) of the data points. In other words, the data analysis by PMF can be described as to minimize the objective function  $Q$ ,

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

under constraints  $g_{ik} \geq 0$ ,  $f_{kj} \geq 0$ , and with  $s_{ij}$  as the error estimate (estimated uncertainty) for  $x_{ij}$ . The estimates of source contributions and source profiles are obtained by a unique algorithm in which both matrices  $\mathbf{G} = [g_{ik}]$  and  $\mathbf{F} = [f_{kj}]$

**Table 2. Summary of PM<sub>2.5</sub> and 28 species mass concentrations at Orange used for PMF analysis**

Species	Concentration (µ/gm <sup>3</sup> )				Number of BDL <sup>b</sup> values (%)	Number of missing values (%)
	Geometric mean <sup>a</sup>	Arithmetic mean	Minimum	Maximum		
PM <sub>2.5</sub>	10719	12050	150	60400	1 (0.34)	29 (10)
Al	16	54	4.6	1420	146 (49.8)	29 (10)
As	0.59	0.82	0.1	4.2	146 (49.8)	29 (10)
Ba	7.1	9	2	84.6	199 (67.9)	29 (10)
Br	1.9	2.9	0.2	49.3	68 (23.2)	29 (10)
Ca	47	61	1.5	824	2 (0.68)	29 (10)
Cl	8.2	63	0.75	1580	106 (36.2)	29 (10)
Co	0.4	0.45	0.28	6.9	244 (83.3)	29 (10)
Cr	0.78	1.2	0.35	23.2	162 (55.3)	29 (10)
Cu	1.6	2.9	0.5	40.8	112 (38.2)	29 (10)
Fe	73	116	0.43	861	2 (0.68)	29 (10)
K	72	87	2.2	656	1 (0.34)	29 (10)
Mg	14	24	3.7	343	180 (61.4)	29 (10)
Mn	1.5	2.6	0.31	24.5	121 (41.3)	29 (10)
Na	67	129	11	1280	118 (40.3)	29 (10)
Ni	0.72	0.9	0.34	11	138 (47.1)	29 (10)
Pb	1.8	2.3	0.55	26	147 (50.2)	29 (10)
Se	0.5	0.6	0.1	21	238 (81.2)	29 (10)
Si	87.6	177	3	2477	5 (1.7)	29 (10)
Sr	1.2	1.6	0.35	15.3	145 (49.5)	29 (10)
Ti	3.1	6.4	0.55	90	96 (32.8)	29 (10)
V	1.8	2.6	0.1	11	77 (26.3)	29 (10)

<sup>a</sup>Data below the limit of detection were replaced by half of the reported detection limit values for the geometric mean calculations

<sup>b</sup>Below detection limit

are adjusted in each iteration step. The process continues until convergence occurs (Paatero, 1997; Polissar *et al.*, 1998).

The application of PMF requires the estimated uncertainty for each data value to be carefully selected so that it reflects the quality and reliability of each data point. This important feature of PMF enables us to properly handle any below detection limit and missing data values. The uncertainty estimate provides a useful tool to decrease the weight of any below detection limit and missing data values when searching for the minimum of *Q* in Eq. (2). In this study, the procedure of Polissar *et al.* (1998) was adopted as follows: (i) the

concentration value was the actually measured concentration, and the sum of the analytical uncertainty and one third of the detection limit value was used as the estimated uncertainty if was a determined value; (ii) the concentration value was replaced by half of the detection limit value, and five sixths of the detection limit value was used as the estimated uncertainty if was below detection limit; (iii) the concentration value was set equal to the geometric mean of all the measured values of for element *j*, and its corresponding uncertainty was set equal to four times of this geometric mean value if was a missing data value. Half of the average detection limits

were used for below detection limits values in the calculation of the geometric means. Furthermore, the estimated uncertainties of OC and SO were increased by a factor of three because of its magnitude compared to the lower concentration species.

It is well-known that extreme data values as well as true outliers can distort the least-squares estimation profoundly. A delicate handling of these data values is important, and PMF offers a robust mode to properly weigh these data points in the process of searching for the minimum of  $Q$ . The robust factorization based on the Huber influence function (Huber, 1981) is a technique of iterative reweighing of the individual data values. The least squares approach with the robust factor analysis leads now to the equations:

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

Where:

$$h_{ij}^2 = \begin{cases} 1 & \text{if } |e_{ij} / s_{ij}| \leq \alpha, \\ |e_{ij} / s_{ij}| / \alpha & \text{otherwise,} \end{cases}$$

and  $\alpha$  is the outlier threshold distance. The value of  $\alpha = 4.0$  was chosen in this study. Because of the mass closure violation noted previously, the measured particle mass concentration was included as an independent variable in the PMF modeling to directly obtain the mass apportionment instead of using a regression analysis (Kim *et al.*, 2003; 2004). The estimated uncertainties of the PM<sub>2.5</sub> mass concentrations were set at four times of their values to reduce their weight in the model fit so that the magnitude of PM mass will not skew the analysis. When the measured particle mass concentration is included as an independent variable, the PMF apportions a mass concentration for each source based on its temporal variation without using a multiple linear regression. The results of PMF modeling are then normalized by the apportioned particle mass concentrations so that the quantitative source contributions are obtained. Specifically,

$$x_{ij} = \sum_{k=1}^p (c_k g_{ik})(f_{kj} / c_k), \quad (4)$$

where  $c_k$  denotes directly apportioned mass concentration by PMF for the  $k$ th factor.

## RESULTS & DISCUSSION

An essential step in PMF analysis is to determine the number of factors and source apportionment. For determination of the number of factors, the basic consideration is to obtain a good fit of the model to the original data, and the model can well explain the physical meaning of the data. If there is goodness of fit, the theoretical  $Q$  value in Eq. (3) should be approximately equal to the number of degrees of freedom or approximately equal to the number of entries in the data array provided that correct values of  $s_{ij}$  have been used (Yakovleva *et al.*, 1999). In a well-fit model, the residuals  $e_{ij}$  and the error estimates  $s_{ij}$  should not be too much different in size, and the ratio  $(e_{ij} / s_{ij})$  should fluctuate between  $\pm 3$ . Junnto and Paatero (1994) recommended values of  $\pm 2$  for the ratio. Based on the criterion of obtaining the most physically meaningful solution with the calculated  $Q$  value ( $Q = 8012$  and  $9597$  for the Hamshire and Orange sites, respectively) close to the theoretical  $Q$  value ( $Q = 7801$  and  $8497$  for the Hamshire and Orange sites, respectively), the PMF identified nine common source types by trial and error with different numbers of factors. We termed these factors as sulfate-rich secondary aerosol, cement/carbon-rich, wood smoke, motor vehicle/road dust, metal processing, nitrate-rich secondary aerosol, soil, sea salt, chloride depleted marine aerosol. The contributions of these factors towards the PM mass, i.e.,  $c_k$  in Eq. (4) from PMF at the two sites are summarized in (Table 3). To study the spatial variations contributed by different factors between the two sites, the square of correlation coefficient ( $R^2$ ) was calculated from the estimated source contributions with respect to the common factor for the two sites. (Table 4) presents the summary of the squared correlation coefficients for the factors.

Both (Figs. 2 & 3) show a relationship between the reconstructed PM<sub>2.5</sub> mass contributions from all sources and the measured PM mass concentrations. It is clear that the resolved sources effectively reproduce the

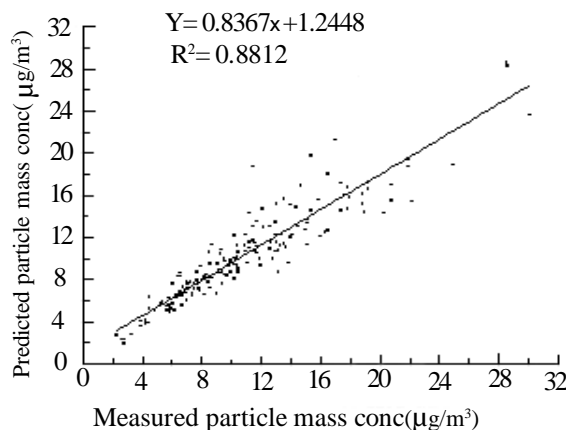
measured values and account for most of the variation in the PM mass concentrations (slope = 0.84 and  $r = 0.88$  for the Hamshire site; slope = 1.03 and  $r = 0.81$  for the Orange site). (Fig. 4 through Fig. 12) present the time series plots of estimated source contributions to PM mass concentrations for each source, the identified source profile at the two sites, and polar plot of the factor contributions with wind direction for the two sites. The seasonal variations in the time series plots may be explained by variation in source strength, atmospheric transport, and possible chemical reactions in the atmosphere, or a combination of the three. To estimate the uncertainties of source profiles obtained from PMF, a bootstrapping technique (Efron and Tibshirani, 1993) combined with a method to account for the rotational freedom in the solution was used. The middle panels of (Fig. 4 through Fig. 12) display the lower and upper limit of a 90% confidence interval for the mean profiles as well as the profile estimates. The smaller magnitude the whisker, the more consistent the estimate is and the larger magnitude the whisker, the less consistent the result is. In other words, the smaller magnitude the whisker, the smaller associated error the estimate has and the larger magnitude the whisker, the larger associated error the estimate has. Among the nine factors identified, the first common factor is sulfate-rich secondary aerosol. Fig. 4. shows the contribution and factor profile resolved by PMF for this factor at the two sites. The source has a high concentration of carbon,  $\text{SO}_4^{2-}$ , and  $\text{NH}_4^+$ .

**Table 3. Possible source types and factor contributions (%) obtained by PMF**

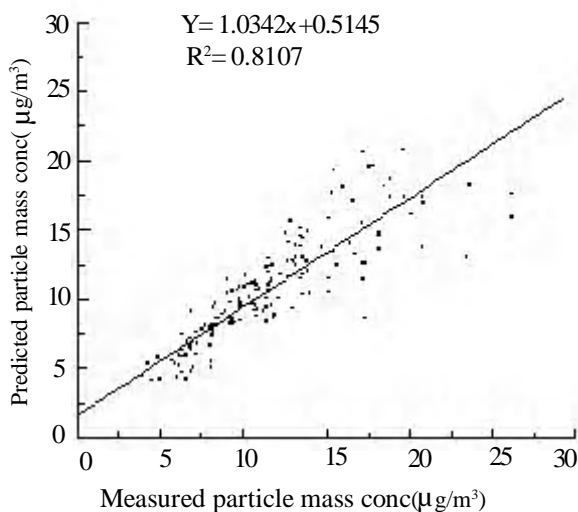
Source type	Hamshire site	Orange site
Sulfate-rich	42.0	43.4
Cement/carbon-rich	13.7	10.6
Wood smoke	10.9	15.5
Motor vehicle/road dust	8.7	8.2
Metal processing	8.3	5.9
Nitrate-rich	4.2	3.2
Soil	3.7	4.2
Sea salt	3.2	2.3
Cl depleted marine aerosol	5.3	6.7

**Table 4. Hamshire vs. Orange  $R^2$  for the factors by PMF at the two sites**

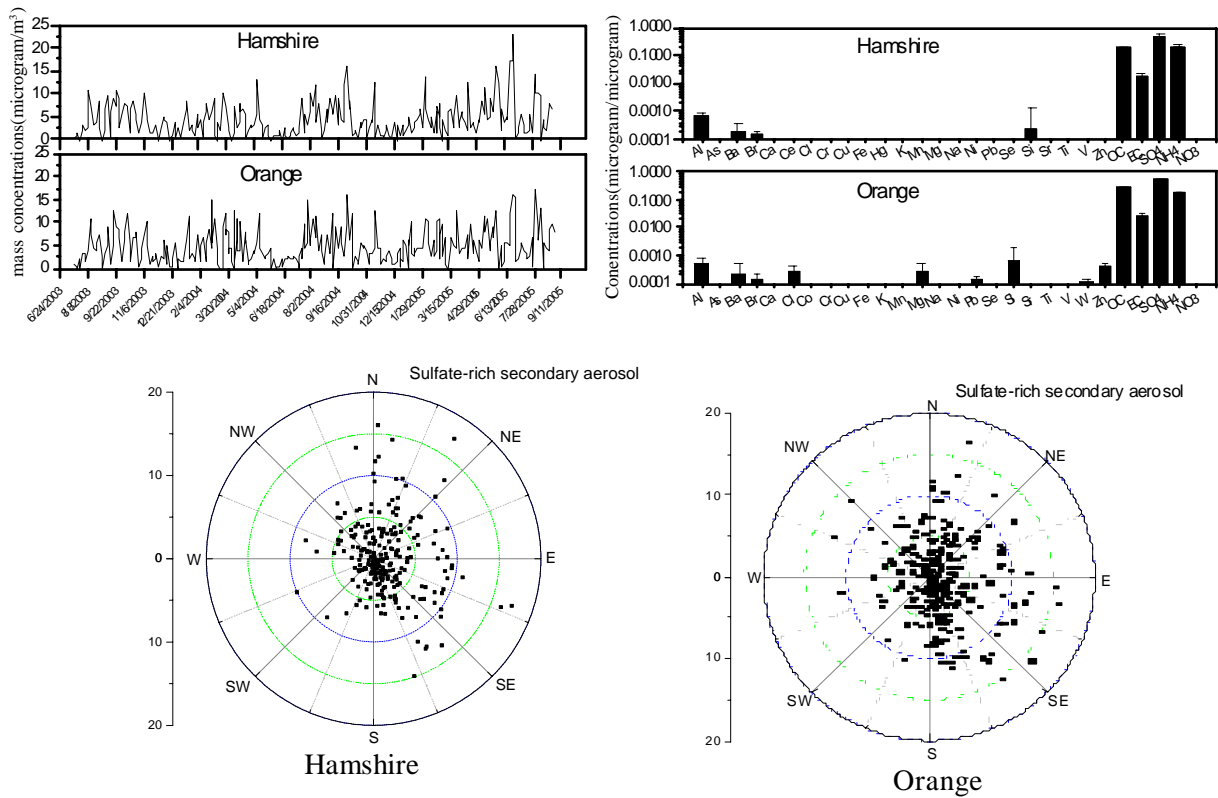
Source type	Squared correlation coefficient
Sulfate-rich	0.67
Cement/carbon-rich	0.04
Wood smoke	0.15
Motor vehicle/road dust	0.90
Metal processing	0
Nitrate-rich	0.32
Soil	0.95
Sea salt	0.62
Cl depleted marine aerosol	0.29



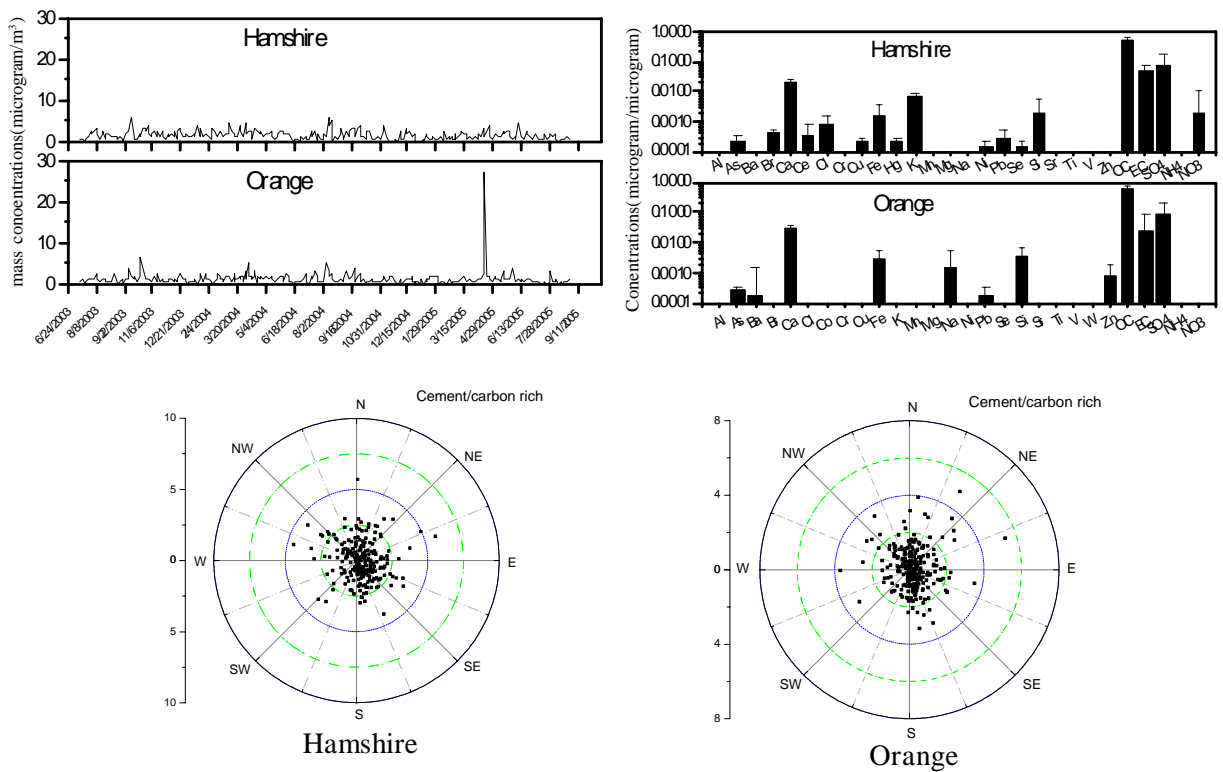
**Fig. 2. Measured versus predicted PM<sub>2.5</sub> mass concentration for Hamshire**



**Fig. 3. Measured versus predicted PM<sub>2.5</sub> mass concentration for Orange**



**Fig. 4.** Contributions by and profiles of the sulfate factors resolved by PMF analysis at the two sites; the polar plot of the factor contributions with wind direction for Hamshire and Orange sites



**Fig. 5.** Same as Fig. 4. but for the cement/carbon-rich factors



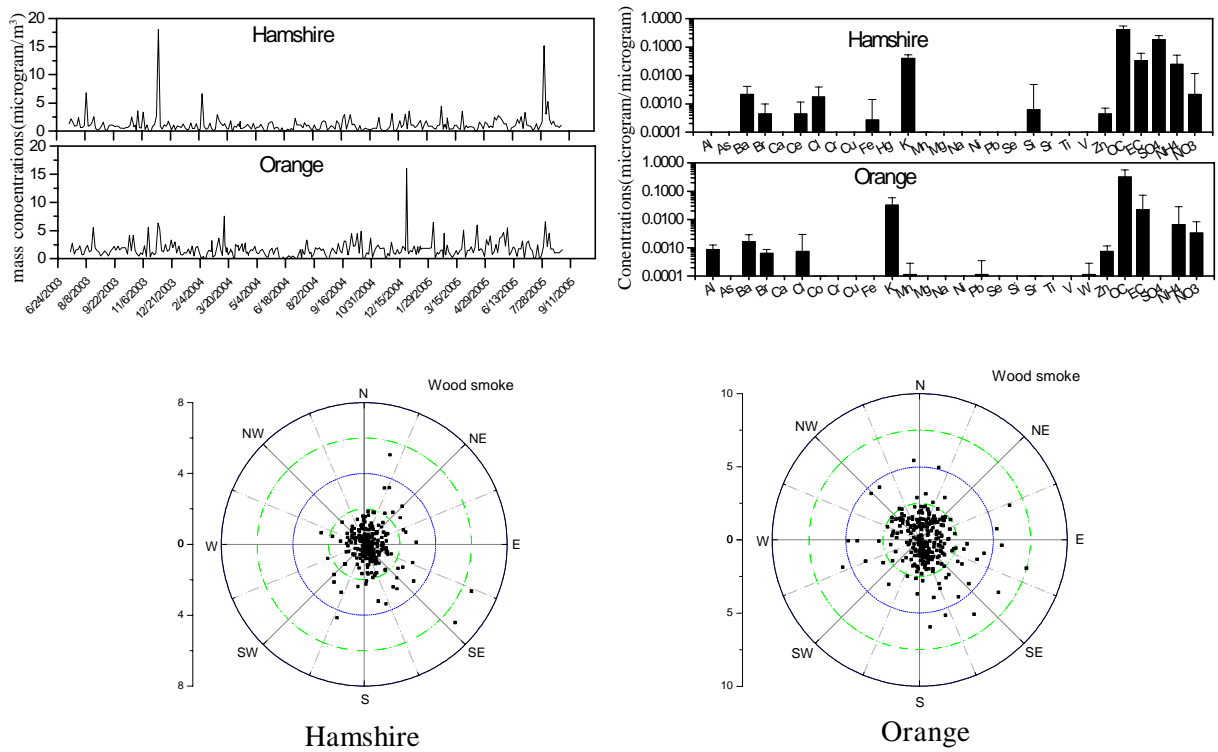


Fig. 6. Same as Fig. 4 but for the wood smoke factors

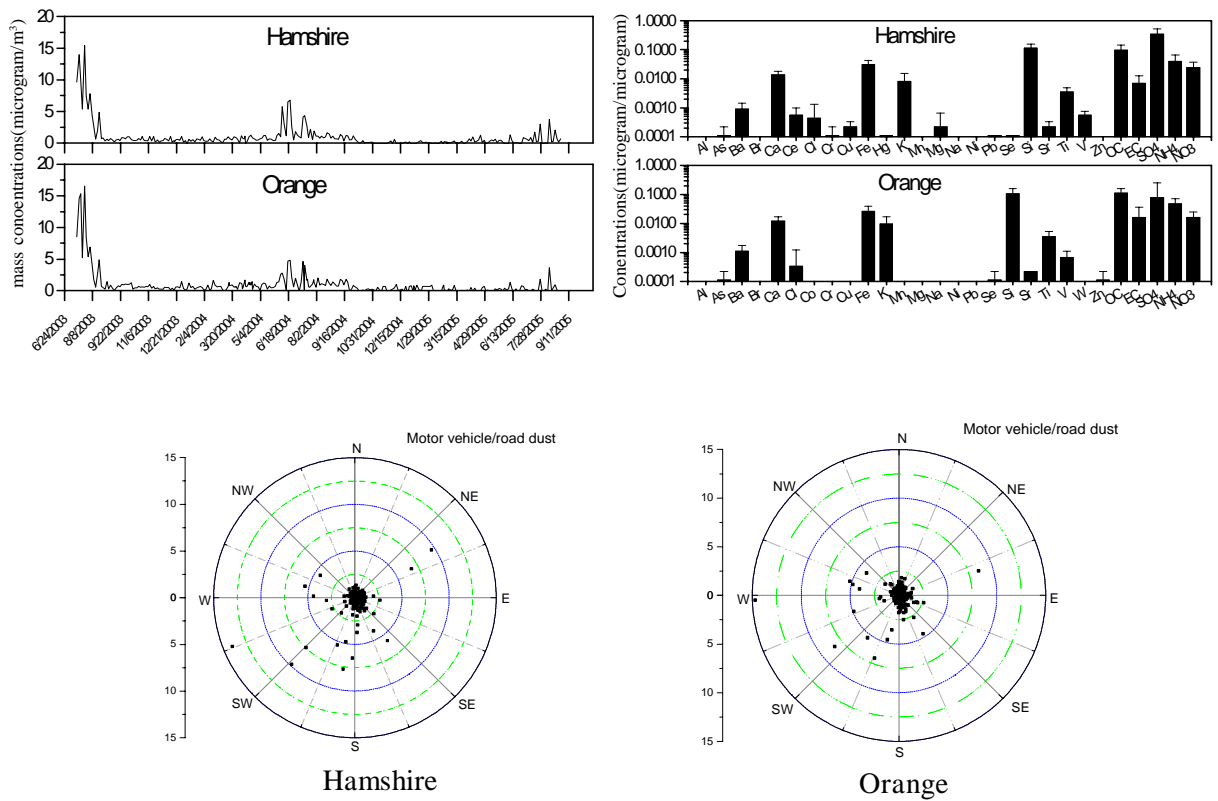


Fig. 7. Same as Fig. 4 but for the motor vehicle/road dust factors

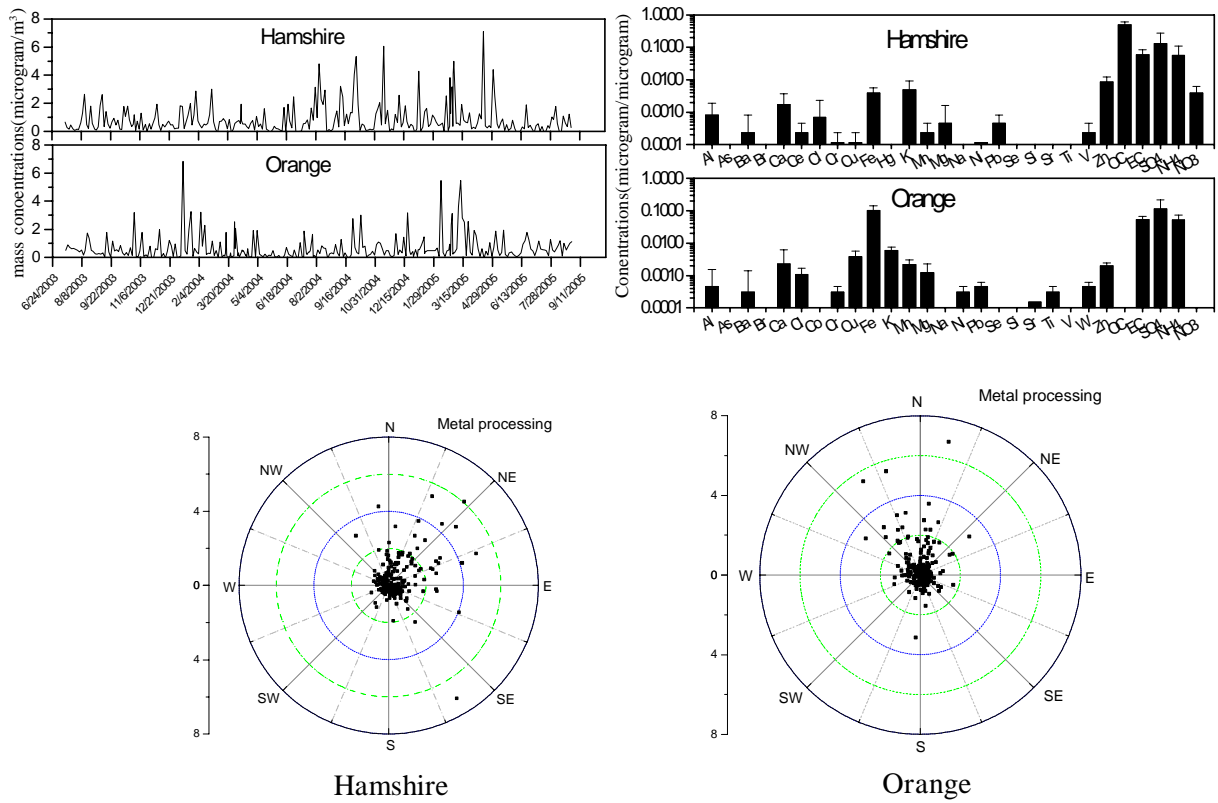


Fig. 8. Same as Fig. 4 but for the metal factors

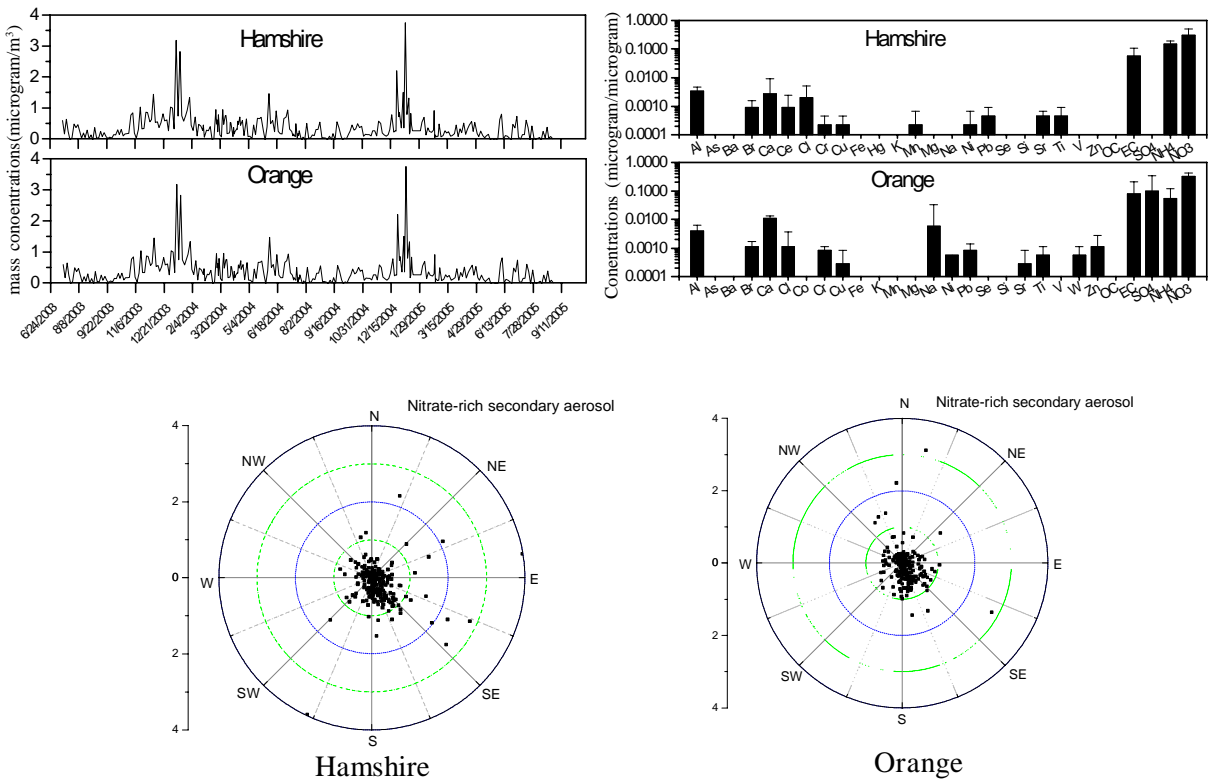


Fig. 9. Same as Fig. 4 but for the nitrate factors

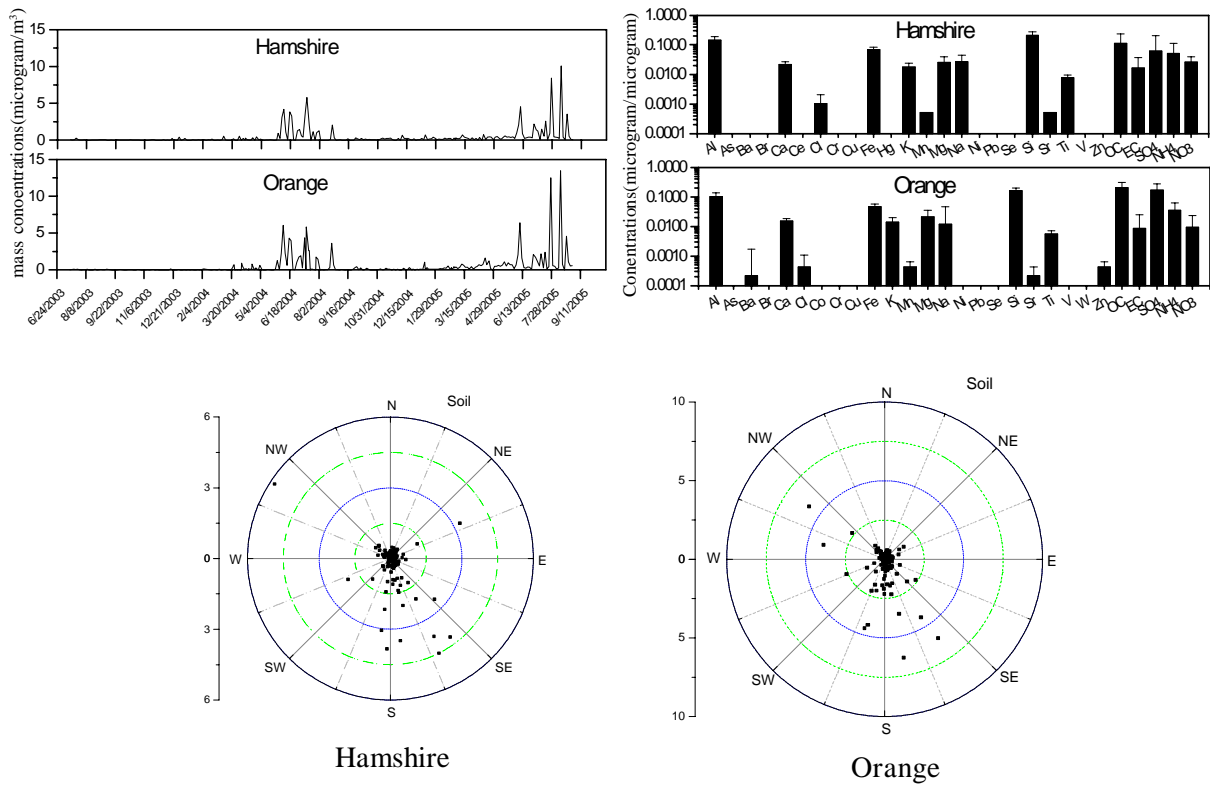


Fig. 10. Same as Fig. 4 but for the soil factors

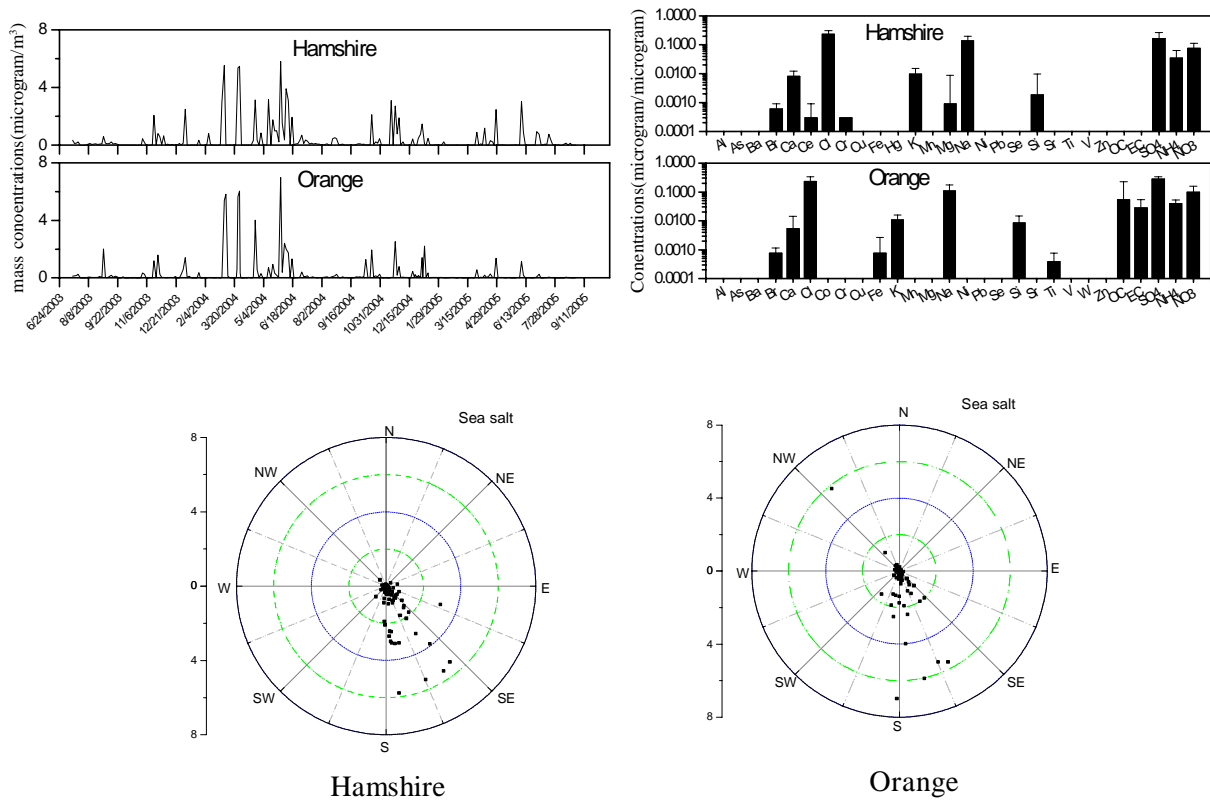


Fig. 11. Same as Fig. 4 but for the sea salt factors

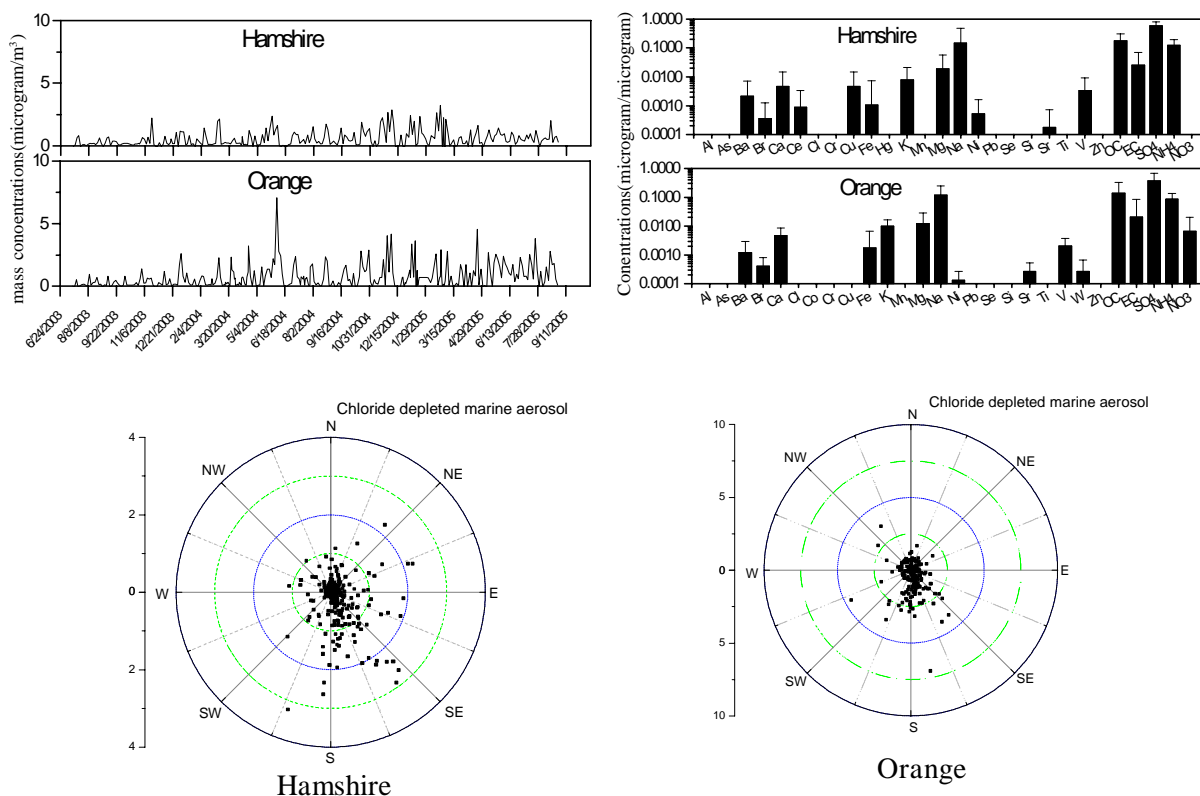


Fig. 12. Same as Fig. 4 but for the chloride depleted marine factors

OC and EC were associated with this factor. The significant OC association was consistent with several previous studies (Kim *et al.*, 2004; Ramadan *et al.*, 2000). The mixed EC concentration probably reflects that the resolved factor by PMF may not merely represent one source. Molar ratio of ammonium to sulfate for this factor was 2.1 at the Hamshire site and 1.9 at the Orange site. Because of the possible evaporation of ammonium during sample analysis and/or the uncertainty of the PMF estimate, sulfate is likely present mainly as ammonium sulfate at the two receptor sites. An analysis on the PM<sub>2.5</sub> data for the molar equivalents between  $[Na^+] + [K^+] + 2[Ca^{2+}] + [NH_4^+]$  and  $2[SO_4^{2-}] + [NO_3^-] + [Cl^-]$  was performed. It revealed the sum of the cations is equivalent to the sum of anions (normalized molar ratios of cations to anions were 0.94 and 1.00 for Hamshire and Orange, respectively). The cation is dominated by ammonium (82% on a normalized molar basis for both sites) and sodium (13% and 12% at the Hamshire and Orange sites, respectively). The anion is dominated by sulfate (93% and 94% on a

normalized molar basis at the Hamshire and the Orange sites, respectively). All other ionic constituents were less than 5% by mole. Correlation of the ionic compositions indicates that sulfate is present primarily as ammonium sulfate and chloride is present mainly as sodium chloride. Nitrate is negatively correlated to ammonium at both sites, but weakly correlated with sodium, potassium, and calcium. This indicates that nitrate is less likely to be associated with ammonium. Instead, it may be bound with a number of cations in the PM. Among the nine factors identified, sulfate-rich secondary aerosol has the highest source contribution to PM mass concentration with 42.0% and 43.4% at the two sites, respectively. Carbon and trace elements usually become associated with the secondary sulfate aerosol in the atmosphere (Kim *et al.*, 2004). The source shows slightly higher concentrations in spring and early fall when the photochemical activity is still high in the region (Polissar *et al.*, 2001; Song *et al.*, 2001). In the polar plot of this factor with wind directions at the two sites (bottom of Fig. 4), the contributions from southeast indicate the sulfate aerosol was influenced by the sources along the Gulf Coast under prevailing flow from the

southeast. However, the contributions from northeast were likely to include eastern U.S. and east Texas sources under less frequent northeasterly wind conditions. The sulfate-rich secondary aerosol source accounts for 42.0% and 43.4% of the PM mass concentration at the two sites, respectively. This is similar to the study of three northeastern US cities which identified its contributions of 47%, 55%, and 51% to the PM mass concentration (Song *et al.*, 2001). The top panel in Figure 4 shows highly similar seasonal variations at the two sites. This highly similar seasonal variations at the two sites and a significant squared correlation coefficient of = 0.67 for this factor between the two sites imply that this factor is a Golden Triangle regionally related factor.

The second common factor is related to a cement/carbon-rich source characterized by Ca, OC, and EC (Kim *et al.*, 2004; US EPA, 2002). The middle panel in Figure 5 shows the factor profiles at the two sites. SO was associated with this factor at the two sites while a small amount of K was associated with this factor at the Hamshire site. It contributes 13.7% and 10.6% to the PM mass concentration at the sites, respectively, and likely includes contributions from a couple of construction material sites and an unknown carbon-rich source possibly from chemical plants in Golden Triangle. The high carbon concentration of this source indicates that the cement and a carbon-rich source are co-located and daily emission patterns are similar (Kim *et al.*, 2004). The source shows slightly higher concentrations in spring. The top panel in Figure 5 shows the contribution of this factor at the sites. It did not suggest any seasonal dependence and seasonal variations for this factor. There was an extremely high peak on April 16, 2005 at the Orange site but not at the Hamshire site. This factor was possibly dominated by weekday-weekend local activity such as reduced activity at the construction material sites over weekends. The factor contribution plot for the two sites also shows little similar temporal variability between the two sites. In the polar plot of this factor with wind directions at the two sites (bottom of Fig. 5). the slightly higher contributions from east and southeast indicate the carbon aerosol was possibly influenced by the chemical and refinery

plants along the coast in Port Arthur and Orange under prevailing flow from the southeast. The factor contribution peaks and temporal variability did not match between the two sites with a value of 0.04. This indicates that the two sites may be influenced by some different local sources.

The third common factor was identified as wood smoke source which is characterized by K, OC, and EC (Kim *et al.*, 2004; Liu *et al.*, 2005; Watson *et al.*, 2001). It contributes 10.9% and 15.5% to the PM mass concentration at the two sites, respectively. The middle panel in Fig. 6. shows the factor profiles at the two sites. SO was associated with this factor at the Hamshire site and a small amount of NO and NH were associated with the factor at the two sites. The wood smoke probably comes from residential wood burning, local agricultural biomass burning, and occasional forest fires. For the two sites, this factor has a slightly higher trend in winter season possibly related to residential wood burning for heating. The short-term peaks in spring and summer shown at the top panel of Fig. 6. were probably due to forest fires, and/or biomass burning from Central America. The polar plot of this factor at the two sites (bottom of Fig. 6). indicated there were slightly higher contributions from east and southeast. However, the squared correlation coefficient of this factor between the two sites was = 0.15 and it is considered substantially low. Furthermore, the plot of factor contribution shows little similar temporal variability between the two sites. It indicates that the two sites are influenced by different types of local sources such as forest fires and/or different annual and semiannual wood burning activity.

The fourth common factor was not as readily interpreted as the other factors; however, it was identified as motor vehicle/road dust source characterized by higher concentration of Si, OC, SO, NH along with Ca, Fe, K, and EC (Chueinta *et al.*, 2000). The middle panel in Figure 7 shows the quite similar factor profiles at the two sites. The association of NH and NO with this factor might be directly related to NO and NH emissions from (catalytic) motor vehicles. This source might be accounted for the mixing of sources such as vehicles on interstate highway I-10, state highway SH 124 and SH 62, road dust, and transported summer soil. Highway I-10 and SH 124 are nearly

parallel in the proximity of Hamshire, and the monitoring site is located about 2 km north of SH 124 and 3 km south of I-10. The Orange site is located about 6 km north of I-10 and 0.5 km west of SH 62 (Fig. 1). It has short-term peaks in June and July, and shows a summer-high seasonal trend possibly due to the higher concentration of soil dust during the period. This source accounts for 8.7% and 8.2% of the PM mass concentration at the two sites, respectively, and it shows slightly higher concentrations in summer. At the top panel in Fig. 7. the large peaks of annual cycle indicated the seasonal dependence of source formation possibly from the summer soil dust. The plot of factor contributions shows highly similar seasonal variations at the two sites due to the summer dust and the monitoring sites being in the surrounding area of Golden Triangle. The polar plot of this factor with wind directions at the two sites (bottom of Fig. 7). did not suggest any particular higher contributions from any direction. The apparently similar temporal variability and a squared correlation coefficient of = 0.90 between the two sites imply that this source is highly influenced by the summer soil dust and area traffic sources.

The fifth common factor suggested a source of metal processing because of the profile characterized by its high concentration of Zn along with OC, EC, SO, and NH at the Hamshire site and high concentration of Fe associated with EC, SO, and NH at the Orange site (Kim *et al.*, 2004; US EPA, 2002). The middle panel in Figure 8 shows the factor profiles at the two sites. This factor at Hamshire is highly likely to include contributions from a couple of metal processing facilities in Beaumont as Beaumont is located northeast of the Hamshire receptor site. One of the facilities, 20 km northeast of the Hamshire receptor, uses the process of galvanization to coat steel and iron with zinc. The polar plot of this factor at the two sites (bottom of Fig. 8). indicated there were slightly higher contributions from northeast likely to include eastern U.S. and east Texas sources under less frequent northeasterly wind conditions. A major steel mill, 1 km south of I-10 and 10 km east of Beaumont, is also likely to contribute to this factor at the Orange site. The source showed a slightly reduced seasonal trend at the receptors in the spring possibly due to the locations of metal processing facilities and

southerly winds during the season. This source accounts for 8.3% and 5.9% of the PM mass concentration at the two sites, respectively. The top panel in Fig. 8. shows the contribution for this factor at the two sites. The peaks with high periodicity might correspond to the annual and semiannual cycle possibly reflecting the seasonal variation, and those with low periodicity were related to a weekly cycle. The weekly high peaks suggested this factor was dominated by weekday-weekend local activity. However, the plot of factor contributions shows little similar temporal variability between the two sites. The much different time variations at the two sites and a squared correlation coefficient of = 0 between the two sites emphasize this factor may come from different local metal sources or common sources with different impacts on the two sites.

The sixth common factor resolved at the two sites mainly consists of ammonium and nitrate. The nitrate-rich secondary aerosol is identified by its high concentration of NO and NH. Fig. 9 .shows the factor contribution and profile results for this source. SO was mixed with this factor at the Orange site. EC and a small amount of trace metals were also associated with this factor possibly from the transported metallic aerosol. This source includes NH that becomes associated with the secondary nitrate aerosol in the atmosphere. Molar ratios of ammonium to nitrate were 1.7 and 0.6 for the Hamshire and Orange sites, respectively. Because of the possible evaporation of ammonium during sample analysis and/or the uncertainty of the PMF estimate, nitrate is probably present mainly as ammonium nitrate at the two receptor sites. Nitrate is formed in the atmosphere mostly through the oxidation of NO depending on ambient temperature, relative humidity, and the presence of ammonia (Liu *et al.*, 2005). It has short-term peaks and higher trend in cool seasons possibly indicating that low temperature and high humidity foster the formation of nitrate aerosol in the region as discussed in the study for Atlanta (Kim *et al.*, 2004) and three northeastern US cities (Song *et al.*, 2001). In the polar plot of this factor at the two sites (bottom of Fig. 9). the contributions from southeast indicate the nitrate aerosol was influenced by the sources along the Gulf Coast under prevailing flow from the southeast. The seasonal dependence of nitrate

formation is reflected by high peaks with high periodicity in the top panel of Fig. 9. The plot also shows the similar seasonal variations of ammonium nitrate at the two sites. It reflects the regional characteristic of ammonium nitrate formation and transport. The local characteristic of this source in the area was reflected by the small monthly or weekly peaks with low periodicity, and apparently the local impact was limited. The value of 0.32 between the two sites is not as high as that of sulfate factor possibly due to the shorter lifetime of NO than SO. The source accounts for 4.2% and 3.2% of the PM mass concentration at the Hamshire and the Orange sites, respectively. The seventh common factor was identified as soil source represented by Al, Ca, Fe, K, Mg, Na, Si, and Ti (Kim *et al.*, 2004; Watson *et al.*, 2001). It contributes 3.7 % and 4.2% to PM mass concentration at the two sites, respectively. The crustal particles could be contributed by unpaved roads, construction sites, and soil dust. The contribution and factor profile in Figure 10 show the highly similar results of this factor between the two sites. The airborne soil shows seasonal variation with higher concentrations in the summer. The short-term peaks in summer of 2004 and 2005 likely reflect the intercontinental dust transport as indicated in several analyses across the eastern US (Liu *et al.*, 2005). Prospero (2001) showed that the summer trade winds carry African dusts into US from the direction of southeast which is consistent with what the polar plot of this source (bottom of Figure 10) indicated. Sahara dust typically has relatively lower calcium than US or Asian dust. The Al to Ca ratio of 6.7 in this source at the two sites (about 3.8 ratios in US or Asian dust) suggests that this source might have been influenced by African dusts. The mixed OC, EC, SO, NO, and NH concentration in this factor imply that this source was mixed with some other sources during the long-range transport. The large peaks with high periodicity of annual cycle indicated the seasonal variations of this factor. The extremely small peaks with low periodicity suggest that weekly activity has limited impact on this factor at the two sites. The top panel in Figure 10 shows highly similar seasonal trends at the two sites. This highly similar seasonal variations at the two sites and an extremely large squared correlation coefficient of = 0.95 for this source

between the two sites imply that this factor is a regionally related factor.

The eighth common factor at the two sites has high concentration of Cl and Na. It is clearly from the marine or sea salt aerosol source (Lee *et al.*, 1999). As the Hamshire and Orange monitoring sites are located 30 km and 55 km north of the Gulf of Mexico, respectively (Fig. 1). the presence of marine-related aerosol is expected. Fig. 11. shows the contribution and factor profile resolved by PMF for this factor. The middle panel in Fig. 11. shows the comparable factor profiles at the two sites. Both nitrate and sulfate were associated with this factor at the two sites probably due to scavenging of nitrate and sulfate during the transport from the coast. Ammonium was associated with this factor at the two sites while OC and a small amount of EC were also associated with this factor at the Orange site. It has slightly higher concentrations in summer possibly due to the southerly winds during the season. This source accounts for 3.2% and 2.3% of the PM mass concentration at the monitoring sites, respectively. The peaks with high periodicity of annual and semiannual cycle indicated the seasonal variations of this factor. The peaks with low periodicity suggest that the Gulf of Mexico has substantial impact on this factor at the two sites. The polar plot (bottom of Fig. 11). clearly shows the relationship of this factor with wind direction from the Gulf of Mexico. The plot of factor contribution shows highly similar seasonal variations at the two sites due to the proximity of the monitoring sites to the Gulf of Mexico. This highly similar seasonal variations at the two sites and a squared correlation coefficient of = 0.62 between the two sites imply that this source is highly influenced by the monitoring site being in the proximity to the Gulf of Mexico.

The ninth common factor was identified as chloride depleted marine aerosol that is related to the sea salt factor (Lee *et al.*, 1999). The middle panel in Fig. 12. shows the comparable factor profiles at the two sites. It has high concentration of Na, SO, NH, and OC. A small amount of K, Mg, Ni, V, and EC were associated with this factor at the two sites. However, a small amount of NO was also associated with this factor at the Orange site. It is originated from sea salt aerosol which

has undergone the chloride loss reactions through acid substitution and yielded a higher loading of SO in the source than sea-salt aerosol (Lee *et al.*, 1999). This chemical reaction usually occurs in the coastal areas with high sulfur loading. The composition of chloride depleted marine aerosol depends on air quality and meteorological conditions, and therefore it was separately identified from sea salt. The higher sulfate loading in chloride depleted marine aerosol compared to sea-salt has led almost no chloride associated with this factor identified by PMF. However, the lower sulfate loading in sea-salt compared to chloride depleted marine aerosol has led a high chloride loading in sea-salt identified by PMF. The content of Ni, V, and high loading of SO in the source suggests that this factor might have been influenced by marine shipping emissions along the Gulf Coast. Ship traffic is increasingly recognized as a significant source of these trace metals and sulfur in coastal areas. In the bottom panel of Figure 12, there are indications of higher concentrations at the two sites from the direction of south which is consistent with the source direction of sea salt aerosol. The chloride depleted marine aerosol source has slightly higher concentrations in spring and summer. This source accounts for 5.3% and 6.7% of the PM mass concentration at the two sites, respectively. The large peaks with high periodicity of annual cycle indicated the seasonal dependence of formation of this source. The peaks with low periodicity suggest that the Gulf of Mexico has substantial impact on this factor at the two sites. The top panel in Figure 12 shows highly similar seasonal variations at the two sites due to the proximity of the monitoring sites to the Gulf of Mexico. This highly similar seasonal variations at the two sites and the squared correlation coefficient of  $r^2 = 0.29$  between the two sites imply that this source is likely influenced by the monitoring site being in the proximity to the Gulf of Mexico.

## CONCLUSIONS

An air quality study has been carried out to identify and compare the sources of particulate pollutants at two EPA monitoring sites in Texas, namely, the Hamshire and the Orange monitoring sites located in Golden Triangle of Texas with about 65 km separation along Interstate Highway

10. The two sites have average annual PM concentrations of 11.00 and 12.05, respectively, which are below the National Ambient Air Quality Standards of 15 for PM. In this study, the collected PM compositional data at two monitoring sites were analyzed using PMF for source attribution. The PMF effectively identified nine possible common source-related factors for PM. The estimated source contributions for the common factor between the two sites were used to analyze spatial differences and to calculate correlations.

Sulfate-rich secondary aerosol was extracted by PMF which had the highest contribution to the PM mass in the region accounting for almost 42% and 44% of the total concentration at the two sites, respectively. Sulfate and nitrate mainly exist as ammonium sulfate and ammonium nitrate at the receptor sites. The soil factor has high source contribution peaks during the summer likely reflecting the intercontinental dust transport. The sea salt factor is clearly seen at the sites from the Gulf of Mexico. The chloride depleted marine aerosol was originated from sea-salt aerosol; however, it was separately identified because of the chloride loss during chemical reactions in the atmosphere. The correlation between the two sites is quite high for motor vehicle/road dust, and is moderate for sea salt and chloride depleted marine aerosol. Sulfate, motor vehicle/road dust, nitrate, soil, sea salt, and chloride depleted marine show regional characteristics with similar seasonal variation patterns at the two sites and significant correlation coefficients for these factors between the two sites. The regional factors account for about 67% and 68% of the PM mass concentration for the two sites, respectively. The correlations between the two sites are poor for cement/carbon-rich and wood smoke, and none for metal processing. A couple of metal processing facilities and a steel mill in Golden Triangle are clearly suggested of being related to the source of metal processing. The local factors on the average contribute 33% and 32% to the PM mass concentration for the two sites, respectively.

## ACKNOWLEDGES

This study was supported in part by the US EPA through project R-07-0159. The authors wish to thank Professor P.K. Hopke of Clarkson



University for helpful e-mail communications. The result of this research represents only the authors' assessments and does not reflect the funding agency's views on the air quality issues in this region.

## REFERENCES

- Anttila, P., Paatero, P., Tapper, U. and Järvinen, O. (1995). Source identification of bulk wet deposition in Finland by positive matrix factorization. *Atmos. Environ.*, **29**, 1705-1718.
- Chiou, P., Tan, W., Lin, C.J., Chu, H.W. and Ho, T.C. (2007). Atmospheric aerosol over a southeastern region of Texas: chemical composition and possible sources. *Environ. Model. Assess.*, doi: 10.1007/s10666-007-9120-8.
- Chueinta, W., Hopke, P.K. and Paatero, P. (2000). Investigation of sources of atmospheric aerosol at urban and suburban residential areas in Thailand by positive matrix factorization. *Atmos. Environ.*, **34**, 3319-3329.
- Efron, B. and Tibshirani, R.L. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.
- Hansen, D.A., Edgerton, E.S., Hartsell, B.E., Jansen, J.J., Kandasamy, N., Hidy, G.M. and Blanchard, C.L. (2003). The Southeastern aerosol research and characterization study: Part 1 – overview. *J. Air Waste Manage. Assoc.*, **53**, 1460-1471.
- Hopke, P.K. (1985). *Receptor Modeling in Environmental Chemistry*. New York: John Wiley & Sons.
- Hopke, P.K. (1991). *Receptor Modeling for Air Quality Management*. Amsterdam: Elsevier Science.
- Huber, P.J. (1981). *Robust Statistics*. New York: John Wiley.
- Juntto, S. and Paatero, P. (1994). Analysis of daily precipitation data by positive matrix factorization. *Environmetrics*, **5**, 127-144.
- Kim, E., Hopke, P.K. and Edgerton, E.S. (2003). Source identification of Atlanta aerosol by positive matrix factorization. *J. Air Waste Manage. Assoc.*, **53**, 731-739.
- Kim, E., Hopke, P.K. and Edgerton, E.S. (2004). Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. *Atmos. Environ.*, **38**, 3349-3362.
- Lee, E., Chan, C.K. and Paatero, P. (1999). Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmos. Environ.*, **33**, 3201-3212.
- Liu, W., Wang, Y.H., Armistead, R. and Edgerton, E.S. (2005). Atmospheric aerosol over two urban-rural pairs in the southeastern United States: chemical composition and possible sources. *Atmos. Environ.*, **39**, 4453-4470.
- Paatero, P. (1997). Least squares formulation of robust, non-negative factor analysis. *Chemomet. Intell. Lab. Syst.*, **37**, 23-35.
- Paatero, P. and Tapper, U. (1993). Analysis of different modes of factor analysis as least squares fit problems. *Chemomet. Intell. Lab. Syst.*, **18**, 183-194.
- Paatero, P. and Tapper, U. (1994). Positive matrix factorization: a non-negative factor models with optimal utilization of error estimates of data values. *Environmetrics*, **5**, 111-126.
- Paterson, K.G., Sagady, J.L., Hooper, D.L., Bertman, S.B., Carroll, M.A. and Shepson, P.B. (1999). Analysis of air quality data using positive matrix factorization. *Environ. Sci. Technol.*, **33**, 635-641.
- Polissar, A.V., Hopke, P.K., Malm, W.C. and Sisler, J.F. (1996). The ratio of aerosol optical absorption coefficients to sulfur concentrations, as an indicator of smoke from forest fires when sampling in polar regions. *Atmos. Environ.*, **30**, 1147-1157.
- Polissar, A.V., Hopke, P.K., Paatero, P., Kaufman, Y.J., Hall, D.K., Bodhaine, B.A., Dutton, E.G. and Harris, J.M. (1999). The aerosol at Barrow, Alaska: long-term trends and source locations. *Atmos. Environ.*, **33**, 2441-2458.
- Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C. and Sisler, J.F. (1998). Atmospheric aerosol over Alaska: 2. Elemental composition and sources. *J. Geophys. Res.*, **103**, 19045-19057.
- Polissar, A.V., Hopke, P.K. and Poirot, R.L. (2001). Atmospheric aerosol over Vermont: chemical composition and sources. *Environ. Sci. Technol.*, **35**, 4604-4621.
- Prospero, J.M. (2001). African dust in America. *Geotimes*, November, 24-27.
- Ramadan, Z., Eickhout, B., Song, X.H., Buydens, L.M.C. and Hopke, P.K. (2003). Comparison of positive matrix factorization and multilinear engine for the source apportionment of particulate pollutants. *Chemomet. Intell. Lab. Syst.*, **66**, 15-28.
- Ramadan, Z., Song, X.H. and Hopke, P.K. (2000). Identification of sources of Phoenix aerosol by positive matrix factorization. *J. Air Waste Manage. Assoc.*, **50**, 1308-1320.

Song, X.H., Polissar, A.V. and Hopke, P.K. (2001). Sources of fine particle composition in the northeastern US. *Atmos. Environ.*, **35**, 5277-5286.

US EPA (2002). *SPECIATE version 3.2*. US Environmental Protection Agency, Research Triangle Park, NC.

Watson, J.G., Chow, J.C. and Houck, J.E. (2001). PM chemical source profiles for vehicle exhaust, vegetative burning, geological material, and coal burning in northwestern Colorado during 1995. *Chemosphere*, **43**, 1141-1151.

Xie, Y.L., Hopke, P.K., Paatero, P., Barrie, L.A. and Li, S.M. (1999). Identification of source nature and seasonal variations of Arctic aerosol by positive matrix factorization. *J. Atmos. Sci.*, **56**, 249-260.

Yakovleva, E., Hopke, P.K. and Wallace, L. (1999). Receptor modeling assessment of PTEAM data. *Environ. Sci. Technol.*, **33**, 3645-3652.

Zheng, M., Cass, G.R., Schauer, J.J. and Edgerton, E.S. (2002). Source apportionment of PM in the southeastern United States using solvent-extractable organic compounds as tracers. *Environ. Sci. Technol.*, **36**, 2361-2371.