A QUANTITATIVE ASSESSMENT OF SOURCE CONTRIBUTIONS TO INHALABLE PARTICULATE MATTER POLLUTION IN METROPOLITAN BOSTON

GEORGE D. THURSTON* and JOHN D. SPENGLER

Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115. U.S.A.

Abstract—In this paper, source apportionment techniques are employed to identify and quantify the major particle pollution source classes affecting a monitoring site in metropolitan Boston, MA. A Principal Component Analysis (PCA) of particulate elemental data allows the estimation of mass contributions for five fine mass particle source classes (soil, motor vehicle, coal related, oil and salt aerosols), and six coarse particle source classes (soil, motor vehicle, refuse incineration, residual oil, salt and sulfate aerosols). Also derived are the elemental characteristics of those source aerosols and their contributions to the total recorded elemental concentrations (i.e. an elemental mass balance). These are estimated by applying a new approach to apportioning mass among various PCA source components: the calculation of Absolute Principal Component Scores, and the subsequent regression of daily mass and elemental concentrations on these scores.

One advantage of the PCA source apportionment approach developed is that it allows the estimation of mass and source particle characteristics for an unconventional source category: transported (coal combustion related) aerosols. This particle class is estimated to represent a major portion of the aerosol mass, averaging roughly 40 per cent of the fine mass and 25 per cent of the inhalable particle mass at the Watertown, MA site. About 45 per cent of the fine particle sulfur is ascribed to this one component, with only 20 per cent assigned to pollution from local sources. The composition of the coal related aerosol at this site is found to be quite different from particles measured in the stacks of coal-fired power plants. Sulfates were estimated to comprise a much larger percentage of the ambient coal related aerosol than has been measured in stacks, while crustal element percentages were much reduced. This is thought to be due to primary particle deposition and secondary aerosol accretion experienced during transport. Overall, the results indicate that the application of further emission controls to local point sources of particles would have less influence on fine aerosol and sulfate concentrations than would the control of more distant emissions causing aerosols transported into the Boston vicinity.

Key word index: Aerosols, dichotomous sampler, elemental composition, inhalable particles, particles, pollution sources, pollution transport, principal component analysis, selenium, source apportionment, sulfur, trace metals.

1. INTRODUCTION

In the air pollution field, Principal Component Analysis (PCA) and other related multivariate techniques such as Factor Analysis (FA) have been applied to arrays of pollution variables, to aerosol elemental composition data, or to spatial pollutant distributions in order to derive information about pollution sources influencing the data. Blifford and Meeker (1967) applied rotated FA to both particle elemental composition and mass data from 30 U.S. cities, thereby identifying independent components of particle variance as due to specific source classes (e.g. automotive exhaust). Hopke *et al.* (1976) applied FA to Boston total suspended particle elemental composition data in order to identify particle contributors in that city. Alpert and Hopke (1980) reanalyzed these Boston data using target transformation factor analysis to derive

quantitative estimates of the source contributions to total particle mass. Henry and Hidy (1979, 1982) have applied PCA to a mixture of air quality and meteorological variables in order to explain sulfate variations over time in four U.S. cities. Peterson (1970) used eigenvectors to determine characteristic occurrences of spatial sulfur dioxide patterns in St. Louis, while Cox and Clark (1981) have similarly identified spatial ozone patterns over the Eastern U.S. using FA.

In this work, PCA is applied to inhalable particle (IP) fine and coarse elemental concentration data collected over a two year period in Watertown, Massachusetts (a suburb of Boston). Based upon these PCA results, fine and coarse particle pollution sources affecting the monitoring site are identified. The particle mass contribution of each identified source is then estimated using a new technique: the computation of Absolute Principal Component Scores (APCS) for each sample, followed by the regression of sample particle mass concentrations on these APCS to derive each identified source's estimated mass contribution. The pollution source elemental profiles implied by the

* Present address: Institute of Environmental Medicine, New York University, Long Meadow Road, Tuxedo, NY 10987, U.S.A.

estimated source impacts are then derived (using linear regression) and compared with literature profiles. Subsequently applying the derived source elemental profiles to the source mass contributions at the monitor allows an elemental mass balance to be estimated. This indicates the contribution of each source to the ambient concentrations of each element considered in the analysis. As an aid to the reader, the overall PCA/regression source apportionment and source profile determination procedures are demonstrated using a simple synthetic data set created using two hypothetical sources and tracers (see Appendix). Conclusions are drawn regarding the PCA/regression technique employed in this work, as well as concerning the quantity and character of particle pollution sources affecting the Watertown, Massachusetts monitoring site.

2. METHODOLOGY

2.1. Data collection

The fine and coarse inhalable particle (IP) samples examined in this work were collected using a Beckman virtual impactor dichotomous sampler (Cat. No. 675900). Fine fraction aerosols collected had an aerodynamic diameter (d_a) less than 2.5 μm, while coarse particles had 2.5 μm < d_a < 15 μm. These samples were analyzed for 30 elements by X-ray fluorescence (XRF) spectrometry (Dzubay and Stevens, 1975). Fine and coarse masses were determined by Beta-gauge (Jaklevic et al., 1981). The 332 fine and coarse samples analyzed here were 24-h samples collected at least every other day between 4 June 1979 and 5 June 1981. The Watertown sampling site is situated in a high school athletic field located in a suburban neighborhood. As shown in Fig. 1, Watertown is within the Boston metropolitan area, approximately 10 km WNW of downtown Boston, MA. This particle sampling was conducted as part of the Harvard School of Public Health's ongoing prospective epidemiologic Air Pollution Respiratory
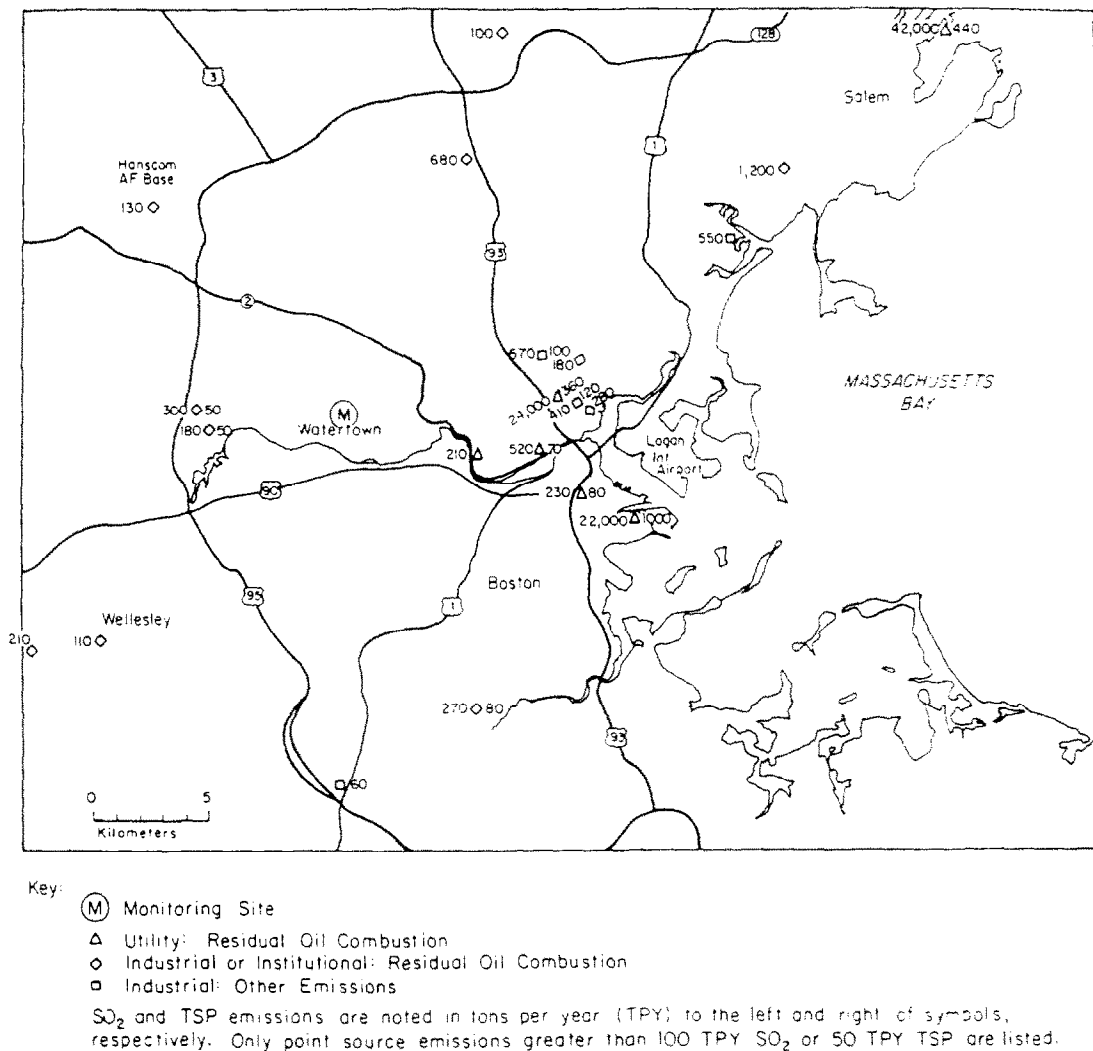

Key:

(M) Monitoring Site

△ Utility: Residual Oil Combustion
◇ Industrial or Institutional: Residual Oil Combustion
□ Industrial: Other Emissions

SO$_2$ and TSP emissions are noted in tons per year (TPY) to the left and right of symbols, respectively. Only point source emissions greater than 100 TPY SO$_2$ or 50 TPY TSP are listed.

Fig. 1. Map of Metropolitan Boston noting major local air pollution sources and Watertown monitoring site.

Health Study (Ferris et al., 1979). The health study particle sampling network and analysis techniques have been described in detail previously (Spengler and Thurston, 1983).

Prior to statistical analysis of the elemental and mass data, the samples were checked for validity. The sample records from this site were checked and flagged (where appropriate) for sampling problems including inappropriate flow rates, inappropriate sampling duration, and samples noted as suspicious by the operator. The dichotomous sampler quality control process, results, and criteria employed have been described in detail previously (Briggs et al., 1982). As a result of eliminating select samples flagged by the quality control procedures, the number of sample days was reduced from a total of 389 to the 332 measurement days ($m = 332$) considered in this analysis.

Elements thought to be unreliably determined by XRF at this site were also eliminated from the data set. This determination was made on each size fraction independently, and was based upon a qualitative knowledge of XRF elemental interference problems and a quantitative determination of the significance of elemental concentrations relative to their reported uncertainties. For the latter of these two criteria, elements having a mean concentration less than 1.65 times their reported mean standard deviation of measurement uncertainty were eliminated from the analysis, as they were deemed not able to be measured at the site (i.e. $P > 0.05$ that $\mu = 0.0$). Based upon these assessments of the reliability of measurement of the various XRF elements, 14 elements were selected for analysis of the fine data, while 15 were selected for the coarse fraction analysis. The elements selected, their respective mean concentrations, standard errors of the mean (S.E.) and mean standard deviations (S.D.) of measurement uncertainty are presented, by particle size fraction, in Table 1.

Since the time when the samples considered in this paper were collected and analyzed, one apparent bias in the coarse mass and elemental concentrations has been documented: the loss of coarse particles from filters during shipment from the field to the E.P.A. in North Carolina. A recent paper by Dzubay and Barbour (1983) has revealed that repeated dropping of packaged Teflon filter trays results in coarse particle mass losses of 10-53%, but no measurable losses from fine mass samples. In our own analysis of hand-carried coarse and fine samples, preliminary results confirm that the masses of shipped coarse particle samples average 30-35%

less than hand-carried samples. However, no such effect was found for fine mass samples (Thurston, 1983). To the extent that coarse particles have been lost during shipment, the results reported herein will understate the concentrations of the coarse aerosol, although fine aerosol results are apparently unaffected.

2.2. Statistical methods

Pollution sources affecting the Watertown site were statistically identified using varimax rotated Principal Component Analysis (PCA) of the XRF elemental concentration data. PCA is a statistical technique which can be applied to a set of variables in order to reduce their dimensionality. That is, to replace a large set of intercorrelated variables with a smaller number of independent variables. These new variables (components) are derived from the original variables, and are simply linear combinations of those variables. The first step in this analysis is to transform the elemental data into a dimensionless standardized form

$$Z_{ik} = \frac{(C_{ik} - \overline{C}_i)}{\sigma_i}, \qquad (1)$$

where $i = 1, 2, \ldots n$, the total number of elements in the analysis; $k = 1, 2, \ldots m$ the total number of observations; Z_{ik} (often termed the Z-score) is the standardized value of element i for observation k; C_{ik} is the concentration of that element for that observation; $\overline{C}_i$ is the mean concentration for the ith element over all observations; and, σ_i is the standard deviation of the distribution of concentrations of the ith element. It is important to note at the outset that, for multivariate analyses such as proposed in this work, sufficient degrees of freedom should be available in the model. Thus, the dataset employed must have many more observations than variables (e.g. $m \gtrsim n + 50$) if stable results are to be derived.

The PCA assumes that the total concentration of each element is made up of the sum of elemental contributions from each of j pollution source components. Hence,

$$Z_{ik} = \sum_{j=1}^{p} W_{ij} P_{jk}, \qquad (2)$$

where P_{jk} is the jth component's value for observation k; $j = 1, \ldots, p$, the number of pollution sources influencing the data and W_{ij} is the coefficient matrix of the components.

Table 1. Mean Watertown, MA fine and coarse aerosol elemental and mass concentrations (ng m^{-3})

	Fine aerosol ($m = 332$)			Coarse aerosol ($m = 332$)		
	Mean	S.E. of mean	Mean S.D. of meas. uncert.	Mean	S.E. of mean	Mean S.D. of meas. uncert.
Al	*	*	17.1	366.	25.6	51.7
Si	104.	6.96	13.7	1021.	66.2	158.
S	1840.	74.6	108.	201.	9.53	120.
Cl	86.2	11.1	7.12	306.	45.8	66.2
K	76.2	2.74	3.94	112.	6.62	10.3
Ca	41.9	2.29	2.31	214.	15.2	13.2
Ti	*	*	4.29	25.9	2.08	5.22
V	22.1	1.10	2.89	3.34	0.33	2.46
Mn	3.61	0.17	1.14	5.81	0.403	1.29
Fe	74.7	3.87	4.15	281.	17.1	15.5
Ni	8.57	0.39	0.86	2.44	0.13	0.66
Cu	16.1	1.01	1.40	10.4	0.69	1.28
Zn	26.5	1.05	1.71	12.2	0.64	1.30
Se	0.595	0.046	0.273	*	*	0.208
Br	85.7	4.39	4.18	25.7	1.55	1.94
Pb	326.	13.2	15.6	75.6	4.57	5.95
Mass (μg m^{-3})	17.4	0.52	1.60	8.57	0.45	1.60

* Mean concentration below detection (1.65 × mean S.D. of measurement uncertainty).

Henry and Hidy (1979) have noted that Equation (2) may be inverted, yielding (in matrix terms)

$$[P]_{j \times t} = [B]_{j \times i}[Z]_{i \times t}$$ (3)

where

$$[B]_{j \times i} = [W]_{j \times i} \; \lambda_j$$ (4)

and λ_j is the eigenvalue associated with P_j.

The PC scoring matrix $[B]$ is derived so that the first principal component (PC1) explains as large a per cent of the original variables' total variance as possible. The coefficients for the second principal component (PC2) are, in turn, chosen so that it explains as large a per cent of the remaining variance in the original variables (i.e. not explained by PC1), subject to the restriction that PC1 and PC2 are uncorrelated. In general, the coefficients for PC(j) are chosen so that PC(j) explains as much of the remaining variance (i.e. not explained by PC1–PC($j-1$)), subject to the constraint that PC(j) be uncorrelated with PC1–PC($j-1$).

The PC equation coefficients $[B]$ are mathematically derived from the correlation matrix

$$[R]_{i \times i} = [Z]_{i \times t}[Z]'_{t \times i}$$ (5)

Since the objective of PCA is to find orthogonal (uncorrelated) components, the correlation matrix $[R]$ is diagonalized. The diagonalization finds a matrix Q such that,

$$[Q^{-1}]_{i \times i}[R]_{i \times i}[Q]_{i \times i} = [\Lambda]_{i \times i},$$ (6)

where $[\Lambda]$ is a diagonal matrix (i.e. a matrix with no off diagonal cross-correlation terms) of eigenvalues arranged in descending order of magnitude; and Q contains the corresponding eigenvectors which diagonalize the correlation matrix. By definition, these eigenvectors are the matrix $[B]$, which can be used in (3) to derive the PC score matrix $[P]$ from the $[Z]$ matrix.

The primary objective of applying PCA is to derive a small number of components which explain a maximum of the variance in the data. Initially, the PCA results in as many PCs as there are original variables (n). Usually, however, only a limited number of these uncorrelated PCs (e.g. five or six) are required to explain virtually all of the variance in a data set of fifteen or more original (intercorrelated) variables. In order for this reduction in the dimensionality to be useful, the new variables (components) must have simple substantive interpretations. Empirically, it has been found that unrotated PCs are often not readily interpretable since they each attempt to explain all remaining variance in the data set (Harris, 1975). This calculation results in a number of sources of variance being grouped together. For this reason, a limited number of components ($p < n$) are usually subjected to rotation using a criteria such as varimax. After PCA rotation, the resulting components have been found to often be more representative of individual underlying sources of variation (e.g. Walsh and Richman, 1981). This, in turn, results in more interpretable and useful PCs. The merits of PCA rotation in analyzing air pollution data have been discussed in the literature by Thurston (1981), Henry and Hidy (1981) and Hopke (1982).

Based on the above considerations, the varimax rotated PCs of the fine and coarse data have been used in this determination of the fine and coarse particle sources affecting our Watertown site. The p components rotated had, as recommended by Hopke (1982), eigenvalues greater than 1.0 after rotation. The varimax rotation constitutes a maximization of the variance of the communality-normalized loadings (correlations) presented in the columns of a PC pattern. Such rotations tend to drive variable loadings toward either zero or one on a given component. They preclude the emergence of a strong general component by making the variance explained by the individual components more equal, resulting in orthogonal components which are often more readily identifiable as specific source components.

An examination of the rotated component loadings on the original elements allowed the identification of the PCs as

pollution sources affecting the data (e.g. oil combustion, automotive emissions, etc.). Subsequently, the mass contribution (in $\mu g \, m^{-3}$) of each source component to the Watertown site's total particle pollution levels was estimated. The first step in the derivation of source impacts is to calculate component scores for each sample. Rotated PC coefficients, B^*, are calculated by applying the rotation transformation matrix $[T]$ to $[B]$.

$$[B]^*_{p \times n} = [B]_{p \times n}[T]_{n \times n}$$ (7)

Rotated PC scores are computed using this transformed $[B]$ matrix,

$$[P]^*_{p \times m} = [B]^*_{p \times n}[Z]_{n \times m}$$ (8)

These PC scores are correlated with their respective pollution sources impacting the site (i.e. a higher component score, P^*_{jk}, implies a higher pollution impact by the pollution source j during observation k). However, because they are computed from the normalized elemental concentrations, Z_{ik}, they too are normalized. Each component indicates deviations from the mean source impact, with each having a mean of zero. Thus, while the PC scores are correlated with the identified pollution source impacts, they are not proportional to these pollution impacts.

Henry and Hidy (1979) showed that the regression of a dependent variable Y_k on the daily scores of components P_{jk} gives the formula

$$Y_k = \overline{Y} + \sum_{j=1}^{p} \zeta_j P_{jk}.$$ (9)

where $\overline{Y}$ equals the mean of Y_k. If the dependent variable Y_k is the total mass (in $\mu g \, m^{-3}$), then ζ_j are the conversion coefficients of the non-dimensional PC score deviations into mass deviations from the mean source impact. Since the components are not scored as deviations from zero, but instead as deviations from the mean, this results in the presence of $\overline{Y}$ in the equation. This, in turn, prevents a direct apportionment of the total particle mass at the site to the normalized pollution source components.

Previous work by Alpert and Hopke (1980) has addressed the problem of deriving absolute components related to absolute zero by initializing a correlation about the origin (instead of about the mean) and normalizing the data to the mean of all elements in each sample (Q analysis) instead of to the mean of all samples for each element (R analysis). This allowed a retention of the information regarding the relative size of each element in absolute terms (i.e. distance from zero). The resultant vectors were then target transformed to align (to a least-squares fit) with known pollution source elemental profile vectors. Henry (1977) also attempted to target factors to assumed single element trace compositions, but the data employed were very limited (14 observations). However, existing Target Transformation Factor Analysis (TTFA) techniques are not readily executed with conventional statistical packages, which do not provide for target rotations.

In this work, we have applied a simple PCA pollution apportionment procedure which can be executed on conventional, generally available statistical packages, such as SAS (Barr et al., 1979). The PCA has been conducted using a conventional R analysis of elemental correlations about their means. However, the 'absolute zero' PC score has subsequently been estimated for each PC by separately scoring an extra 'day' wherein all the elemental concentrations are zero. This is accomplished by deriving the Z-score for absolute zero concentrations,

$$(Z_0)_i = \frac{0 - C_i}{\sigma_i} = \frac{-C_i}{\sigma_i}$$ (10)

and then calculating the rotated absolute zero PC scores, P^*_0, for each of p components,

$$P_{0_p}^* = \sum_{i=1}^{n} B^*_{pi}(Z_0)_i$$ (11)

These estimates of the PC scores for each component at absolute zero are then used to estimate Absolute PC Scores (APCS) for each component on each day as follows

$$[APCS]^*_{p \times j} = [P]^*_{p \times j} - [P_0]^*_{p \times j} \qquad (12)$$

where the j columns of $[P_0]^*$ are all identically equal to the values calculated in (11). It can be proved in a straightforward manner that the calculation of $[APCS]^*$ according to (12) gives the exact score which would be achieved had the original scoring been executed using unnormalized data.

Regressing daily fine and coarse mass data on these APCS gives estimates of the coefficients which convert the APCS into pollutant source mass contributions (in $\mu g\ m^{-3}$) for each sampling day, as follows

$$M_k = \zeta_0 + \sum_{j=1}^{p} \zeta_j APCS^*_{jk}, \qquad (13)$$

where M_k is the particle mass recorded (in $\mu g\ m^{-3}$) during observation k; $APCS^*_{jk}$ is the rotated absolute component score for component j on observation k; $\zeta_j APCS^*_{jk}$ is the particle mass contribution on observation k (in $\mu g\ m^{-3}$) made by the pollution source identified with component j; and ζ_0 is the particle mass contribution (in $\mu g\ m^{-3}$) made by sources unaccounted for in the PCA. If the PCA is successful, ζ_0 should tend toward zero. This regression approach to source apportionment is analogous to previous work by Kleinman et al. (1980), except that multi-element, orthogonal PCs are used as tracers of pollution impacts, instead of single element tracers.

For each of the pollution sources identified in the fine and coarse mass PCA, it is possible to derive information regarding the elemental composition of particles from that source. As shown by Dattner et al. (1982), the individual weighted regression of each element's recorded daily concentrations on the predicted mass contributions from all sources during those days yields estimates of the content of that element in each source's mass, as follows

$$C_k = a_0 + \sum_{j=1}^{p} a_j S_{jk}, \qquad (14)$$

where C_k is the concentration of an element i during observation k; S_j is the mass concentration of source j on day $k [= \zeta_j APCS^*_{jk}$ from Eqn. (13)]; and a_j is the mean mass fraction of source j's particles represented by the element. Repeating this weighted least square regression for each of the $i = 1, 2, \ldots n$ elements considered in this analysis, by size fraction, allows an estimation of the elemental source profiles of each pollution source identified in each particle size

category (a_{ij}). Moreover, the product of these estimates of the mean elemental mass fractions and the prior estimates of the mean component mass impacts yields estimates of the contribution by the jth source to the ambient concentrations of element i during observation $k (= a_{ij} S_{jk})$. Thus, estimates can be derived for both the elemental composition of the identified source particles and their contributions to the ambient elemental concentrations.

3. RESULTS

3.1. Principal Component Analysis

3.1.1. *Fine mass elemental PCA.* Varimax rotated PCA was applied to the fine elemental data, yielding the components presented in Table 2. This component pattern contains the correlation (r) of each element with each component. These six components account for 87 per cent of the variance in the data set, with roughly 30 per cent (3.8/13 variables) accounted for by the first component.

The first component appears to represent soil aerosol impacts at this site, since it is highly correlated with the crustal elements Si, Fe and Ca. The interpretation of the second component is still clearer, since it is correlated with both Br and Pb, two elements highly enriched in emissions from leaded automobiles (e.g. see Cooper and Watson, 1980).

The third component in Table 2 is most dependent upon selenium and sulfur, and appears to incorporate the influence of coal combustion transport products. In the U.S., coal combustion is the primary anthropogenic source of atmospheric selenium emissions, and Se is found in higher ambient concentrations where coal is burned (NAS, 1980) (Spengler and Thurston, 1983). Indeed, emission inventories of anthropogenic Se emissions indicate that coal combustion emissions account for between 85% (Eimutis et al., 1978) and 62% (NAS, 1976) of U.S. Se emissions with smelters accounting for most of the remainder (9–26%). Residual oil also contains Se, though roughly an order of magnitude less than coal (NAS, 1976; Hashimoto

Table 2. Varimax rotated Principal Component Analysis of Watertown fine mass elemental data
($m = 332$)

Element	Soil aerosols	Motor vehicle aerosols	Coal related aerosols	Residual oil aerosols	Salt aerosols
Si	0.97	0.05	−0.01	0.03	0.12
S	0.05	0.04	0.86	0.08	−0.11
Cl	0.26	0.06	−0.09	0.14	0.93
K	0.66	0.33	0.24	0.34	0.26
Ca	0.92	0.02	0.00	0.14	0.11
V	0.18	0.33	0.14	0.86	0.17
Mn	0.68	0.40	0.33	0.19	0.03
Fe	0.95	0.10	0.14	0.14	0.10
Ni	0.20	0.27	0.13	0.89	0.04
Zn	0.18	0.48	0.56	0.32	0.23
Se	0.11	0.10	0.88	0.08	0.01
Br	0.16	0.93	0.00	0.24	0.10
Pb	0.06	0.90	0.20	0.29	−0.05
Eigenvalue	3.8	2.2	2.1	2.0	1.1

and Winchester, 1967), and the emission inventories noted above indicate oil to contribute only 1–5 °$_0$ of U.S. anthropogenic Se emissions. This Se factor has appeared both in past work with Boston area aerosols (Hopke et al., 1976) and in analysis of rural New York State aerosols (Parekh and Husain, 1981). For the former of these two references, the actual observations were obtained from the authors. During the high Se days, high values were uniformly recorded at sites throughout the Boston metropolitan area, and five of the six highest days occurred during days of winds from the west or southwest, according to surface observations at Boston's Logan Airport. These results would indicate the source not to be localized within Boston, and a source who's pollution impacts the Boston area during W–SW winds. In the latter of these references, the Se–S factor was shown to also be associated with arsenic and to be highest during days when the sampled air parcel had passed over the Ohio Valley region. The authors further identified this component as coal combustion related. Navarre et al. (1981) also noted an S–Se–As association in aerosols collected in rural Belgium, identifying coal combustion as one possible source. However, there was essentially no coal combustion occurring in the local Boston metropolitan area during the Watertown sampling. In light of all of the above insights, this Se–S component seems likely to be associated with transported fine particles of coal combustion origins.

The final two components in Table 2 can be interpreted, based upon their correlations with key trace elements, as residual oil combustion (Ni, V) and salt (Cl). It is interesting to note that the residual oil component clearly separated itself from the Se–S component, indicating these element groups to originate from separate sources. Both selenium and sulfur are poorly correlated with vanadium within this data set (r = 0.25 and 0.19, respectively). In traditional PCA applications, components 4 and 5 would often be eliminated from consideration, as they had unrotated

eigenvalues less than 1.0 (i.e. explain less variance than the average single original variable). However, they are displayed here because they attain eigenvalues exceeding 1.0 after rotation. Originally, elemental copper concentrations were also included in the PCA, resulting in another component (loaded on copper alone). This element was eliminated from the fine aerosol PCA analysis, however, as it was deemed overly influenced by an 'artificial' Cu source previously documented by Patterson (1980): the routine wearing of copper motor brushes in collocated high-volume samplers.

Before progressing to a consideration of the coarse mass source components, it is important to note that not all of the mass attributable to these source components need be derived solely from the source for which an elemental tracer is measured. For example, although the motor vehicles component was identified based upon tracers for leaded fuel combustion, it is recognized that other motor vehicle emissions will likely impact the monitor simultaneously, so the Pb and Br concentrations become tracers for all motor vehicle impacts. Similarly, sources which impact the monitor simultaneously with the tracer thought to be derived from coal emissions (Se) will have their mass impacts assigned to this component. As a result, the Se associated component has been called coal related aerosols in this analysis, rather than primary coal combustion aerosols.

3.1.2. Coarse mass elemental PCA. The varimax rotated coarse mass elemental data PCA component structure is presented in Table 3. The six components noted account for over 95 per cent of the total variance in the original fifteen variables. The interpretation of many of these components is similar to that for the fine data, as indicated by the column headings. In this case, copper from the hi-volume sampler brushes may be a smaller percentage of coarse Cu, as the PCA is able to differentiate a coarse Cu–Zn particle component which appears to be refuse incineration, a source known to be enriched in Zn (Greenberg et al., 1978)

Table 3. Varimax rotated Principal Component Analysis of Watertown coarse mass elemental data (m = 332)

Element	Soil aerosols	Motor vehicle aerosols	Refuse burning aerosols	Residual oil aerosols	Salt aerosols	Sulfate aerosols
Al	0.97	0.06	0.06	0.14	−0.01	0.07
Si	0.98	0.04	0.06	0.12	0.00	0.05
S	0.17	0.19	0.10	0.23	0.07	0.93
Cl	0.00	0.04	0.00	0.13	0.99	0.06
K	0.97	0.05	0.04	0.11	0.05	0.10
Ca	0.94	0.05	0.05	0.14	0.02	0.05
Ti	0.97	0.04	0.05	0.16	0.00	0.04
V	0.25	0.20	0.07	0.89	0.10	0.13
Mn	0.97	0.04	0.07	0.11	0.00	0.03
Fe	0.97	0.07	0.09	0.17	0.00	0.08
Ni	0.35	0.27	0.24	0.77	0.11	0.20
Cu	0.00	0.06	0.95	0.00	0.01	0.00
Zn	0.21	0.16	0.87	0.24	−0.01	0.12
Br	0.05	0.97	0.09	0.15	0.03	0.09
Pb	0.08	0.96	0.12	0.20	0.02	0.11
Eigenvalue	6.8	2.0	1.8	1.7	1.0	1.0

and to be a major source of atmospheric Zn emissions (Nriagu and Davidson, 1980). Both Cu and Zn have been suggested as elemental tracers for incineration particles in New York City (Kleinman *et al.*, 1980). A Zn component was previously identified in Boston aerosol by Hopke *et al.* (1976). Since incineration is not now permitted in the Boston area, it was surprising to still find evidence of this practice. However, conversations with local pollution inspectors revealed that refuse incineration is still conducted at a number of housing projects in the Boston area (DEQE, 1982a). Finally, the coarse sulfate component is identified as such since there are no other detectable trace elements associated with coarse S. However, it is known that a fraction of fine mass aerosol "crosses over" to the coarse mass filter in a dichotomous sampler (e.g. see McFarland *et al.*, 1979). Since Table 1 indicates coarse S to average only one tenth the fine S value, it is suspected that cross-over contamination of fine sulfur onto the coarse filter sample may be a major portion of the noted "coarse" sulfur.

3.2. PCA *Source apportionment of particle mass*

For each fine and coarse elemental component identified above, the PC score was computed for each sample. These were then adjusted to Absolute Principal Component Scores (APCS) as detailed above [Eqns (10)–(12)]. Mass loadings for the samples (in $\mu g\,m^{-3}$) were then regressed on the APCS according to (13). The resultant regression coefficients were then employed to convert the daily absolute factors scores to estimates of the PC mass contributions $(= \zeta_j APCS^*_{jk}$ in $\mu g\,m^{-3})$.

The weighted least squares regression of the coarse masses on the elemental APCS yielded significant regression coefficients $(P < 0.05)$ for all components considered. The model yielded an intercept (i.e. impact of unidentified sources) of $0.7\,\mu g\,m^{-3}$ and an overall correlation (r) of 0.97. A similar analysis of the fine

fraction aerosol indicated that this aerosol was not fit quite as well by the identified sources, having an unexplained mass of $3.8\,\mu g\,m^{-3}$ and an overall r of 0.90. The salt component in the fine aerosol was not found to contribute significantly to the aerosol mass. Inserting the mean APCS for each dataset into these regression equations results in the mean source contributions noted in Table 4 for the fine, coarse and total IP masses.

It should be noted that the estimated standard errors (S.E.) of the mean source impacts presented in Table 5 are likely to underestimate the true standard errors. This is because these S.E. estimates are derived in standard statistical regression packages by assuming that the independent variables (the APCS) are known without error. However, they are estimated quantities derived from the elemental concentrations, and are indeed subject to uncertainty. Thus, the S.E. estimates should be considered estimates of the minimum possible standard errors. The estimation of the standard errors of PCA score calculations and their incorporation into the calculation of subsequent regression coefficient standard errors is a step which would yield larger, more accurate, estimates of the confidence intervals around these source contribution estimates.

The results in Table 4 indicate that a majority of the IP mass is in the fine fraction (67%, on average). It must be remembered that, as discussed above, initial investigations of particle losses during shipping indicate that the coarse particle losses may range from 10 to 53%. Of the fine mass, the largest portion is indicated as being attributable to the coal related component (40% of fine mass, 27% of IP). As shall be determined later, this mass impact is primarily due to the sulfates associated with this component. Soil, auto emissions, and oil combustion each average between 10 and 15 per cent of the fine mass. The estimated coarse mass impacts noted in Table 4 are dominated by the soil component. Soil makes up 57 per cent of the

Table 4. Mean fine, coarse and inhalable mass contributions from PCA identified sources $(m = 332)$

| | Mean particle mass source contributions ($\mu g\,m^{-3} \pm$ S.E. of the mean) | | |
	Fine	Coarse	IP
Soil	1.9 (± 0.1)	4.9 (± 0.4)	6.8 (± 0.4)
Motor vehicles	2.5 (± 0.2)	0.6 (± 0.05)	3.2 (± 0.2)
Oil	2.1 (± 0.2)	0.2 (± 0.08)	2.3 (± 0.2)
Refuse burning	*	0.5 (± 0.03)	0.5 (± 0.03)
Salt	†	0.4 (± 0.08)	0.4 (± 0.08)
Coal related	6.9 (± 0.4)	*	6.9 (± 0.4)
Coarse sulfate	*	1.2 (± 0.1)	1.2 (± 0.1)
Unexplained	3.8 (± 0.5)	0.7 (± 0.2)	4.5 (± 0.5)
Total predicted	17.3 (± 0.5)	8.5 (± 0.4)	25.8 (± 0.7)
Total observed	17.4 (± 0.5)	8.6 (± 0.4)	25.9 (± 0.8)
Pred.–Obs. mass correlation	0.90	0.97	0.94

* Not identified by the PCA as a particle source for this size range.
† Not statistically significant contributor (at $P = 0.05$) to mass regression.

Table 5. PCA-derived trace element source profiles (% of mass ± S.E.)

	Motor vehicle aerosols		Oil aerosols		Soil aerosols		Coal related aerosols	Refuse aerosols	Salt aerosols	Sulfate aerosols
	fine	coarse	fine	coarse	fine	coarse	fine	coarse	coarse	coarse
Al		1.8 ± 0.6		4.7 ± 0.31	4.4 ± 0.1	5.8 ± 0.2		3.6 ± 0.5		2.0 ± 0.3
Si	0.20 ± 0.04	5.2 ± 0.8	0.18 ± 0.04	8.7 ± 0.4	3.9 ± 0.6	14.7 ± 0.4		13.0 ± 0.8		4.7 ± 0.5
S	6.7 ± 1.0	3.8 ± 0.1	5.4 ± 0.9	2.6 ± 0.03	0.55 ± 0.1	0.41 ± 0.004	12.3 ± 0.5	2.4 ± 0.04	0.95 ± 0.04	11.5 ± 0.05
Cl	0.23 ± 0.09	3.4 ± 0.2	0.25 ± 0.07	6.0 ± 0.1					46. ± 1.	2.8 ± 0.1
K	0.44 ± 0.05	0.81 ± 0.07	0.32 ± 0.04	1.1 ± 0.1	1.4 ± 0.02	1.6 ± 0.02	0.16 ± 0.02	1.1 ± 0.1	0.49 ± 0.06	0.93 ± 0.05
Ca		1.2 ± 0.2	0.15 ± 0.02	1.9 ± 0.1	1.55 ± 0.05	3.0 ± 0.1	0.022 ± 0.007	2.6 ± 0.3		1.0 ± 0.1
Ti		0.073 ± 0.03		0.35 ± 0.02		0.48 ± 0.01		0.2 ± 0.05		0.053 ± 0.023
V	0.20 ± 0.01	0.12 ± 0.01	0.54 ± 0.01	0.35 ± 0.01	0.13 ± 0.01	0.018 ± 0.001	0.033 ± 0.003	0.039 ± 0.013	0.041 ± 0.007	0.048 ± 0.006
Mn	0.034 ± 0.003	0.038 ± 0.007	0.016 ± 0.003	0.057 ± 0.005	0.085 ± 0.004	0.095 ± 0.002	0.012 ± 0.001	0.091 ± 0.011		0.022 ± 0.005
Fe	0.17 ± 0.02	2.2 ± 0.2	0.27 ± 0.02	2.9 ± 0.1	2.4 ± 0.1	4.1 ± 0.1	0.14 ± 0.01	4.4 ± 0.2		1.4 ± 0.1
Ni	0.069 ± 0.003	0.066 ± 0.004	0.19 ± 0.004	0.10 ± 0.003	0.59 ± 0.003	0.011 ± 0.001	0.015 ± 0.001	0.085 ± 0.007	0.016 ± 0.003	0.032 ± 0.003
Cu	0.084 ± 0.02	0.059 ± 0.016			0.05 ± 0.02		0.029 ± 0.006	1.9 ± 0.04		
Zn	0.25 ± 0.02	0.18 ± 0.02	0.25 ± 0.01	0.17 ± 0.01	0.16 ± 0.02	0.034 ± 0.002	0.12 ± 0.01	1.3 ± 0.05		0.095 ± 0.011
Se							0.0089 ± 0.0003			
Br	1.9 ± 0.1	3.0 ± 0.03	0.55 ± 0.03	0.27 ± 0.01	0.40 ± 0.03	0.018 ± 0.001	0.054 ± 0.008	0.43 ± 0.02	0.056 ± 0.009	0.19 ± 0.01
Pb	7.2 ± 0.2	7.6 ± 0.2	2.2 ± 0.1	1.2 ± 0.04	1.1 ± 0.1	0.10 ± 0.01	0.73 ± 0.04	1.6 ± 0.1	0.10 ± 0.04	0.60 ± 0.043

coarse mass. The next largest contributor is sulfates, which averages 14 per cent of the coarse mass. All other sources each contribute 8 per cent or less of the coarse mass. The possible problem of coarse particle losses would not be expected to drastically change these mass source contribution percentages.

3.3. Estimation of source particle characteristics

The percentage elemental compositions resulting from the weighted regression of elemental concentrations on the source mass contributions [i.e. Eqn (14)] are displayed in Table 5. The profiles were esimated for both the fine and coarse size fractions. Elements missing from a profile were not found to have statistically significant regression coefficients when regressed on that source category's mass impacts.

Comparing the fine and coarse motor vehicle profiles in Table 5, it is seen that the coarse profile has significantly higher percentages of the crustal elements (e.g. Si, Ca and Fe). Since the source impacts input into (14) were also derived via regression (13), they also reflect any mass impacts correlated with the identified source component. In this case, it seems probable that coarse road dust suspended by traffic is correlated with the motor vehicle impacts, increasing the total predicted auto impact and raising the crustal element percentages.

The Pb: Br: Cl ratios for the derived fine and coarse motor vehicle profiles are 1.0:.26:.03 and 1.0:.39:.45, respectively, as compared to typical leaded auto exhaust ratios of 1.0:.38:.17 (Biggins and Harrison, 1979). The relatively low fine Cl and Br values are likely due to volatilization of these elements off particles after emission (Robbins and Snitz, 1972). Volatilization occurs especially rapidly in this size fraction because fine particles have a higher surface to volume ratio than coarse particles. In contrast, the coarse automotive component appears to contain excess Cl, which may reflect an impact of road salting similar to that previously noted for road dust. Coarse Cl has been shown to be seasonal at this site, with the highest impacts during the winter (road salting) months (Spengler and Thurston, 1983).

Comparison of the derived fine and coarse road vehicle elemental profiles with those in the published literature indicate qualitative and quantitative agreement. For example, total particle data by Pierson and Brachaczek (1983) in Table 6 indicate a Pb:Br:Cl ratio of 1.0:.34:.03, which lies between that derived for the fine and coarse fractions for Br, and equals the value derived for fine fraction Cl/Pb. Moreover, the motor vehicle Pb fraction indicated by the literature is 7 per cent, very close to the average 7.4 derived in this work. These values are considerably less than the per cent of Pb indicated in Table 6 for leaded gasoline burning automobiles (21 per cent). These differences are due to the influences of fugitive road dust, diesel vehicles, and unleaded gasoline fueled autos (which collectively generate a mixture of road dust and exhaust particles enriched in carbon and sulfates, but

no Pb, Br or Cl). This tends to lower the percentages of Pb, Br and Cl in the total motor vehicle related aerosol (while raising others, such as $°_oS$), in agreement with the results in Table 5 (except coarse Cl, as discussed above). It is interesting to note that analogous tracer regression work by Kleinman et al. (1980) on New York City ambient aerosols indicated a total automotive particle lead content of 9.0 per cent in 1975. The quantitative differences which are noted between the results of this analysis and leaded auto exhaust literature data are attributable to the fact that this statistical approach incorporates all aerosol mass generated by vehicular traffic, not just the primary leaded auto exhaust for which we have a measurable elemental tracer.

The oil component profiles shown in Table 5 indicate lower percentages of V and Ni (residual oil tracers) than reported in literature profiles (Table 6). The V:Ni ratio for both coarse and fine profiles is 1.0:0.35, which is higher than that reported by Mroz (1976) (1.0:0.17), but lower than that reported by Watson (1979) (1.0:1.6 for fine aerosol, and 1.0:0.78 for coarse). The absolute V percentages reported by these two sources are 7.0 per cent (Mroz, 1976) and 3.4, 1.8 for fine, coarse emissions (Watson, 1979). These values are higher than the 0.54%, 0.35% V for fine and coarse particles found in this work. Similarly, the Ni percentages reported in Table 6 are higher than the fine and coarse percentages for Ni estimated here. Thus, the estimated percentage values for the primary residual oil tracers (V and Ni) appear low, but the Ni V ratio appears to lie within the range of values reported, and the shift in per cent V and Ni from fine to coarse agrees with the literature trends.

The reason for the low percentages of V and Ni was investigated. Discussions with the Massachusetts Department of Environmental Quality Engineering revealed that the majority of residual fuel burned in Eastern Massachusetts during the study period was low sulfur fuel (0.5–1.0% S) (DEQE, 1982b). Fuel oil desulfurization lowers not only the fuel sulfur content, but also reduces trace element concentrations in the fuel (NAS, 1974). Analyses obtained for fuels burned in Boston indicated that 2.2% S residual has roughly a 290 ppm V content, while 1.0% S has an average 44 ppm V (DEQE, 1982b). Thus, for a 55 per cent reduction in sulfur content, there is about an 85 per cent reduction in V content. Kleinman et al. (1980) note this reduction in fuel oil V with reductions in S content, and derive a regression coefficient indicating a 0.97% V content in oil particles in New York City in 1975. The removal of trace elements during fuel desulfurization would appear to explain the low trace element percentages estimated for residual oil particles in the Boston area.

A comparison of the fine and coarse soil components' elemental profiles in Table 5 with the literature values in Table 6 indicates that the coarse profile agrees with expected values better than the fine profile does. Size separated soil data by Watson (1979)

Table 6. Literature trace element source profiles (% of mass)*

	Motor vehicles		Residual oil			Soil			Road dust		Coal	Refuse	Salt
	All traffic	Leaded auto											
	Total mass[1]	Total mass[2]	Fine mass[3]	Coarse mass[3]	Total mass[4]	Fine mass[3]	Coarse mass[3]	Total mass[5]	Fine mass[3]	Total mass[3]	Total mass[6]	Total mass[7]	Total mass[2]
Al	0.8	0.043	0.53	1.47		11.7	7.2	5.6	8.8	6.6	13.0	1.4	
Si	1.4	0.075	0.96	3.7		25.4	30.		22.	28.			
S	2.2	0.07	16.	2.4		0.069			0.37				3.0
Cl	0.2	5.4		0.74	3.7			0.018			0.026	20.	55.
K	0.2		0.28		0.1	1.0	2.2	1.1	1.0	3.0	1.7		1.0
Ca	1.1		1.6	3.6	2.4	0.93	1.4	0.45	2.44	1.0	1.2	0.24	1.0
Ti	0.013		0.11	0.14		0.76	0.82	0.36	0.64		0.78		
V	0.0003		3.4	1.8	7.0	0.025	0.017	0.0073	0.23	0.027	0.039		
Mn	0.065		0.046	0.027	0.03	0.20	0.085	0.09	0.12	0.10	0.039	0.073	
Fe	0.54	0.25	3.0	1.5	0.84	6.8	4.8	3.5	6.0	5.7	9.1	0.65	
Ni			5.4	1.4	1.2	0.0092		0.0031	0.009		0.026	0.015	
Cu	0.04	0.004	0.075	0.086	0.24	0.02		0.0061	0.03		0.052	0.17	
Zn	0.125	0.021	0.40	0.05	0.49	0.041		0.0078	0.11		0.065	12.0	
Se					0.01			0.00001			0.02	0.0037	
Br	2.4	8.2	0.013		0.02			0.00067	0.02	0.008	0.052	0.12	0.55
Pb	7.0	21.1	0.11		0.12	0.026		0.0011	0.37	0.51	0.052	8.1	

* These data are for elements emitted as particles. Gaseous emissions which condense or react in the atmosphere to form particles were generally not included.
1. Pierson and Brachaczek (1983) 2. Cass and McCrae (1983) 3. Watson (1979) 4. Mroz (1976) 5. Thomae (1977) 6. Kowalczyk et al. (1978) 7. Greenberg et al. (1978)

indicate the highest mass percentages to be in the crustal elements (e.g. Si, Al, Fe, K, Ca and Mn). The estimated coarse soil profile in Table 5 shows a similar trend, though the estimated crustal percentages are lower for all noted elements but Ca and Mn. In the fine mass fraction, lower crustal element percentages are estimated than for the coarse profile, but significantly higher S, Br and Pb percentages are estimated. The reason for these high values [orders of magnitude higher than those reported for the fine soil fraction of Watson (1979)] is not clear, but they may, in part, be manifestations of collinear road dust impacts at this site. Indeed, the fine particle road dust profile reported in Table 6 indicates enrichments of S, Br and Pb above fine soil levels, with percentages closer to (though still less than) the percentages estimated for the fine soil component. These results concur with a separate meterological analysis of this data set (Thurston and Spengler, 1985) which indicated that the coarse soil component correlated with the effects of precipitation and season (as expected), but the fine fraction soil component did not. The source of the fine aerosol component initially identified as soil requires further investigation, as it neither appears or behaves exactly like a 'typical' soil aerosol.

The profile for the source tentatively identified in Table 5 as coal combustion related aerosols has a chemical composition quite different from particles in primary coal combustion emissions. This is true with regard to the crustal elements (such as Al, Si, Ca and Fe), which are indicated to be much lower percentages of the total mass than has been reported based on stack testing (e.g. Lee et al., 1975; Gladney et al., 1976), as well as volatile elements such as S, Zn and Pb, which have higher percentages than indicated by the literature. These trends were expected, however, since Natusch et al. (1974) have shown that certain volatile elements such as Se, As, Zn and Pb (gaseous in the high temperature flues of coal fired power plants) rapidly condense onto fine particles as the plume cools downwind. Moreover, since there was little coal combustion in the Boston area during the study period, it is believed that this pollution component represents particles transported into the vicinity. During transport, the character of the pollution would modify due to particle deposition and the addition of secondary aerosols generated in the atmosphere (for example, by photochemical processes) from gaseous pollutants (e.g. sulfur oxides, nitrogen oxides and hydrocarbons). As the pollution ages, these aerosols (including sulfates, nitrates and particulate carbon) would, therefore, represent an ever increasing percentage of the plume's remaining particle mass, in turn reducing the particulate mass percentage accounted for by the primary emission elements. Thus, an aged coal combustion aerosol might reasonably be expected to have characteristics qualitatively similar to those found for this coal related' component profile.

In their seminal work regarding Chemical Mass Balance (CMB) aerosol apportionment techniques,

Miller et al. (1972) recognized that source profiles would be modified between source and receptor. In their original model, they stated

$$C_i = \sum_{j=1}^{p} \alpha_{ij} f_{ij} S_j, \qquad (15)$$

where f_{ij} is the mass fraction of element i in particles from source j at the source, and α_{ij} is the coefficient of fractionation of element i between the source and the receptor. In practice, most users of (15) assume that, within an airshed, $\alpha_{ij} = 1.0$ and $f_{ij} = a_{ij}$, where a_{ij} is the mass fraction of element i in particles from source j at the receptor. However, for transported aerosols this will clearly not be the case, with some primary pollution elements being depleted by deposition while other secondary aerosol elements experience accretion in the plume.

Comparison of the 'Coal Related' profile in Table 5 with other aerosol analyses indicates support for the identification of this component as aerosols transported to Boston. In a comprehensive trajectory analysis of this data set (Thurston et al., 1982), air masses which had passed over the Ohio Valley/Midwest were found to record the highest mass fraction of Se. These Ohio Valley/Midwest air masses recorded an average Se percentage of 0.006%, approaching the 0.0089% value derived in this work for the coal related component (Table 5). Also, plots of the highest and lowest 10 impact days by this pollution component show a strong directionality, with the highest days' trajectories enveloping the industrialized Midwest, while the lowest days' trajectories passed primarily over Eastern Canada (Thurston and Spengler, 1984). Moreover, ambient aerosol measurements made between May 1980 and August 1981 at three rural sites located within the Ohio Valley Region revealed elevated Se contents in those aerosols when compared to typical Watertown contents (Shaw and Paur, 1983). The measurements, which were collected and analyzed in a manner similar to those employed in this work, indicated that the rural quarterly mean Se mass percentages ranged from 0.0082 to 0.021%. Considering our results in light of these independent data, it would appear that the coal burning Ohio Valley Region is characterized by an aerosol enriched in Se and, when air masses are transported across this region to southern New England, the Se (and its associated sulfate pollution) is also transported into southern New England.

Only limited data are available for comparison to the refuse incineration profile. Data presented in Table 6 for refuse incinerators would indicate a higher Zn concentration, lower copper percentage, and much higher chlorine percentage than reported here. The higher estimated Cu percentage may, in part, be due to errors induced by Cu emissions from the collocated high-volume samplers at this site. It is reassuring to note, however, that regressions by Kleinman et al. (1980) indicated a 1.9% Cu content for incineration aerosols, a value identical to that reported in Table 5.

The salt profile presented in Table 5 is no doubt influenced by road salt and marine aerosols impacting this coastal metropolitan location. The salt component profile agrees qualitatively with the limited available data on profiles for salt aerosols. Kowalczyk et al. (1978) present a profile derived from sea water data which indicates an aerosol dominated by Na (31%) and Cl (56%). Similarly, Taback (1979) indicate sea salt to have 31% Na and 55% Cl (see Table 6). Although Na data are not available from X-ray fluorescence, the estimate of 46 per cent Cl content for salt aerosols does appear to be in reasonable agreement with the limited 'source' composition data available.

Of special note with respect to the comparisons of the results in Table 6 with literature data are comparisons between the elemental profiles derived in this work and those statistically derived by Alpert and Hopke (1980) for particle samples collected in Boston in 1970. The Alpert and Hopke target-transformed factor analyses results were very nearly identical to the literature values input (which were themselves the same or similar to the literature values reported in Table 7). Since the Alpert and Hopke data were total suspended particle samples collected before the advent of Clean Air Act emission changes (e.g. the extensive use of low sulfur fuel oil) and before the increase in non-leaded gasoline, the noted discrepancies with our results may, in part, be due to real changes in aerosol characteristics over time. However, the nearly exact agreement between the Alpert and Hopke profiles and the target literature values input into the analysis has caused researchers to question the practical advantages of TTFA over conventional CMB approaches (Cooper and Watson, 1980). Thus, the present analysis may represent not only a technique more readily applied by available statistical packages, but also some improvement, in that it is not biased by prior assumptions about the quantitative nature of local source profiles. Recent target transformation factor analyses by Roscoe and Hopke (1981) and Alpert and Hopke (1981), however, have addressed the above noted problem by targeting the factors to unique vectors (all

elements set equal to zero except one, which is set equal to 1.0) and iteratively converging on estimates of source profiles. This approach was not applied to the Boston aerosol data, however, and it is not known how different the resultant profiles would be from the originally calculated profiles. Only by conducting proper source sampling in analyzed air sheds can such questions regarding source profiles be resolved with more certainty.

3.4. *Elemental apportionments for fine and coarse Watertown aerosols*

As discussed in Section 2, once the mean source impacts and elemental profiles are estimated, the elemental mass balances can be calculated by taking their product. Table 7 contains the fine elemental impacts distributed among the sources, as calculated from the mass contribution estimates in Table 4 and the elemental profiles in Table 5. For the fine elemental data, Cl and Cu are the elements most underpredicted, on average, by the source apportionment. This result could be anticipated, in light of the Cu impacts associated with collocated high-volume sampler operations, and the fact that the Cl component failed to achieve significance in the mass regressions (13). The correlations of the predicted and observed values (r^2) are quite high, except for Cl and Cu. Taken as a whole, these components account for 75 per cent of the elemental variance, close to the 87 per cent accounted for by the original rotated components (which included the salt component). As can be seen in Table 7, the largest contribution of sulfur mass is associated with the coal component, but nearly 35 per cent of the sulfur mass, on average, is unaccounted for by the sources identified here. As might be expected, a majority of the fine V is associated with the oil component, while most of the Br and Pb are contributed by the motor vehicle component. The soil component dominates the crustal element impacts (Si, K, Ca, Mn and Fe).

Table 8 indicates that the source components ac-

Table 7. Mean source contributions to fine elemental concentrations (ng m^{-3})

	Motor vehicle	Oil	Soil	Coal related	Sum of estimated mean contributions	Observed mean concentration	Estim.–Obs. r^2
Si	5.16	3.95	83.6		92.7 ± 5.6	104. ± 7.0	0.92
S	170.	117	74.6	844.	1206. ± 50.	1840. ± 75.	0.73
Cl	7.71	6.42	12.1		26.3 ± 1.1	86.2 ± 11.1	0.41
K	11.2	6.84	26.9	10.9	55.8 ± 2.1	76.2 ± 2.7	0.66
Ca		3.14	29.5	1.55	34.2 ± 2.0	41.9 ± 2.3	0.80
V	5.12	11.5	2.46	2.29	21.4 ± 1.0	22.1 ± 1.1	0.95
Mn	0.87	0.35	1.61	0.82	3.65 ± 0.13	3.61 ± .17	0.89
Fe	4.28	5.73	45.3	9.60	64.9 ± 3.1	74.7 ± 3.9	0.86
Ni	1.73	4.05	1.12	1.01	7.92 ± 0.39	8.57 ± 0.39	0.92
Cu	2.13		0.96	2.02	5.11 ± 0.19	16.1 ± 1.0	0.15
Zn	3.04	3.06		8.32	20.7 ± 0.7	26.5 ± 1.0	0.73
Se				0.61	0.61 ± 0.03	0.59 ± 0.05	0.79
Br	47.1	11.9	7.52	3.73	70.3 ± 3.1	85.7 ± 4.4	0.82
Pb	183.	47.7	21.6	50.2	302. ± 12.	326. ± 13	0.95

Table 8. Mean source contributions to coarse elemental concentrations (ng m⁻³)

	Motor vehicles	Oil	Soil	Sulfate	Refuse	Salt	Sum of estimated mean contributions	Observed Mean concentration	Estim.-Obs. r^2
Al	11.5	11.2	282.	23.	17.1		345. ± 23.	366. ± 26.	0.86
Si	33.0	20.8	719.	55.1	62.9		891. ± 59.	1021. ± 66.	0.89
S	24.0	6.33	19.8	134.	11.6	3.78	199. ± 9.	201. ± 9.	0.89
Cl	21.7	14.3			33.1	183.	252. ± 36.	306. ± 46.	0.93
K	5.16	2.67	80.6	10.8	5.5	1.94	107. ± 7.	112. ± 6.	0.97
Ca	7.98	4.52	147.	12.2	12.8		184. ± 12.	214. ± 15.	0.91
Ti	0.47	0.83	23.2	0.62	0.94		26.1 ± 1.9	25.9 ± 2.1	0.95
V	0.79	0.85	0.90	0.56	0.19		3.45 ± 0.31	3.34 ± 0.33	0.92
Mn	0.24	0.13	4.62	0.25	0.44		5.69 ± 0.38	5.81 ± 0.40	0.95
Fe	14.3	6.93	200.	16.8	21.1		260. ± 17.	281. ± 17.	0.94
Ni	0.42	0.26	0.53	0.37	0.41	0.06	2.06 ± 0.11	2.44 ± 0.13	0.87
Cu	0.37				9.41		9.79 ± 0.64	10.4 ± 0.7	0.93
Zn	1.16	0.41	1.65	1.10	6.10		10.4 ± 0.5	12.2 ± 0.6	0.80
Br	19.4	0.66	0.87	2.26	2.07	0.22	25.5 ± 1.6	25.7 ± 1.6	0.99
Pb	48.7	2.93	4.96	6.92	7.79	0.40	71.8 ± 4.1	75.6 ± 4.6	0.97

count for the coarse elemental concentrations better than was the case for the fine aerosol. All elements are accounted for, on average, within 20 per cent (the worst being Cl, averaging 17% underpredicted), and the predicted–observed correlations r^2 all equal or exceed 0.80. As was true for the fine aerosol, the coarse soil component dominates the crustal element impacts, a result consistent with our prior knowledge of the elemental composition of these sources. Surprisingly, however, the coarse oil component is not estimated to be the predominant V and Ni contributor, with the soil component equaling its V impact and exceeding the coarse oil component Ni impact. Both of these coarse V and Ni impacts are small, however, compared to the V and Ni impacts by the oil component in the fine mass. As expected, coarse Zn and Cu are contributed primarily by refuse incineration, while automotive emissions are estimated to be the largest contributor of ambient coarse Pb and Br.

While the major features of the results in Tables 7 and 8 are confirmed by known information regarding the possible elemental sources, a careful scrutiny of the results reveals that some of the less major sources of elements are larger than expected. One such problem was noted above for the V and Ni coarse concentrations. Also, the estimate in Table 7 that motor vehicles account for half as much fine V as oil combustion appears suspect. Although the content of V has dropped in residual oils in recent years, and oil constitutes a smaller portion of the fine mass than motor vehicles, this estimate seems large. The likely reason for this result is tied to the fact that these sources are of local origins, and their mass impacts at the Watertown site are not totally independent of one another (all local sources tending to rise during poor local dispersion). Indeed, a separate evaluation of this elemental data set with respect to meteorological interactions (Thurston and Spengler, 1984) showed that both oil and motor vehicle impacts are moderately correlated with poor local dispersion. It seems likely that this shared covariance is manifesting itself in the

elemental profile regressions by distributing the elemental masses (e.g. V and Pb) between these sources, causing some bias. This secondary statistical effect aside, however, the major features of the elemental profiles (Table 6) and elemental contributions (Tables 8 and 9) do agree with available information regarding their expected values.

4. CONCLUSIONS

This analysis has presented a particle source apportionment technique which requires a minimum of inputs regarding source characteristics, yet provides quantitative information regarding both source particle characteristics and impacts. While the theoretical aspects of the technique are somewhat complicated, it is a simple technique to execute in practice, requiring only Principal Component and regression procedures routinely available on standard statistical computer packages.

In this particular application, the absolute PCA source apportionment approach has indicated that traditional local primary emission sources such as automotive, oil and refuse incineration emission contribute only roughly one fourth of the 26 μg m⁻³ mean inhalable particle (IP) mass at the sampling site in Watertown, MA. Fugitive particle sources (soil and salt aerosols) account for nearly 30 per cent of the IP mass. Almost half the IP mass was not attributable to identified sources in the local area. Indeed, nearly 30 per cent of the IP mass was ascribed to a coal combustion related component believed to be transported into the Boston area.

The identification of a major coal related aerosol which is apparently transported to this site (primarily from the industrial Midwest) completes a chain of evidence leading to this conclusion. This component is strongly correlated with S and Se. Selenium is known to be highly concentrated in air pollution emissions from coal combustion (Natusch et al., 1974) (Klein et

al., 1975) (Wangen, 1981). Also, analyses of particle samples collected within the Ohio Valley coal burning region indicate that ambient aerosols in that region are highly enriched in selenium (Spengler and Thurston, 1983; Shaw and Paur, 1983). Statistical and wind trajectory analyses of aerosol samples collected in rural New York State have indicated that ambient selenium and sulfur concentrations increase at sites in that state during air mass transport from the coal burning Midwest to those sites (Parekh and Husain, 1981). The results contained in this and related papers (Thurston *et al.*, 1982; Thurston and Spengler, 1984) document a similar pattern for this Eastern Massachusetts site. These results indicate that, using trace elements, the particle pollution resulting from emissions in the industrial Midwest can be detected as far away as the East Coast of the U.S.

The apportionment of sulfate aerosol mass among possible sources is of special concern from both human health and environmental standpoints. Of the mean 1.8 μg m^{-3} FP sulfur mass measured, only 20 per cent was attributable to local emission sources, while 45 per cent of the sulfur mass was ascribed to the pollution component identified as coal related aerosols transported into the Boston area, on average. It seems unlikely that the 35 per cent of sulfur left unapportioned by the regressions is due to some large unconsidered source(s). Instead, it is more probable that fluctuations in sulfur dioxide to sulfate transformation rates cause day to day variations in the sulfate portion of the source elemental profiles. During the sulfur regression procedure, such variations (random errors) tend to frustrate the apportionment process, causing the model intercept to be increased disproportionately. Thus, attempts to apply source apportionment techniques to sulfate aerosols might well be improved by incorporating such variations in the sulfur portion of source aerosol elemental profiles.

We were encouraged by the consistency found between the elemental profiles derived separately for the fine and coarse components, as well as the qualitative agreement between derived profiles and literature profiles. In general, there were moderate quantitative differences between these profile values, as might be expected given the limited nature of the available literature data, the effects of aerosol fractionation on source elemental profiles, and the errors which may exist in statistical derivation techniques. However, the nature of the coal-related component's elemental profile was estimated to be appreciably different from the character of primary coal combustion emissions. Elements dominating primary emissions were found in much lower concentrations, while secondary pollution elements (e.g. S, Se which are released primarily as gases in high temperature exhausts) are found in much higher concentration. This result implies that there is a substantial influence of fractionation between source and receptor in the case of transported (aged) aerosols. If such changes in the composition of aerosols are ignored, significant errors can be expected in the source

identification and quantification of transported aerosols.

Source apportionment techniques appear, in general, to represent viable and cost-effective public policy decision making tools. The absolute principal component regression approach presented here is seen to be a useful hypothesis generating tool: one which makes few emissions or dispersion assumptions, but which deduces quantitative estimates of air pollution characteristics and culpability. Moreover, this multivariate source apportionment and characterization technique apparently represents a promising means of assessing the composition and impacts of aerosols which have experienced considerable transport and, hence, fractionation. After conducting such analysis, the decision maker would have an indication as to which sources need further elemental emission characterization, as well as an estimate of the extent to which controlling various source classes may achieve ambient particle concentration reductions. This initial application would appear to indicate that, in the case of Metropolitan Boston, MA, the control of traditional local particle sources would have less influence on fine aerosol and sulfate concentrations than would the control of emissions causing secondary aerosols transported to the Boston vicinity.

REFERENCES

Alpert D. J. and Hopke P. K. (1980) A quantitative determination of sources in the Boston urban aerosol. *Atmospheric Environment* **14**, 1137–1146.

Alpert D. J. and Hopke P. K. (1981) A determination of the sources of airborne particles collected during the Regional Air Pollution Study. *Atmospheric Environment* **15**, 675–687.

Barr A. J. *et al.* (1979) *SAS User's Guide*. SAS Institute, Raleigh, NC.

Biggins P. D. E. and Harrison R. M. (1979) Atmospheric chemistry of automotive lead. *Envir. Sci. Technol.* **13**, 558–565.

Blifford I. H., Jr. and Meeker G. O. (1967) A factor analysis model of large scale pollution. *Atmospheric Environment* **1**, 147–157.

Briggs S. L. K., Thurston G. D., Dumyahn T. S., Evans J. S., Spengler J. D. and Turner W. A. (1982) The effectiveness of dichotomous sampler quality assurance procedures. Presented at the 75th Annual Meeting of the A.P.C.A., New Orleans, LA.

Cass G. R. and McRae G. J. (1983) Source–receptor reconciliation of routine air monitoring data for trace metals: an emission inventory assisted approach. *Envir. Sci. Technol.* **17**, 129–139.

Cooper J. A. and Watson J. G. (1980) Receptor oriented methods of air particulate source apportionment. *J. Air Pollut. Control Ass.* **30**, 1116–1124.

Cox W. M. and Clark J. (1981) Ambient ozone concentration patterns among eastern U.S. urban areas using factor analysis. *J. Air Pollut. Control Ass.* **31**, 762–766.

Dattner S., Russel P. and Hasner L. (1982) Multivariate analysis techniques used during the TACB's participation in the Quail Roost II Receptor Model Validation Study. EPA Workshop on Receptor Model Validation. Rougemont, NC.

DEQE (1982a) Communication with Air Pollution Inspector Bill Vickers, Mass. Department of Air Quality Engineering.

DEQE (1982b) Data provided by Don Squires, Mass. Department of Environmental Quality Engineering.

Dzubay T. G. and Stevens R. K. (1983) Ambient air analysis with dichotomous sampler and X-ray fluorescence spectrometer. *Envir. Sci. Technol.* **9**, 663.

Dzubay T. G. and Barbour R. K. (1983) A method to improve the adhesion of particles on Teflon filters. *J. Air Pollut. Control Ass.* **33**, 692.

Eimutis E. C., Quill R. P. and Rinaldi G. M. (1978) Source assessment: noncriteria pollutant emissions. Monsanto Research Corp. for U.S. E.P.A. Industrial and Environmental Research Laboratory, EPA-600/2-78-004t. Research Triangle Park, NC.

Ferris B. G., Speizer F. E., Spengler J. D., Dockery D. W., Bishop Y. M. M., Wolfson M. and Humble C. (1979) Effects of sulfur oxides and respirable particulates on human health: methodology and demography of populations in study. *Am. Rev. Resp. Disease* **120**, 767.

Gladney E. S., Small J. A., Gordon G. E. and Zoller W. H. (1976) Composition and size distribution of in-stack particulate material at a coal-fired power plant. *Atmospheric Environment* **10**, 1071.

Greenberg R. R., Gordon G. E., Zoller W. H. and Jacko R. B., Neuendorf D. W. and Tost K. J. (1978) Composition of particles emitted from the Nicosia municipal incinerator. *Envir. Sci. Technol.* **12**, 1329.

Harris R. J. (1975) A Primer of Multivariate Statistics. Academic Press, New York.

Henry R. C. (1977) The application of factor analysis to urban aerosol source identification. *Proc. Fifth Conference on Probability and Statistics in Atmospheric Sciences*. Am. Met. Soc. Boston, MA.

Henry R. C. and Hidy G. M. (1979) Multivariate analysis of particulate sulfate and other air quality variables by principal components —I. Annual data from Los Angeles and New York. *Atmospheric Environment* **13**, 1581.

Henry R. C. and Hidy G. M. (1981) Authors reply. Multivariate analysis of particulate sulfate and other air quality variables by principal components—I. Annual data from Los Angeles and New York. *Atmospheric Environment* **15**, 425.

Henry R. C. and Hidy G. M. (1982) Multivariate analysis of particulate sulfate and other air quality variables by principal components—II. Salt Lake City, Utah and St. Louis, Missouri. *Atmospheric Environment* **16**, 929.

Hopke P. K., Gladney E. S., Gordon G. E., Zoller W. H. and Jones A. G. (1976) The use of multivariate analysis to identify sources of selected elements in the Boston urban aerosol. *Atmospheric Environment* **10**, 1015–1025.

Hopke P. K. (1982) Discussion: trace element concentrations in summer aerosols at rural sites in New York State and their possible sources, and seasonal variations in the composition of ambient sulfate—containing aerosols in the New York area. *Atmospheric Environment* **16**, 1379.

Jaklevic J. M., Gatti R. C. and Goulding F. S. (1981) A beta-gauge method applied to aerosol samples. *Envir. Sci. Technol.* **15**, 680.

Klein D. H., Andren A. W., Carter J. A., Emery J. F., Feldman C., Fulkerson W., Lyon W. S., Ogle J. C., Talmi T., Van Hook R. I. and Bolton N. (1975) Pathways of thirty-seven trace elements through coal-fired power plant. *Envir. Sci. Technol.* **9**, 973–979.

Kleinman M. T., Pasternack B. S., Eisenbud M. and Kneip T. J. (1980) Identifying and estimating the relative importance of sources of airborne particulates. *Envir. Sci. Technol.* **14**, 62–65.

Kowalczyk G. S., Choquette C. E. and Gordon G. E. (1978) Chemical element balances and identification of air pollution sources in Washington, DC. *Atmospheric Environment* **12**, 62–65.

Lee R. E., Crist H. L., Riley A. E. and MacLeod K. E. (1975) Concentration and size of trace metal emissions from a power plant, a steel plant, and a cotton gin. *Envir. Sci. Technol.* **9**, 643–647.

McFarland A. R., Ortiz C. A. and Rodes C. E. (1979) Characteristics of aerosol samplers used in ambient air monitoring. Presented at the 86th Annual Meeting of the American Institute of Chemical Engineering. Houston, TX.

Miller M. S., Friedlander S. K. and Hidy G. M. (1972) A chemical element balance for the pasadena aerosol. *J. Colloid Interface Sci.* **39**, 165–176.

Mroz E. J. (1976) The study of the elemental composition of particulate emissions from an oil-fired power plant. Ph.D. thesis, University of Maryland, College Park, MD.

NAS (1974) *Medical and Biological Effects of Environmental Pollutants—Vanadium.* Committee on Biological Effects of Atmospheric Pollutants, National Academy of Sciences, Washington, DC.

NAS (1976) *Medical and Biological Effects of Environmental Pollutants—Selenium.* Committee on Medical and Biological Effects of Environmental Pollutants, National Academy of Sciences, Washington, DC.

NAS (1980) Controlling airborne particles. Committee on Particulate Control Technology, National Academy of Sciences, Washington, D.C.

Natusch D. F. S., Wallace J. R. and Evans C. A. (1974) Toxic trace elements: preferential concentration in respirable particles. *Science, Wash.* **183**, 202–204.

Navarre J. L., Priest P. and Ronneau C. (1981) Relations between sulfur and heavy elements in rural atmosphere. *Atmospheric Environment* **15**, 221–227.

Nie N. H. (1975) Statistical Package for the Social Sciences. (2nd Edn) McGraw-Hill, New York.

Nriagu J. O. and Davidson C. I. (1980) Zinc in the atmosphere. In *Zinc in the Environment—I. Ecological Cycling* (edited by Nriagu J. O.). John Wiley, New York.

Parekh P. P. and Husain L. (1981) Trace element concentrations in summer aerosols at rural sites in New York State and their possible sources. *Atmospheric Environment* **15**, 1717–1725.

Patterson R. K. (1980) Aerosol contamination from high-volume sampler exhaust. *J. Air Pollut. Control Ass.* **30**, 169–171.

Peterson J. T. (1970) Distribution of sulfur dioxide over Metropolitan St. Louis, as described by empirical eigenvectors and its relation to meteorological parameters. *Atmospheric Environment* **4**, 501–518.

Pierson W. R. and Brachaczek W. W. (1983) Particulate matter associated with vehicles on the road—II. *Aerosol Sci. Technol.* **2**, 1–40.

Robbins J. A. and Snitz F. L. (1972) Bromine and chlorine loss from lead halide automobile exhaust particulates. *Envir. Sci. Technol.* **6**, 164–169.

Roscoe B. A. and Hopke P. K. (1981) Comparison of weighted and unweighted target transformation rotations in factor analysis. *Comput. Chem.* **5**, 1–7.

Shaw. R. W. and Paur R. J. (1983) Composition of aerosol collected at rural sites in the Ohio River Valley. *Atmospheric Environment*

Spengler J. D. and Thurston G. D. (1983) Mass and elemental composition of fine and coarse particles in six U.S. cities. *J. Air Pollut. Control Ass.* **33**, 1162.

Taback H. J., Brienza A. R., Macko J. and Brunetz N. (1979)

Fine particle emissions from stationary and miscellaneous sources in the South Coast—Air Basin. KVB 5806-783. KVB, Inc. Tustin, CA.

Thomae S. C. (1977) Size and composition of atmospheric particles in rural areas near Washington. DC. Ph.D. thesis, University of Maryland. College Park, MD.

Thurston G. D. (1981) Discussion: multivariate analysis of particulate sulfate and other air quality variables by principal components—I. Annual data from Los Angeles and New York. *Atmospheric Environment* **15**, 424.

Thurston, G. D., Spengler J. D. and Samson P. J. (1982) An assessment of the relationship between regional pollution transport and trace elements using wind trajectory analysis. *Receptor Models Applied to Contemporary Pollution Problems* (edited by Dattner S. and Hopke P.). The Air Pollution Control Association. Pittsburgh. PA.

Thurston G. D. (1983) A source apportionment of particulate air pollution in metropolitan Boston. Doctoral dissertation. Dept. of Environmental Health Sciences, School of Public Health, Harvard University. Boston, Massachusetts.

Thurston G. D. and Spengler J. D. (1985) A multivariate assessment of meteorological influences on inhalable particle source impacts. *J. Climate app. Met.* (submitted).

Walsh J. E. and Richman M. B. (1981) Seasonality in the associations between surface temperatures over the Unied States and the North Pacific Ocean. *Mon. Weath. REv.* **109**, 767.

Wangen L. E. (1981) Elemental composition of size-fractionated aerosols associated with a coal-fired power plant plume. *Envir. Sci. Technol.* **15**, 1080–1088.

Watson, J. G. (1979) Chemical element balance receptor model methodology for assessing the sources of fine and total suspended particulate matter in Portland, Oregon. Ph.D. thesis, Oregon Graduate Center, Beaverton, OR.

APPENDIX: A DEMONSTRATION OF THE ABSOLUTE PCA SCORE SOURCE IMPACT AND ELEMENTAL PROFILE ESTIMATION TECHNIQUES USING A TWO ELEMENT SYNTHETIC DATASET

In this appendix, the source apportionment techniques developed in this paper are applied to a two element concentration dataset in which the concentrations for 100 observations are generated from two predefined source impact distributions and elemental profiles. The analysis techniques are then employed to back calculate the original 'known' source impacts and source elemental profiles used to create the dataset.

The first step in the development of this example is the artificial creation of source mass and elemental impact distributions. Two normal mass distributions having the characteristics:

$$\overline{S}_1 = 4.80 \ \mu g \ m^{-3} \quad \sigma_{S_1} = 1.07 \quad m = 100 \quad \text{(A1)}$$

$$\overline{S}_2 = 11.89 \ \mu g \ m^{-3} \quad \sigma_{S_2} = 0.937 \quad m = 100 \quad \text{(A2)}$$

have been generated using the SAS statistical package (Barr *et al.*, 1979). The first distribution has been assigned to represent crustal mass impacts and defined as having a 13 per cent Al content and a 0.53 per cent V content. The second distribution is assigned as representing residual oil impacts, and defined as being 0.04 per cent Al and 3.44 per cent V. Thus, the generated 'ambient' total mass, Al and V concentration distributions ($k = 1, 2, \ldots 100$) are defined to be

$$\text{Mass}_k = S_{1k} + S_{2k} \quad \text{(A3)}$$

and

$$\text{Al}_k = 0.13S_{1k} + 0.0053S_{2k} \quad \text{(A4)}$$

and

$$V_k = 0.0004S_{1k} + 0.0344S_{2k}. \quad \text{(A5)}$$

Inserting the 100 values of S_{1k} and S_{2k} yields the distributions:

$$\overline{\text{Mass}} = 16.7 \ \mu g \ m^{-3} \quad \sigma_{\text{mass}} = 1.42 \quad m = 100 \quad \text{(A6)}$$

$$\overline{\text{Al}} = 0.687 \ \mu g \ m^{-3} \quad \sigma_{\text{Al}} = 0.139 \quad m = 100 \quad \text{(A7)}$$

$$\overline{V} = 0.411 \ \mu g \ m^{-3} \quad \sigma_V = 0.0322 \quad m = 100 \quad \text{(A8)}$$

The only information entered into the synthetic dataset source apportionment is the final total mass, Al and V concentration distributions above: data analogous to that routinely available from ambient monitoring. The total Al and V concentrations are first analyzed via varimax rotated PCA, yielding

Table A1. Rotated PCA of the synthetic elemental data set ($m = 100$)

Element	PC1	PC2
Al	0.99999	0.00493
V	0.00493	0.99999
Eigenvalue	1.0	1.0

Based upon their high correlations with Al and V, respectively, the rotated components are identified as crustal particles (PC1) and residual oil particles (PC2). The rotated component scores for these k observations ($k = 1, 2, \ldots 100$) are then calculated (as per Equation (8) of the text)

$$\text{PC1}_k^* = 1.00004 \ Z_{k_{Al}} - 0.00493 \ Z_{k_V} \quad \text{(A9)}$$

$$\text{PC2}_k^* = -.00493 \ Z_{k_{Al}} - 1.00004 \ Z_{k_V} \quad \text{(A10)}$$

where [from (1)]

$$Z_{k_{Al}} = \frac{(\text{Al}_k - \overline{\text{Al}})}{\sigma_{Al}} \quad \text{(A11)}$$

and

$$Z_{k_V} = \frac{(V_k - \overline{V})}{\sigma_V}. \quad \text{(A12)}$$

The absolute zero of each PC is estimated by inserting the absolute zero Z values [as in (10)]

$$Z_{0_{Al}} = -\frac{\overline{\text{Al}}}{\sigma_{Al}} = \frac{-0.6873}{0.1392} = -4.937 \quad \text{(A13)}$$

and

$$Z_{0_V} = -\frac{\overline{V}}{\sigma_V} = \frac{-0.4110}{0.0322} = -12.764 \quad \text{(A14)}$$

into the PC scoring equations [(A9), (A10)] above, yielding [as in (11)]

$$\text{PC1}_0^* = 1.00004(-4.937) - 0.00493(-12.764) = -4.876 \quad \text{(A15)}$$

$$\text{PC2}_0^* = -0.00493(-4.937) + 1.00004(-12.764) = -12.737. \quad \text{(A16)}$$

Using these estimates of absolute zero of the PC score distributions, the Absolute Principal Component Scores (APCS) are calculated [as in (12)]

$$\text{APCS1}_k^* = \text{PC1}_k^* - (-4.876) \quad \text{(A17)}$$

$$\text{APCS2}_k^* = \text{PC2}_k^* - (-12.737). \quad \text{(A18)}$$

Estimates of the 100 crustal and residual oil source impacts (in $\mu g \ m^{-3}$) can be derived from these APCS by regressing total mass on these APCS [as in (13)], yielding

$$\text{Mass}_k = -5 \times 10^{-5} - 1.063(\text{APCS1}_k^*) + 0.904(\text{APCS2}_k^*) \quad \text{(A19)}$$

Thus, inserting the absolute PC scores for each of the $k = 1, 2, \ldots 100$ observations into the following equations:

$$E(S_{1k}) = 1.063(APCS1_k^*) \qquad (A20)$$

$$E(S_{2k}) = 0.904(APCS2_k^*) \qquad (A21)$$

yields estimates of the source impact distributed

$$E(\overline{S}_1) = 5.18 \quad E(\sigma_{S_1}) = 1.06 \quad m = 100 \qquad (A22)$$

$$E(\overline{S}_2) = 11.51 \quad E(\sigma_{S_2}) = 0.904 \quad m = 100 \qquad (A23)$$

and

$$E(\overline{Mass}) = 16.7 \quad E(\sigma_{mass}) = 1.40 \quad m = 100 \qquad (A24)$$

which are in reasonable agreement with the original distributions presented in (A1), (A2) and (A6).

The results summarized in Equations (A22) and A23 can be employed to estimate the elemental profiles (i.e. $\%_o Al$, $\%_o V$) of the particle sources they represent. This is accomplished be regressing the 100 total Al and V concentrations [created via (A4) and (A5)] on the 100 estimates of crustal and residual oil impacts. This yields the regression equations

$$Al = 4 \times 10^{-4} + .131(E(S_{1k})) + .000759(E(S_{2k})) \qquad (A25)$$

and

$$V = 1 \times 10^{-4} + .000149(E(S_{1k})) + .0356(E(S_{2k})). \qquad (A26)$$

which compare well with the true values in (A4) and (A5) for the Al fraction of crustal courses ($E(S_1$ fraction) = 0.131 vs true value of 0.13) and the V fraction for residual oil ($E(S_2$ fraction) = 0.0356 vs true value of 0.0344). The estimates for elements not primarily concentrated in these sources (V fraction for crustal sources and Al fraction for residual oil), however, are seen to be less accurately predicted, in both cases being underestimates of the true elemental fractions presented in (A4) and (A5).

This synthetic dataset analysis has provided an example of the steps involved in the use of absolute PCA score analyses to derive source impacts and elemental profiles. Further, this example indicated that, in this simple case, the estimated source impacts and source profiles approximated the major features of the true distributions used to generate the dataset.

REFERENCE

Barr A. J., Goodnight J. H., Sall J. P., Blair W. H. and Chilko D. M. (1979) *SAS User's Guide*, 1979 *Edition*. SAS Institute, ISBN 0-017382-06-3, Raleigh, N.C.