MANE-VU
Mid-Atlantic/Northeast Visibility Union

Source Apportionment Analysis of Air Quality Monitoring Data: Phase II

Final Report

APPENDIX 6.1

Principal Components Analysis

Prepared by
Desert Research Institute

March 2005

for the
Mid-Atlantic/Northeast Visibility Union
and
Midwest Regional Planning Organization

Submitted to the United States Environmental Protection Agency, Region III under EPA Agreement X-983384-01-0 and to the United States Environmental Protection Agency, Region V under EPA Agreement A-97522601-1.

Desert Research Institute

For copies of this report contact:

MARAMA
Mid-Atlantic Regional Air Management Association
711 West 40th Street
Suite 318
Baltimore, MD 21211

TEL: 410-467-0170
FAX: 410-467-1737
http://www.marama.org/

SOURCE APPORTIONMENT ANALYSIS OF AIR QUALITY MONITORING DATA: PHASE II

This report was prepared by the Desert Research Institute, under contract to MARAMA. Funds for preparation thereof were provided to MARAMA by both MANE-VU and the Midwest RPO.

The primary authors of the report were:

Desert Research Institute

Dr. Johann Engelbrecht (PI)
Dr. Mark Green
Dr. Hampden Kuhns
Dr. Richard Tropp

2215 Raggio Parkway
Reno, NV 89512-1095
Tel. 775 674 7027
Fax. 775 674 7009
e-mail johann@dri.edu

TABLE OF CONTENTS

Page

TABLE OF FIGURES

LIST OF TABLES

6. RECEPTOR MODELING

Various receptor were explored, and this report summarizes the results from the Principal Components Analysis (PCA), Positive Matrix Factorization (PMF), as well as the UNMIX modeling. The sampling periods and numbers of samples (observations) for each of the four sights are given in Table 6.1-1.

Table 6.1-1. Sampling periods and total number of samples (observations) for the four investigated sites.

	Sampling Period		Total Number of Samples
	From	To	
Boundary Waters Canoe Area, MN	08/14/1991	02/25/2002	953
Lye Brook Wilderness Area, VT	09/21/1991	02/25/2002	1001
Shenandoah National Park, VA	03/02/1988	02/25/2002	1399
Washington DC	08/31/1988	02/25/2002	1344

The species concentrations from the IMPROVE data sets were first analyzed by PCA to assess which variables may be of importance for PMF and UNMIX modeling. For the PCA no outliers were distinguished or eliminated, but data sets were adjusted by replacing all non-positive (i.e. zero and negative) concentrations by missing values.

By analyzing data sub-sets, and later applying wind direction and back trajectory analysis, it was possible to determine *de facto* profiles attributable to specific sources (e.g., metallurgical processing) or a generalized source category (e.g., "urbanized area downwind influence").

6.1 Principal Components Analysis (PCA)

Most standard statistical techniques such as linear regression analysis and analysis of variance are concerned with mean values of data sets (confidence intervals about the mean, or if several means are equal, etc.). PCA does not fall in this category. It is a multivariate statistical technique, which considers the structure of the variability and interdependence amongst random variables in a multidimensional data set (Le Maitre, 1982). The PCA procedure calculates a new set of independent variables (principal components, eigenvectors) from the original data set of independent variables (chemical species), usually from the covariance matrix (as in this case). A geometric interpretation of PCA is that it recalculates the original variable concentrations, into principal components, in decreasing order of variance (eigenvalues) so that most of the variance is described by the first few (three or four) principal components, so also reducing the dimensionality of the data set. In order to minimize the dimensions no Varimax rotations were performed on the data sets. The first eigenvector therefore represents the assemblage of chemical species, which accounts for the largest proportion of the variance in the data set. The sample scores can furthermore be graphically represented by two- or three-dimensional orthogonal plots, without sacrificing much information from the higher dimension original data set. It should be borne in mind that PCA remains a qualitative exploratory tool, being a descriptive technique, with no statistical model being assumed.

In this project, PCA was applied individually to each of the four multivariate data sets (Boundary Waters Canoe Area, Lye Brook Wilderness Area, Shenandoah National Park and Washington DC). It was, furthermore, applied to various data subsets from each of the four sites, including monthly subsets. In order to better resolve point and local sources, other data subsets, representative of different meteorological conditions, or identified by pronounced variance in minor and trace element compositions, are explored.

The SAS® System Version 8.2 software package for Personal Computers was used for general data reduction and PCA.

Histogram plots and other exploratory statistics were generated for each of the variables from each of the four data sets. This provided information on the general structure of the data, and established which parameters would be appropriate for further data analysis. It was found that nearly all the variables showed Gaussian to lognormal data distributions, with concentration ranges often spanning up to four orders of magnitude. The logarithmic transformation of the data sets was considered, but since PCA does not require this, it was not done. However, because of the wide concentration ranges amongst the major and trace species, and in order to weigh the variables equally in the PCA, each variable was standardized to its site mean, *i.e.* each new variable had a mean value of unity (1). This is of particular importance when calculating the principal components from the covariance matrix. PCA procedures eliminate samples (observations) with missing values from the calculation.

PCA was subsequently performed on various combinations of variables, for the complete sets as well as on subsets of monthly samples (combination of each month's data for all years).

The Figures 6.1-1, 6.1-3, 6.1-5 and 6.1-7 graphically portray re-calculated values of the first eight PCA eigenvectors (principal components), as well as the sum of the eight eigenvectors. The species variances shown in the diagrams were calculated from the squared values of the coefficients of the eigenvectors (which sum to unity), in this fashion also converting the negative eigenvector coefficients to positive values. The SUM column is sum of the re-calculated eight principal components, weighted by the eigenvalue. This column can be interpreted to represent the variance of that species in the average ambient sample for that site.

From the PCA, AmmNitr can be seen to be the single most important species at all four sites, being prominent in Winter and Fall but also in Spring. In June 1996 glycerin was added to the denuders for the Module B samplers, resulting in a drop in the mean (1 $\mu g/m^3$ to 0.5 $\mu g/m^3$ for Shenandoah) nitrate concentration levels as well as the variance, more pronounced in the Eastern USA. PCA of the two subsets, prior and after June 1996 do not reflect these statistics, and does not disqualify nitrate as the most important species, showing the largest amount of variance. Al is the second most important species, especially prominent during the Spring and Summer Seasons. Other prominent species during the Winter and Fall Seasons are EC3 and OC1. Washington DC differs from the other sites, in that OPT dominates during the Winter and Fall seasons.

Species identified by PCA were included in the subsequent PMF and UNMIX receptor modeling procedures.

Summary tables (Tables 6.1-2, 6.1-7, 6.1-12 and 6.1-17) list and figures (Figure 6.1-1, 6.1-3, 6.1-5 and 6.1-7) portray the results of the PCA for the total period.

The monthly calculated PCA results are portrayed in Figures 6.1-2, 6.1-4, 6.1-6 and 6.1-8. Tables 6.1-4, 6.1-9, 6.1-14, 6.1-19 summarize the monthly results, grouped by season, of the species in the first three principal components, which contribute most to the variance. The species with principal component (eigenvector) coefficients greater than 0.3 are listed, and of these, those with coefficients greater than 0.7 are highlighted. These coefficient values were arbitrarily selected and the value of greater than 0.3 represents those species, which contribute more than approximately 10% (0.32%) to that principal component, while 0.7, represents that species which contribute more than approximately 50% (0.7^2%) to that principal component. For each sample set the first principal component (eigenvector) accounts for most of the variance (35–85%) followed by the second (10–30%) and the third (5-20%).

6.1.1 Boundary Waters Canoe Area

The Principal components (eigenvectors) for the Boundary Waters Canoe Area are given in Table 6.1-2 and re-calculated values, as elaborated upon in the previous paragraph, are shown in Figure 6.1-1. Species of importance in the first few principal components include OC1 (biogenic), Si and Al (crustal), as well as ammonium nitrate. Other species of note include EC3 in Prin5 and Fe in Prin7.

Table 6.1-2. Eigenvectors and eigenvalues (as proportions of total variance) for Boundary Waters Canoe Area. The coefficients greater than 0.7 are highlighted in yellow and those greater than 0.3 and less than 0.7 are highlighted in turquoise.

Eigenvectors	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
Eigenvalue proportion	0.3671	0.2201	0.1791	0.0612	0.0556	0.0360	0.0216	0.0130
AmmNitr	0.1999	-0.0787	0.8839	-0.2510	-0.0125	-0.2995	-0.0750	-0.0899
AmmSulf	0.1109	0.0004	0.1705	0.1345	-0.0423	0.5994	-0.1176	-0.0304
OC1	0.3099	0.5847	-0.1137	-0.5360	-0.3720	0.0855	0.1177	0.2010
OC2	0.2263	0.3291	-0.0435	0.1121	-0.0814	-0.0301	-0.0552	-0.2364
OC3	0.1983	0.2325	-0.0425	0.4550	-0.0584	-0.4212	-0.0093	-0.0090
OC4	0.1816	0.0980	0.0503	0.4084	0.0283	-0.1335	0.0023	0.2236
OPT	0.1654	0.2473	-0.0006	0.2003	0.0649	0.0200	-0.1274	-0.3525
EC1	0.1960	0.2116	0.0612	0.2034	0.0398	0.0457	-0.0808	-0.0884
EC2	0.0965	0.1059	-0.0119	-0.0295	0.1733	0.1017	0.0749	-0.3042
EC3	0.0964	0.2442	-0.0548	-0.2345	0.8920	-0.0086	-0.0974	0.1299
Al	0.5583	-0.4345	-0.2531	-0.1726	-0.0432	-0.0280	-0.4011	-0.2322
Si	0.3289	-0.2354	-0.1587	-0.0511	0.0180	-0.0157	0.0277	0.0170
Ca	0.2897	-0.1799	-0.0612	0.0042	0.0100	-0.0326	0.1114	0.4208
K	0.1835	-0.0406	-0.0098	0.1156	0.0312	-0.1414	0.0453	0.3885
Fe	0.2516	-0.1582	0.0051	0.0145	0.1130	0.0154	0.8617	-0.2308
Se	0.1440	0.0007	0.1878	0.1931	0.0250	0.4733	0.0666	-0.0400
Zn	0.1494	-0.0051	0.1875	0.1566	0.0243	0.2972	-0.0509	0.4075
Total	4.0537	1.1421	0.9757	0.7841	0.8656	0.5636	0.3146	0.1876

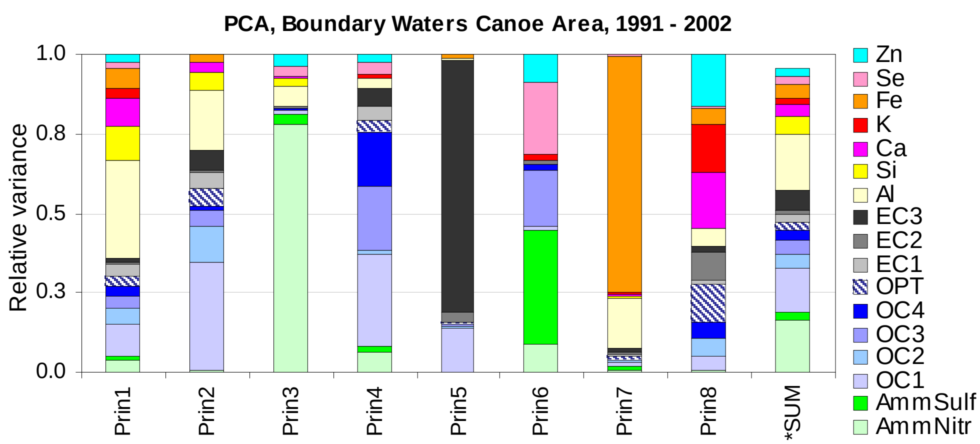

Figure 6.1-1. Principal component plot for Boundary Waters Canoe Area, showing the relative variance for each of the species, for the first eight Principal components. Also shown as SUM is the total weighted variance for each species for the sampling period.

Figure 6.1-2 below portrays the monthly PCA for the Boundary Waters Canoe Area. The low $PM_{2.5}$ loadings (4627 – 7217 ng/m^3) and the small variability is seen to be characteristic of this remote sampling site. The fractions of species do not vary much by season, with AmmNitr being abundant during the winter months, and the crustal material (Al, Si) being more abundant during April, May and September. A high coefficient for OC1 was analyzed for August.

Table 6.1-3. Sample subset applied in monthly PCA, for Boundary Waters Canoe Area.

	Sampling Period From	Sampling Period To	Total Number of Samples
Boundary Waters Canoe Area, MN	08/14/1991	02/25/2002	296 (953)

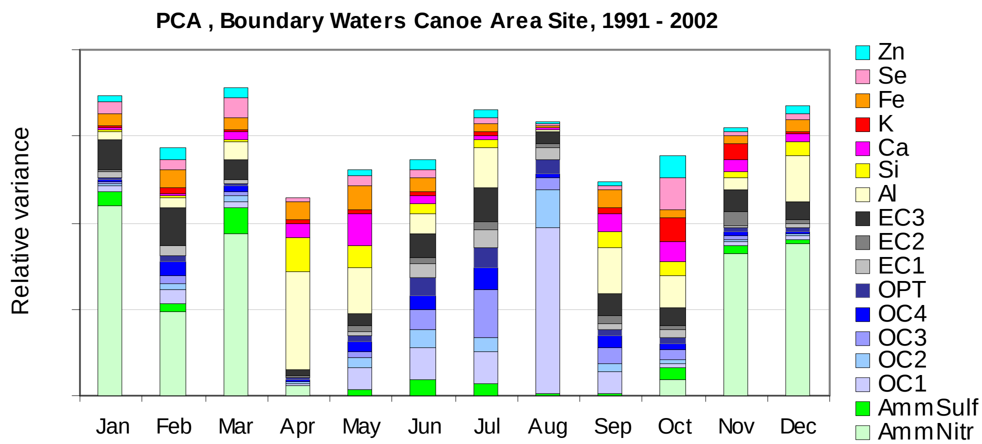

Figure 6.1-2. Principal component plot for Boundary Waters Canoe Area, showing the monthly fluctuation in relative variance for each of the species. The bar heights represent the monthly mean $PM_{2.5}$ masses.

Table 6.1-4. Summary of species of greatest significance in monthly PCA, grouped by season, for Boundary Waters Canoe Area. Only those species with eigenvector coefficients greater than 0.3 are listed and those with coefficients greater than 0.7 are highlighted in yellow.

Boundary Waters Canoe Area - Species Grouped by Season				
Season	Winter	Spring	Summer	Fall
Months	12, 1, 2	3, 4, 5	6, 7, 8	9, 10, 11
Principle Component				
Prin1	AmmNitr,	AmmNitr, Si, Al, Ca	EC1, OC1, OC2, OC3, OC4, OPT	AmmNitr, OC1, Si, Al, Ca, K
Prin2	EC3, Si, Al, Ca	AmmSulf, AmmNitr,, EC3, OC1, OC2, Al, Fe	EC3, OC1, OC3, OC4, OPT, Si, Al, Fe	AmmSulf, , EC1, EC2, EC3, OPT, Al, Ca, K
Prin3	Al, Ca, Fe, EC3, OC1, OC4, Fe, AmmNitr	AmmSulf, Al, Ca, EC3, OC1, Fe	EC3, OC1, OC3, OC4, Al	AmmNitr, K, EC3, Fe

A minor sample subset (Tables 6.1-5, 6.1-6) was run with a greater number of species. However, this resulted in many more observations (samples) being eliminated from the PCA, as a result of missing values.

Table 6.1-5. Minor sample subset with trace elements included in PCA, for Boundary Waters Canoe Area.

	Sampling Period From	Sampling Period To	Total Number of Samples
Boundary Waters Canoe Area, MN	08/14/1991	02/25/2002	31 (953)

Table 6.1-6. Eigenvectors ("source profiles") and eigenvalues ("variances") for a minor sample subset from Boundary Waters Canoe Area.

				Boundary Waters Canoe Area				
Eigenvectors								
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
EC	0.1862	0.0391	0.0188	0.1348	0.0965	0.0551	-0.0528	-0.0408
EC1	0.2344	-0.1025	0.0750	0.1616	-0.0223	0.1208	0.0598	0.0706
EC2	0.0636	0.0303	0.1540	-0.0076	0.2862	0.0828	0.1284	-0.0397
EC3	0.0403	-0.1080	0.6221	0.1396	0.3141	0.1054	-0.0257	-0.5370
OC	0.2937	-0.2305	-0.0486	0.0922	0.0172	-0.0675	0.0191	0.0315
OC1	0.0107	-0.1611	0.1700	0.1652	0.2564	-0.0326	0.4328	0.3570
OC2	0.2064	-0.1824	0.0637	0.1184	0.1346	0.0178	0.0651	0.1398
OC3	0.4424	-0.3285	-0.2351	0.0367	-0.0703	-0.2308	-0.1127	-0.0921
OC4	0.3480	-0.2163	-0.1873	0.0812	-0.0260	-0.1604	-0.1375	-0.0818
OPT	0.2126	-0.1748	0.1663	0.1309	-0.0096	0.1605	0.1658	0.0997
AmmSulf	0.0693	0.0867	-0.0281	0.1313	-0.1037	0.2408	0.1365	0.2209
AmmNitr	0.1560	0.5831	0.0638	0.5999	0.0761	-0.2944	-0.3201	0.0897
Si	0.1960	0.1478	-0.0692	-0.2244	0.1613	0.1120	0.0421	0.0665
Al	0.2905	0.2031	-0.0979	-0.3871	0.3296	0.1478	-0.1161	-0.0167
Ca	0.2548	0.2183	-0.1291	-0.2003	0.1469	0.2648	-0.0232	0.0068
K	0.1619	0.0446	-0.0849	-0.0154	0.1588	0.1028	-0.0903	-0.0936
Fe	0.2081	0.3835	0.0692	-0.2512	-0.0560	-0.4588	0.5443	-0.2079
Cr	0.2675	-0.0080	0.4252	-0.1944	-0.6020	-0.1020	-0.0553	-0.0274
Pb	0.0409	0.1285	-0.1124	0.1506	-0.3377	0.4852	0.1172	-0.3632
S	0.0697	0.0810	-0.0119	0.1345	-0.1096	0.2293	0.1232	0.2031
Se	0.1487	0.1679	0.0241	0.1034	-0.1350	0.1127	0.3076	0.1914
V	0.1044	0.0830	0.4221	-0.2756	-0.0651	0.1094	-0.3967	0.4154
Zn	0.1272	0.1105	-0.0986	0.1235	-0.0878	0.2458	-0.0174	-0.1831
Eigenvalue	16.37	6.91	5.04	3.88	1.88	1.31	1.15	1.03

Table 6.1-6 above shows the PCA results from Boundary Waters Canoe Area, for the entire sampling period. In this instance minor and trace elements were included in the analysis, and the eigenvectors for principal component 1 to 8 (Prin1-8) were calculated. The eigenvector coefficients greater than 0.3 are highlighted, and represent the species of greatest importance (more than approximately 10% contribution to that principal component). As before, EC3 and NO3- are dominant, together with OC3 and OC4. Principal component 3 as well as some of the minor principal components do have large coefficients for the trace species (Cr in Prin3 and 5; Pb in Prin5 and 6; Se in Prin7; V in Prin3, 7 and 8) Although this PCA may not be representative of the complete data set, it does serve to resolve a minor sample subset, indicative of events or meteorological conditions when variable high concentrations of the mentioned trace species were measured at this site. This information will later be used for the back trajectory analysis.

6.1.2 Lye Brook Wilderness Area

The principal components (eigenvectors) for the Lye Brook Wilderness Area are given in Table 6-7 and re-calculated values, as elaborated upon previously, are shown in Figure 6.1-3. Species of importance in the first few principal components include Al, Si and Fe (crustal), OC1 (biogenic), and EC3. Se, in power plant fly-ash, less obvious at Boundary Waters Canoe Area, is prominent at Lye Brook Wilderness Area.

Table 6.1-7. Eigenvectors and eigenvalues (as proportions of total variance) for Lye Brook Wilderness Area. The coefficients greater than 0.7 are highlighted in yellow and those greater than 0.3 and less than 0.7 are highlighted in turquoise.

Eigenvectors	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
Eigenvalue proportion	0.5285	0.1497	0.0894	0.0750	0.0512	0.0274	0.0211	0.0164
AmmNitr	0.0829	0.1973	-0.4490	0.0864	0.6307	0.0649	0.4453	-0.1754
AmmSulf	0.1891	0.2599	-0.0494	0.0608	-0.3678	-0.5255	0.1814	0.2704
OC1	0.1139	0.3919	0.6361	-0.4650	0.4318	-0.1033	-0.0418	0.0610
OC2	0.1518	0.3158	0.0984	-0.0181	-0.2875	0.0624	0.1719	-0.0523
OC3	0.1398	0.2066	-0.0329	-0.0911	-0.1982	0.6026	-0.2736	-0.2563
OC4	0.1572	0.2944	-0.0844	0.0336	-0.1526	0.1947	-0.1727	0.0364
OPT	0.0952	0.2309	0.0889	0.0959	-0.1933	0.0198	0.3907	-0.3043
EC1	0.1260	0.2986	-0.0413	0.0681	-0.1282	0.0483	0.1492	-0.1739
EC2	0.0378	0.0470	0.1208	0.1898	-0.0339	-0.0600	0.1974	-0.0683
EC3	0.0384	0.0166	0.4602	0.8211	0.1830	0.1067	-0.0705	0.0479
Al	0.7303	-0.4319	0.0821	-0.0586	0.0424	-0.1265	-0.0658	-0.3824
Si	0.3464	-0.1586	0.0040	-0.0179	-0.0143	0.0112	0.1116	0.2821
Ca	0.1784	-0.0122	-0.0464	0.0026	0.0571	0.2984	0.0471	0.5657
K	0.1501	0.0841	-0.0499	0.0147	0.0724	0.1827	-0.0233	0.1714
Fe	0.3265	-0.0213	-0.0916	0.0018	0.0318	0.1104	0.0634	0.3041
Se	0.1483	0.3082	-0.2634	0.1719	0.1651	-0.3678	-0.6144	-0.0891
Zn	0.1140	0.2272	-0.2197	0.0733	0.0901	-0.0178	-0.1137	0.1313
Total	3.6545	2.4041	0.2519	1.0442	0.3799	0.5285	0.4033	0.3846

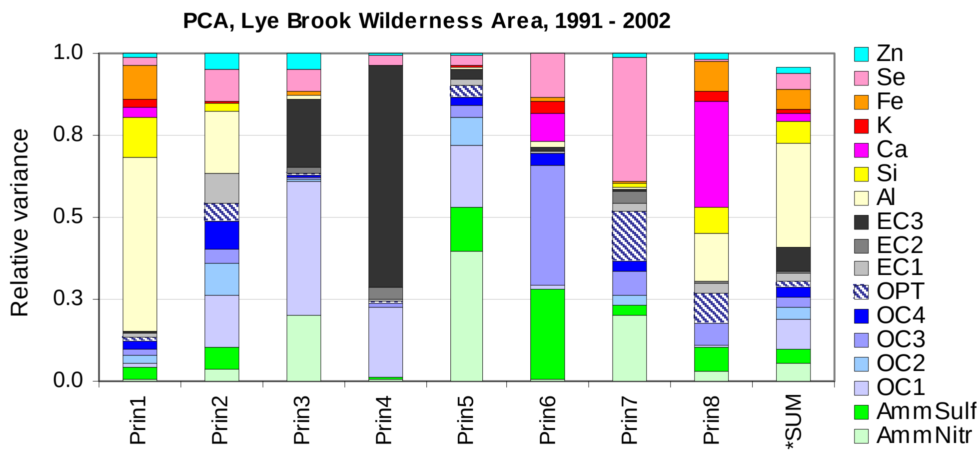

Figure 6.1-3. Principal component plot for Lye Brook Wilderness Area, showing the relative variance for each of the species, for the first eight Principal components. Also shown as SUM is the total weighted variance for each species for the sampling period.

Figure 6.1-4 portrays the monthly PCA for the Lye Brook Wilderness Area. The low $PM_{2.5}$ loadings during the winter months (5164 ng/m^3) are similar to those of the Boundary Waters Canoe Area. However, ascending particulate levels (as high as 11386 ng/m^3 in August) are experienced during the summer months. The crustal species such as Al are prominent during the summer months, more so in June. High values for EC3 are found in July, August and November, and may have been imported from urban areas. Se is prominent during the winter months of November, December, January and February. As with Boundary Waters Canoe Area, AmmNitr was found to be prominent during the winter months of January, March and December while OC1 stands out in February and December.

Table 6.1-8. Sample subset applied in monthly PCA, for Lye Brook Wilderness Area.

	Sampling Period		Total Number of Samples
	From	To	
Lye Brook Wilderness Area, VT	09/21/1991	02/25/2002	302 (1001)

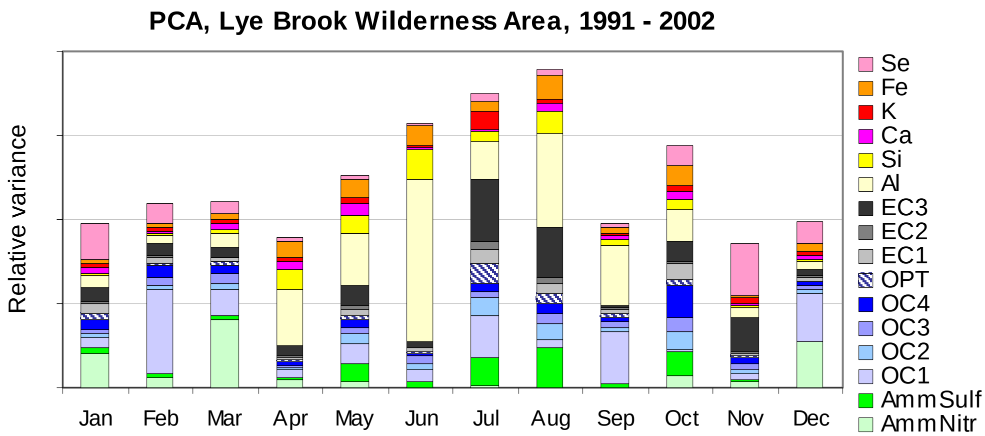

Figure 6.1-4. Principal component plot for Lye Brook Wilderness Area, showing the monthly fluctuation in relative variance for each of the species. The bar heights represent the monthly mean $PM_{2.5}$ masses.

Table 6.1-9. Summary of species of greatest significance in monthly PCA, grouped by season, for Lye Brook Wilderness Area. Only those species with eigenvector coefficients greater than 0.3 are listed and those with coefficients greater than 0.7 are highlighted in yellow.

Lye Brook Wilderness Area - Species Grouped by Season

Season	Winter	Spring	Summer	Fall
Months	12, 1, 2	3, 4, 5	6, 7, 8	9, 10, 11
Principle Component				
Prin1	AmmNitr, EC1, OC1, OC4	AmmNitr, Si, Al, Ca, Fe	AmmSulf, EC3, OC1, OC2, OPT, Si, Al, Fe,	AmmSulf, OC1, OC2, OC4, EC1, EC3, Al, Fe
Prin2	AmmNitr, OC1, OC3, OC4, Al, Ca	AmmSulf, AmmNitr , EC3, OC1, OC2, OC3,, Al	AmmSulf, EC1, EC3, OC1, OC2, OC3	OC1, OC3, OC4, EC3, Al, K, Fe
Prin3	AmmNitr, EC3, Al, Ca	EC3, OC1	EC1, EC3, OC2, OPT, OC1, OC3, Al, Fe	AmmSulf, AmmNitr, EC1, EC3, OC1, OC2, OC3, OPT, Al, Fe

A summary of PCA results from a minor subset of samples run with a greater number of trace species, is given in Tables 6.1-10 and -11.

Table 6.1-10. Minor sample subset with trace elements included in PCA, for Lye Brook Wilderness Area.

	Sampling Period		Total Number of Samples
	From	To	
Lye Brook Wilderness Area, VT	09/21/1991	02/25/2002	49 (1001)

Table 6.1-11. Eigenvectors ("source profiles") and Eigenvalues ("variances") for a minor sample subset from Lye Brook Wilderness Area

				Lye Brook				
Eigenvectors								
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
EC	0.1339	0.1279	-0.0806	0.0586	0.0157	0.1918	0.0997	-0.0258
EC1	0.1311	0.0869	0.0089	0.2002	0.0269	0.2125	0.0389	-0.0204
EC2	0.0392	0.0639	0.0642	-0.0845	0.0302	0.2666	0.0533	0.0316
EC3	0.0219	0.2158	0.1227	-0.6888	0.0717	0.3534	0.4154	-0.2128
OC	0.1321	0.0378	0.0675	0.1938	0.1298	0.1883	0.0791	0.1175
OC1	0.0728	0.1314	0.2396	0.2213	-0.0062	-0.4372	0.7425	0.2998
OC2	0.1480	0.0492	0.1130	0.1982	0.1594	0.2168	0.0951	0.0114
OC3	0.1655	-0.0555	-0.0034	0.1832	0.2327	0.3159	-0.0826	0.3510
OC4	0.1461	0.0881	-0.0215	0.1755	0.1309	0.1879	0.0740	0.0194
OPT	0.0898	0.0398	0.1340	0.2106	0.0430	0.2612	-0.0043	-0.0059
AmmSulf	0.1908	-0.0145	0.0080	0.1784	0.1622	-0.0157	-0.0404	-0.3703
AmmNitr	0.1248	0.2755	-0.5115	0.0612	-0.6068	0.0830	0.0320	0.1813
Si	0.3264	-0.1749	0.0152	-0.1298	-0.0776	-0.0111	-0.0395	-0.0395
Al	0.6585	-0.4066	0.0234	-0.1972	-0.0605	-0.2412	-0.0501	0.0150
Ca	0.1683	0.0076	0.0367	-0.0996	-0.1691	0.1348	0.0953	0.0624
K	0.1164	0.0607	-0.0423	0.0113	-0.0981	0.0793	0.0347	0.0509
Fe	0.3274	-0.0960	-0.0507	-0.0691	-0.0514	0.0433	0.0504	0.0283
Cr	0.2216	0.5801	0.5395	0.0319	-0.2420	-0.1970	-0.3297	-0.1807
Pb	0.0949	0.2092	-0.1344	0.0674	-0.0572	0.0646	0.0013	-0.0708
S	0.1323	0.0421	0.0168	0.1757	0.1337	0.0131	-0.0063	-0.2918
Se	0.1727	0.4017	-0.4804	-0.1361	0.5628	-0.3315	-0.0404	-0.0427
V	0.0533	0.1917	0.1555	-0.2714	0.1729	-0.0031	-0.3187	0.6467
Zn	0.1093	0.1635	-0.2202	0.0754	-0.0785	0.0194	0.0523	-0.0570
Eigenvalue	16.57	5.12	2.61	1.75	1.02	0.84	0.70	0.62

In the case of Lye Brook Wilderness Area, only a small sample subset (49 out of a total of 1001 observations) is resolved. As shown in Table 6.1-11 above, Al, Si and Fe are composed in Prin1, while the trace elements Cr occur in Prin2, 3 and 7; Se in Prin2, 3, 5 and 6; V in Prin7 and 8.

6.1.3 Shenandoah National Park

The principal components (eigenvectors) for Shenandoah National Park are given in Table 6.1-12 and re-calculated values, as elaborated upon previously, are shown in Figure 6.1-5 below. Species of importance in the first few Principal components include Al, Si and Fe (crustal), OC1(biogenic), and EC3 (transported motor vehicle emissions). OPT appears as an important component, as well as AmmNitr, and Zn (smelter). Se (power plant fly-ash) does show up in the lower principal components (Prin7, 8).

Table 6.1-12. Eigenvectors and eigenvalues (as proportions of total variance) for Shenandoah National Park. The coefficients greater than 0.7 are highlighted in yellow and those greater than 0.3 and less than 0.7 are highlighted in turquoise.

Eigenvectors	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
Eigenvector proportion	0.5556	0.1528	0.0992	0.0487	0.0376	0.0250	0.0196	0.0164
AmmNitr	-0.0138	-0.0693	0.0383	0.7350	0.4541	0.2017	-0.2375	-0.3133
AmmSulf	0.0659	0.1982	0.1754	-0.1066	-0.2503	0.5450	0.2563	-0.2844
OC1	0.0913	0.4468	0.3368	-0.3282	0.6545	-0.2419	0.1977	-0.1839
OC2	0.0609	0.2733	0.1562	-0.0140	-0.2433	0.0731	-0.0401	-0.1545
OC3	0.0421	0.2703	0.1318	0.2743	-0.3375	-0.4900	-0.0491	-0.1006
OC4	0.0554	0.2712	0.1766	0.2131	-0.2672	-0.1769	0.0954	-0.1343
OPT	0.0459	0.3455	0.1076	-0.1023	-0.0242	0.2847	-0.6014	0.2227
EC1	0.0368	0.2940	0.1182	0.0615	-0.0492	0.1165	-0.3605	0.1863
EC2	0.0122	0.1295	-0.1183	-0.1117	0.1389	0.1932	-0.0096	0.2506
EC3	0.0161	0.4945	-0.8378	0.0689	0.0504	-0.0055	0.1170	-0.0969
Al	0.7897	-0.1711	-0.0983	-0.0668	-0.0237	-0.0135	-0.0949	-0.2038
Si	0.4284	-0.0437	-0.0179	-0.0359	0.0249	0.0401	-0.0431	0.0380
Ca	0.1790	0.0421	0.0357	0.1605	0.0362	-0.0322	0.2220	0.6172
K	0.1311	0.1091	0.0571	0.1897	-0.0591	-0.2455	0.0077	0.1561
Fe	0.3422	-0.0207	0.0153	0.0914	0.0503	0.0437	0.1047	0.1228
Se	0.0258	0.0925	0.0958	0.1566	-0.1170	0.3281	0.3757	-0.0707
Zn	0.0220	0.1185	0.1266	0.2967	0.1063	0.1495	0.3286	0.3321
Total	2.3311	2.7808	0.4992	1.4822	0.1442	0.7701	0.2691	0.3832

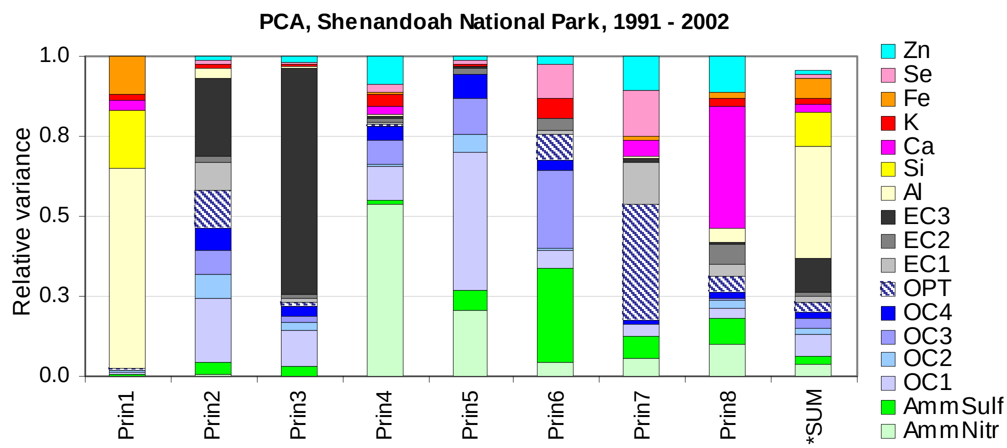

Figure 6.1-5. Principal component plot for Shenandoah National Park, showing the relative variance for each of the species, for the first eight Principal components. Also shown as SUM is the total weighted variance for each species for the sampling period.

Figure 6.1-6 below portrays the monthly PCA for Shenandoah National Park. Shenandoah shows the greatest fluctuation of $PM_{2.5}$ loadings of the sights investigated with total loadings increasing sharply from 5337 ng/m^3 in January, rising sharply to 21833 ng/m^3 in August, and dropping back to 5515 ng/m^3 in December. Species, which show the greatest increase in values, are Al and Si, especially high in July. AmmNitr and EC3 show up in the winter months of February, May, June and September.

Table 6.1-13. Sample subset applied in monthly PCA, for Shenandoah National Park.

	Sampling Period		Total Number of Samples
	From	To	
Shenandoah National Park, VA	03/02/1988	02/25/2002	615 (1399)

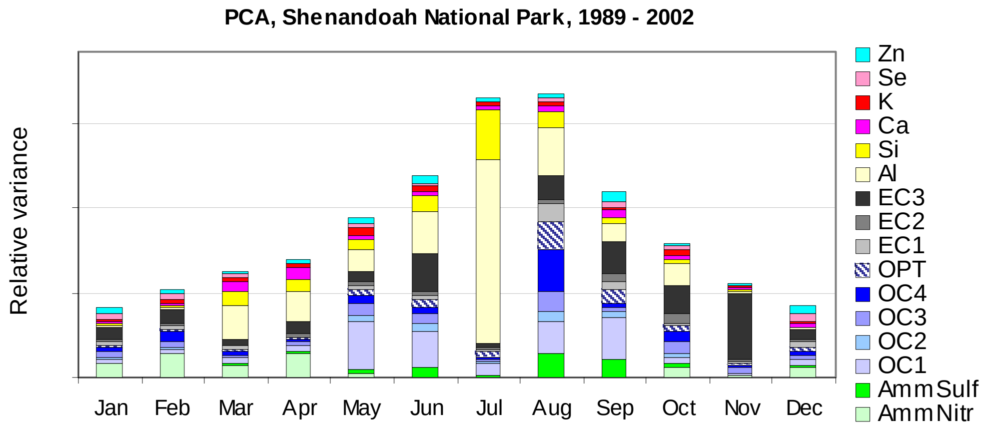

Figure 6.1-6. Principal component plot for Shenandoah National Park, showing the monthly fluctuation in relative variance for each of the species. The bar heights represent the monthly mean $PM_{2.5}$ masses.

Table 6.1-14. Summary of species of greatest significance in monthly PCA, grouped by season, for Shenandoah National Park. Only those species with eigenvector coefficients greater than 0.3 are listed and those with coefficients greater than 0.7 are highlighted in yellow.

Shenandoah National Park - Species Grouped by Season

Season	Winter	Spring	Summer	Fall
Months	12, 1, 2	3, 4, 5	6, 7, 8	9, 10, 11
Principle Component				
Prin1	AmmNitr, EC1, OC3, OC4, K	OC1, Si, Al, Ca, Fe	OC1, OC4, Si, Al, Fe	AmmSulf, EC3, OC1, OC3, OC4, OPT, Al, Ca, K
Prin2	AmmNitr, EC2, EC3, OC3, OC4	AmmNitr, OC1, Si, Al, Fe	EC1, EC3, OC4, OPT, OC1, OPT, Si, Al	EC2,EC3, OC1, OC3 Al, Ca
Prin3	AmmNitr, EC2, EC3, OC1	AmmNitr, EC2, EC3, OC1, OC3, OC4, Al	AmmSulf, EC3, OC1, OC2, OC3, OC4	AmmNitr, EC3, OC1, OC3, OC4, Al, Ca

A minor sample subset (Tables 6.1-15, 6.1-16) of samples was run with a greater number of species.

Table 6.1-15. Minor sample subset with trace elements included in PCA, for Shenandoah National Park.

	Sampling Period		Total Number of Samples
	From	To	
Shenandoah National Park, VA	03/02/1988	02/25/2002	48 (1399)

Table 6.1-16. Eigenvectors ("source profiles") and Eigenvalues ("variances") for a minor sample subset from Shenandoah National Park.

				Shenandoah				
Eigenvectors								
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
EC	0.0262	0.1016	0.2593	-0.0591	0.0156	0.1222	0.1402	0.1634
EC1	0.0171	0.1407	0.1683	-0.0882	-0.0596	0.2142	0.2140	-0.0076
EC2	0.0690	0.0658	-0.1501	-0.0102	0.1102	0.0942	0.0647	0.2128
EC3	0.0439	0.1073	-0.1264	-0.1207	0.2017	0.6458	-0.3243	0.5343
OC1	0.1109	0.2021	-0.0363	-0.0824	-0.2943	0.4417	0.3435	-0.3319
OC2	0.0470	0.1061	0.1740	-0.1026	-0.1912	0.0719	-0.1105	-0.1616
OC3	0.0205	0.0965	0.4480	-0.1206	-0.3342	-0.0425	-0.3834	-0.0357
OC4	0.0232	0.1231	0.3292	-0.0564	-0.1849	-0.0010	-0.2406	0.0422
OPT	0.0227	0.1570	-0.0091	-0.0967	-0.0774	0.2854	0.2257	-0.0924
AmmSulf	0.0325	0.1402	0.0207	-0.0195	-0.0547	0.0791	0.0349	-0.1661
AmmNitr	-0.0213	0.0240	0.3742	0.2219	0.7150	0.2153	-0.1623	-0.4543
Si	0.1619	0.3332	-0.1663	0.0291	0.0445	-0.0831	-0.1144	-0.0523
Al	0.2953	0.5833	-0.2730	0.0740	0.0515	-0.2049	-0.1889	-0.1284
Ca	0.1227	0.2180	-0.0623	0.0448	0.0933	-0.1437	0.0345	0.1404
K	0.0705	0.1562	0.1495	0.0219	-0.0464	-0.0761	-0.1047	0.1282
Fe	0.1035	0.2895	-0.0098	0.0773	0.1158	-0.1168	0.0528	0.0807
Cr	0.6032	-0.2890	0.0218	0.6637	-0.2104	0.1723	0.0042	0.0275
Pb	-0.0094	0.1682	0.2189	0.0373	0.1775	-0.1262	0.5685	0.1983
S	0.0407	0.1323	0.0350	-0.0169	-0.0580	0.0646	0.0280	-0.1578
Se	0.0526	0.0507	0.2235	0.1842	0.0348	-0.1010	0.1181	0.2375
V	0.6797	-0.2692	0.0984	-0.6151	0.2100	-0.1311	0.0576	-0.0214
Zn	0.0189	0.1456	0.3833	0.1043	-0.0295	-0.1048	0.0815	0.2773
Eigenvalue	7.08	3.73	2.36	2.09	1.24	0.88	0.47	0.40

At the Shenandoah National Park 48 out of a total of 1399, observations (Table 6.1-15) form this sample subset. In this case, the trace species have large coefficients for Cr and V in the major Principal Component (Prin1) as well as Prin4. Zn occurs in Prin3 and Pb in Prin7.

6.1.4 Washington DC

The Principal Components (eigenvectors) for Washington DC are given in Table 6.1-17 and re-calculated values, as elaborated upon previously, are shown in Figure 6.1-7 below. Except for EC3 in Prin4, no other species have really high values. What is different to the other three sites is the importance of the optical pyrolysis species OPT. Other high values include Al and Si, possible of crustal origin. AmmNitr and Zn also show up together in Prin3.

Table 6.1-17. Eigenvectors and eigenvalues (as proportions of total variance) for Washington DC. The coefficients greater than 0.7 are highlighted in yellow and those greater than 0.3 and less than 0.7 are highlighted in turquoise.

Eigenvectors	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
Eigenvalue proportion	0.3717	0.2554	0.1163	0.0851	0.0368	0.0311	0.0229	0.0179
AmmNitr	0.0584	0.1730	0.4094	0.0426	-0.3381	-0.6080	0.1830	0.1342
AmmSulf	0.1052	-0.0193	0.0409	-0.0286	0.5674	0.0160	0.3430	0.3753
OC1	0.1634	0.1548	0.3765	0.0914	-0.1687	0.4842	0.5604	-0.4125
OC2	0.1324	0.0895	0.1190	-0.0007	0.4163	0.0389	-0.1418	0.1340
OC3	0.0979	0.1015	0.2987	0.0346	0.2252	0.0210	-0.3105	-0.2436
OC4	0.0855	0.0658	0.2828	0.0755	0.2893	-0.0103	-0.1263	-0.1383
OPT	0.5064	0.6442	-0.3869	-0.3573	-0.0557	-0.0610	-0.0166	-0.0368
EC1	0.1724	0.2227	0.1980	-0.0458	-0.0396	0.1921	-0.1397	0.1967
EC2	0.0281	0.0141	-0.1717	0.1212	-0.1618	0.3068	0.2722	0.4028
EC3	0.1022	0.2437	-0.2881	0.8936	0.0704	-0.1210	0.0114	-0.0555
Al	0.6459	-0.5344	-0.0697	0.0104	-0.0749	-0.2093	0.0622	-0.1526
Si	0.3686	-0.2952	-0.0380	0.0441	0.0037	0.0533	0.0224	0.0879
Ca	0.1481	-0.0726	0.1165	0.0675	-0.1119	0.3232	-0.2367	0.3061
K	0.1525	0.0605	0.1670	0.1238	-0.0060	0.0954	-0.3326	-0.1820
Fe	0.1459	-0.0378	0.1551	0.0907	-0.1633	0.1598	-0.2212	0.2675
Se	0.0619	0.0556	0.1896	-0.0138	0.3171	-0.2112	0.2927	0.0664
Zn	0.0612	0.1129	0.3107	0.0939	-0.2153	-0.1181	-0.0522	0.3679
Total	3.0359	0.9789	1.7098	1.2431	0.5540	0.3517	0.1697	1.1176

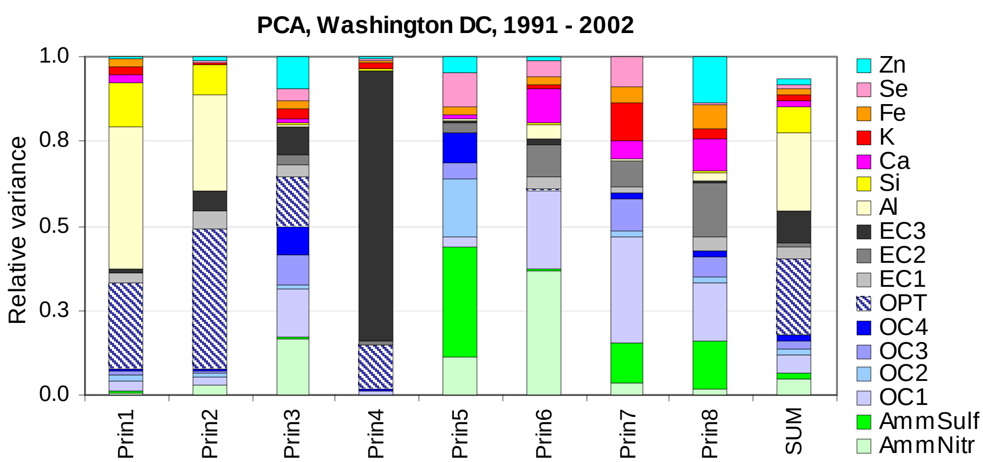

Figure 6.1-7. Principal Component plot for Washington DC, showing the relative variance for each of the species, for the first eight Principal Components. Also shown as SUM is the total weighted variance for each species for the sampling period.

Figure 6.1-8 below portrays the monthly PCA for Washington DC. Washington has variable high $PM_{2.5}$ loadings (13326 – 28875 ng/m^3) showing a sinusoidal pattern, with maxima in July and December. Species of greatest prominence are, as shown previously, OPT showing the greatest prominence during the winter months. Al and Si show the greatest fluctuations in values, together being the largest components in July and August. EC3, possibly from motor vehicle emissions occur throughout the year, but are in greater amounts in December and January but also show up during May and June

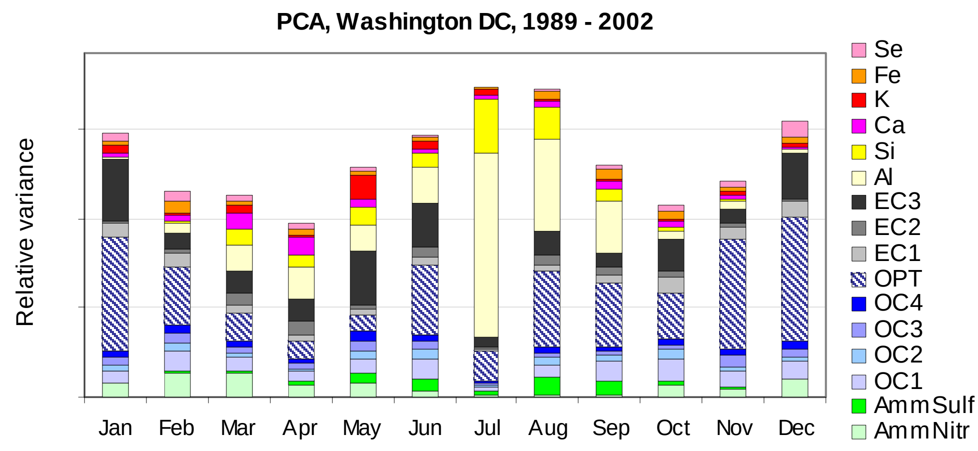

Figure 6.1-8. Principal Component plot for Washington DC, showing the monthly fluctuation in relative variance for each of the species. The bar heights represent the monthly mean $PM_{2.5}$ masses.

Table 6.1-18. Sample subset applied in monthly PCA, for Washington DC.

	Sampling Period		Total Number of Samples
	From	To	
Washington DC	08/31/1988	02/25/2002	537 (1344)

Table 6.1-19. Summary of species of greatest significance in monthly PCA, grouped by season, for Washington DC. Only those species with eigenvector coefficients greater than 0.3 are listed and those with coefficients greater than 0.7 are highlighted in yellow.

Washington DC - Species Grouped by Season

Season	Winter	Spring	Summer	Fall
Months	12, 1, 2	3, 4, 5	6, 7, 8	9, 10, 11
Principle Component				
Prin1	EC1, EC3, OPT	Si, Al, Ca, K	OPT, Si, Al	EC1, OC1, OC3, OC4, OPT, Si, Al
Prin2	AmmNitr, EC3, OC1, OC3, OC4, OPT	AmmNitr, EC2, EC3, OPT, Al	EC3, OPT	AmmNitr, EC3, OC1, OPT
Prin3	AmmNitr, EC3, OC1, OPT	AmmNitr,EC1, OC1, OPT, Al, Ca	AmmSulf, EC3, OC2	EC3, OC1, OPT

A minor sample subset (Tables 5-18, 5-19) was run with a greater number of trace species.

Table 6.1-20. Minor sample subset with trace elements included in PCA, for Washington DC.

	Sampling Period		Total Number of Samples
	From	To	
Washington DC	08/31/1988	02/25/2002	72 (1344)

Table 6.1-21. Eigenvectors ("source profiles") and Eigenvalues ("variances") for a minor sample subset from Washington DC.

			Washington DC					
Eigenvectors								
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
EC	0.1820	0.0872	0.1869	0.1164	0.0489	-0.2925	0.0175	-0.0853
EC1	0.2775	0.0411	0.1176	0.0306	-0.0015	-0.2675	0.0015	-0.0560
EC2	-0.0069	-0.1017	-0.0661	0.0834	0.1782	-0.2346	0.0455	-0.1030
EC3	0.2402	-0.1779	-0.4015	0.6976	0.3677	0.3186	-0.0025	-0.0688
OC	0.1982	0.0309	0.0992	-0.0602	0.0157	0.1223	0.1530	-0.1215
OC1	0.2408	0.0824	0.2362	0.0564	0.1828	-0.2971	0.3510	-0.5014
OC2	0.1425	0.0423	0.0236	-0.0619	0.0304	0.1773	0.1468	-0.0380
OC3	0.1625	0.0694	0.1919	-0.0341	-0.0242	0.1937	0.0891	-0.1599
OC4	0.1196	0.0517	0.1341	-0.0468	0.0225	0.2563	0.1897	-0.0275
OPT	0.6965	-0.3041	-0.3996	-0.3983	-0.1864	-0.0713	-0.0711	0.0912
AmmSulf	0.0283	0.1103	0.0245	-0.1263	0.1914	0.1646	0.3692	0.1303
AmmNitr	0.2027	0.0057	0.2242	0.1458	-0.2548	0.1761	0.1518	0.3784
Si	0.0394	0.0785	0.0564	-0.1478	0.3515	0.0470	-0.2220	0.0538
Al	0.0838	0.1059	0.1292	-0.3198	0.5349	0.2062	-0.3720	0.0971
Ca	0.0738	0.0791	0.0816	-0.0206	0.2150	-0.1337	-0.1987	0.0551
K	0.1381	0.0465	0.1080	-0.0171	0.0739	0.0227	-0.0233	-0.0415
Fe	0.1255	0.0627	0.1233	0.0959	0.1262	-0.1520	-0.2191	-0.0001
Cr	0.0527	0.8257	-0.5105	-0.0196	-0.0679	-0.0976	0.0545	0.0121
Pb	0.1409	0.0855	0.1436	0.1846	0.0790	-0.2805	-0.0074	0.4760
S	0.0261	0.1172	0.0185	-0.1241	0.2126	0.1595	0.3264	0.1140
Se	0.0652	0.0861	0.1655	-0.0166	-0.0125	0.2306	0.1789	0.1788
V	0.2123	0.2685	0.2389	0.1606	-0.3501	0.3383	-0.4481	-0.3084
Zn	0.1592	0.1227	0.1860	0.2686	-0.0522	-0.1333	-0.0556	0.3578
Eigenvalue	11.42	3.00	2.02	1.57	1.41	0.93	0.70	0.49

In the case of Washington DC, the species Cr strongly signifies Prin2 and also Prin3, Pb is in Prin8, V in Prin5, 6 and 7, and Zn in Prin 8 (Table 6.1-21).

In each of the above cases for all four sites, "trace metal" Principal Components were identified. These subsets need to be further investigated with the UNMIX and PMF receptor model, in an effort to resolve quantitative source profiles for point sources such as the metals industries. This information will be correlated with the source inventories and variable meteorological conditions, as well as events such as the Canada fires of August 1998 and the Asian dust storm of April 1998.

9. REFERENCES

Le Maitre, R.W. (1982). Numerical Petrology, Statistical Interpretation of Geochemical Data, Developments in Petrology 8. Elsevier Scientific Publishing Co., 281 pp.