910-R-03-004

Receptor Model Analyses of Aerosol PM2.5 Data from the IMPROVE Monitor at Denali National Park

EPA Technical Report 910-R-03-004

Receptor Model Analyses of Aerosol PM2.5 Data from the IMPROVE Monitor at Denali National Park

EPA Technical Report 910-R-03-004

Authors: Robert Kotchenruther and Rob Wilson, U.S. Environmental Protection Agency Region-10, Office of Environmental Assessment, 1200 Sixth Ave., Seattle, WA 98101.

1. Summary

The UNMIX and PMF receptor models were used to identify sources of $PM_{2.5}$ at Denali National Park. The models used fourteen years of IMPROVE data, from 1988 to 2002. Both models indicated four sources and gave similar results, however, the model performance in both cases was poor due to low filter mass loadings, which lead to high levels of uncertainty in the chemical analyses and relatively poor fitting statistics in the models. Despite the poor fitting statistics in the models, three of four sources were identified. These sources were identified as biomass burning, soil dust, and sulfate and nitrate haze. PMF was found to be better than UNMIX in isolating source signatures and gave results with higher confidence. The fourth and smallest source could not be reliably quantified by UNMIX, but was quantified by PMF. This source remains unidentified.

2. Receptor Models

The two receptor models that were available for this analysis were the U.S. Environmental Protection Agency's (EPA) UNMIX model and the Positive Matrix Factorization (PMF) model. We present below the results of both model analyses. It is advantageous to

compare and contrast the results of both models because their approach to source apportionment is derived from very different mathematical methods, and therefore, when taken together offer solutions with some measure of independence from the mathematics involved. We did not use the EPA's Chemical Mass Balance receptor model due to the lack of available source profile data. We used UNMIX and PMF because they do not require source data and provide a technically valid approach when the number of samples is large.

2.1. The UNMIX Receptor Model

Version 2.3 of the UNMIX (EPA UNMIX 2.3 User Guide, 2002) multivariate receptor model was used in this analysis. Information about, and copies of, the software can be obtained from Gary Norris at the EPA (Norris.Gary@epa.gov). UNMIX uses geometric features in the data called "edges" to constrain the model and determine source apportionment. These "edges" are formed when a particular source contribution falls to zero at the receptor. The edges are, in fact, boundaries in the speciated data that are formed when it is plotted in an n-dimensional 'source space', where n is the number of sources.

As with any model, UNMIX has strengths and weaknesses. The strengths of UNMIX are that it does not require prior knowledge of the number or composition of sources and can independently determine the number of sources. The weaknesses of UNMIX are that it has difficulty identifying ubiquitous sources (where the contribution rarely falls to zero), very infrequent sources, and relatively small sources (contributing less than about 10% to the total mass). Additionally, the UNMIX solution is highly dependant on the species that are selected and UNMIX assumes that source compositions don't change over time. One feature of UNMIX

that is a drawback in this instance, but could be a benefit elsewhere, is that it does not use or require information about measurement uncertainty.

2.2. The Positive Matrix Factorization (PMF) Receptor Model

Positive matrix factorization (PMF) is a form of principal component analysis developed by Pentti Paatero at the University of Helsinki (Paatero and Tapper, 1994) and uses a weighted least squares approach to determine source profiles (User's Guide for Positive Matrix Factorization, 2000). PMF has many of the same strengths and weaknesses as UNMIX, with several important differences. Like UNMIX, PMF does not require prior knowledge of source compositions. However, several advantages of PMF in contrast to UNMIX are that it makes use of the measurement uncertainty to weight data, does not require source contributions to occasionally fall to zero, and is better able to identify small sources. On the other hand, one drawback of using PMF is that one must a priori declare the number of sources prior to running the model. As with UNMIX, PMF also assumes that the source compositions don't change over time.

3. IMPROVE Data and Data Processing

IMPROVE aerosol data from Denali National Park were downloaded from the IMPROVE web site (http://vista.cira.colostate.edu/improve/). The dataset contained 1481 aerosol samples spanning the dates 3/2/88 to 5/26/02. Each sample consisted of chemical and elemental mass analyses for approximately 40 species, $PM_{2.5}$ mass, PM_{10} mass, and numerous derived quantities. Also listed with each species was the associated measurement uncertainty and minimum detection limit (MDL).

Prior to modeling, we analyzed the Denali data looking for the predominant mass constituents. It is widely assumed that the largest sources impacting Denali National Park are biomass burning from summer wildfires, soil dust, and sulfate and nitrate haze (most often associated with emissions from industrial sources). A rough approximation of these source categories can be calculated directly from the IMPROVE dataset, by summing elemental ('EC' in the IMPROVE dataset) and organic carbon ('OMC') for biomass burning, by taking 'SOIL' for soil dust, and by summing ammonium nitrate ('NH_4NO_3') and ammonium sulfate ('$(NH_4)_2SO_4$') for nitrate and sulfate haze. Figures 1a, 1b, and 1c, show the seasonal distributions of 'EC'+'OMC', 'SOIL', and 'NH_4NO_3'+'$(NH_4)_2SO_4$', respectively. Figure 2 shows the monthly average mass distribution of these categories and the distribution of the remaining uncategorized mass. This analysis can only be considered a very rough source apportionment because it assumes that these three source categories have no other significant constituents (e.g., assumes NH_4NO_3 and $(NH_4)_2SO_4$ are the only constituents of haze) and that the above constituents can solely be attributed to these three sources. While this rough source apportionment supports assumptions about the relative importance of sources impacting Denali National Park, a more refined source attribution using receptor models is needed to affirm these results and identify any unexpected sources.

In this receptor modeling analysis, only measurements associated with aerosol fine mass ($PM_{2.5}$) were considered for modeling and are listed in Table 1.

Measurements in the IMPROVE dataset that were reported as less than the minimum detection limit (MDL) were replaced with one half the reported MDL. The uncertainty of the replaced data was also set to one half the MDL, unless the reported analytical uncertainty was larger. Efforts were made to exclude data with large uncertainty relative to the measurement.

While this is not particularly important for PMF modeling, it was important for UNMIX because UNMIX does not make use of measurement uncertainty. First, data were excluded if the uncertainty was more than twice the measured value. After this, the ratio of uncertainty (σ) to measurement (x), σ/x, was calculated for each measurement and the average σ/x computed for each species. Species were excluded from further analyses if the average σ/x exceeded 0.7. Excluding these species had two effects, it eliminated species with high relative measurement error (important for UNMIX), and it excluded species where most of the measurements were replaced with half the MDL (important for both models). Table 1 lists the average σ/x for each species and those species chosen for receptor modeling analyses.

Lastly, aerosol samples were excluded if any of the remaining species had missing values. The resulting dataset retained 1194 of the initial 1481 aerosol samples and 20 of the initial 39 fine mass species. Figure 3a shows the 1194 fine mass measurements plotted sequentially, and Figure 3b shows them plotted with the years overlapping to show the seasonal cycle. The data processing outlined above is similar to that described by Lee et al. (1999) and is consistent with guidelines established in the users' manuals of both PMF and UNMIX.

4. Model Analysis and Results

4.1. UNMIX Analysis

The matrix of 20 species and 1194 samples was input into UNMIX for source-receptor analysis. The model was set to consider fine mass measurements as the total mass, and results were normalized to the fine mass measurements. Data weighting factors were kept at their default values, which decreased the influence of data with the lowest 15% of mass in the model. The model identified four sources, however, as discussed above, the model's diagnostic output

indicated that the model performance was poor. The four-source solution listed a minimum correlation coefficient (r^2) of 0.19, minimum signal to noise ratio of 7.52, and an overall "strength" of 1.33. Recommendations are that these values should be larger than 0.80, 2.0, and 3.0, respectively. Our conclusion was that confidence in the solution should be low. The most likely reason for the poor performance of UNMIX is that $PM_{2.5}$ impacts at Denali National Park are, in general, very small. The mass loadings on many filter samples were low, which caused the relative uncertainty in mass analyses in many cases to be high. Indeed, we eliminated from consideration nearly half (19 of 39) of the available species because of excessively large uncertainty relative to the measurements. A good portion of the remaining data also had high levels of uncertainty (See Table 1). Because UNMIX makes no use of uncertainty information, one would expect model performance to suffer under these conditions.

Table 2 lists the source profiles as mass fractions, the estimated uncertainty in mass fraction, and the relative certainty of each species mass fraction. The relative certainty was calculated as the mass fraction divided by twice the uncertainty. Table 2 shows that despite the model's poor performance, three of the four sources found by the model showed a reasonable amount of confidence in source composition (i.e., high relative certainty).

The relative certainty in the fourth source was less than one for all species, meaning that the uncertainty was more than half the mass fraction. This result is expected because UNMIX has difficulty identifying small sources, those with less than 10% of the total mass (see UNMIX users' manual). The average sample loading was 1.96 $\mu g/m^3$ $PM_{2.5}$ and the average fine mass attributed to each of the sources was 0.87, 0.44, 0.59, and 0.03 $\mu g/m^3$ for sources 1, 2, 3, and 4 respectively. Hence, UNMIX attributed only 1.5% of the total fine mass to Source 4, much less

than 10%. Source 4 was not further considered in the UNMIX analysis because of high uncertainty in the mass fractions and the low attributed percent of total fine mass.

Figures 4a, 4b, and 4c show the fine mass attributed by UNMIX to sources 1, 2, and 3, respectively, with each year plotted overlapping to show the seasonal cycle. Figures 5a, 5b, and 5c show the mass fraction distribution attributed by UNMIX to sources 1, 2, and 3, respectively (please note, the vertical scales differ in many of the figures presented in this report). For a given source, by inspecting the relative abundance of species, which species have the highest relative certainty, and the seasonal distribution of mass, we can surmise a general source category.

The species with the highest mass fractions in Source 1 are EC1, H, OC4, OP, and $(NH_4)_2SO_4$ and those with the largest relative certainties are EC1, H, K, and OC4. These species are indicative of biomass burning, and Figure 4a shows the seasonal pattern of high mass impacts one would expect for wildfires in Alaska (Kasischke et al., 2000).

Source 2 has Al, S, Si, and $(NH_4)_2SO_4$ as the highest mass fractions and Al, Ca, Fe, H, K, Si, and Ti as the species with the highest relative certainty. The EPA Speciate database lists a composite of soil dust having an elemental composition of Si 17.0%, Al 6.3%, Fe 3.0%, and Ca 0.8%, or the elemental ratios Al/Si, Fe/Si, and Ca/Si of 0.37, 0.18, and 0.05, respectively. Roughly the same elemental ratios appear in Source 2: 0.45, 0.29, and 0.12, respectively. UNMIX also adds significant contributions from sulfur and carbon containing species. However, because of the ubiquitous nature of sulfate and carbonaceous species in the data record (see Figure 1) and UNMIX's difficulty with ubiquitous sources (EPA UNMIX 2.3 User Guide, 2002), it is possible that the model is introducing bias in the results for sulfate and carbonaceous species in Source 2.

Figure 4b shows that the highest impacts from Source 2 were between April and June. This corresponds well with the annual spring thaw in this region of Alaska (on average in April, from National Weather Service data), with the months of minimum precipitation (January through May), and with the months of highest wind speed (May and June). However, this time of year also corresponds with the highest frequency of trans-Pacific dust transport events (Husar et al., 1998). Apportioning soil dust between local and trans-Pacific transport sources is beyond the scope of this study.

Source 3 has Na, NH_4NO_3, S, and $(NH_4)_2SO_4$ as the highest mass fractions and Ca, H, K, S, and $(NH_4)_2SO_4$ as the species with the highest relative certainties. The seasonal pattern shown in Figure 4c is indicative of impacts from winter and springtime Arctic haze (Polissar et al., 1999). Arctic haze refers to aerosol, originating from industrial sources in Europe, Asia, and North America, which becomes trapped over the Arctic due to large scale winter and springtime synoptic weather patterns. The mass fractions for Source 3 also shows significant contributions from Na and carbon containing species, which again may be artifacts of the UNMIX model.

4.2. PMF Analysis

The matrix of measurements, 20 species and 1194 samples, and a matrix of measurement uncertainties of the same size and corresponding to the measurements was input into the PMF model for source-receptor analysis. One difficulty in running the PMF model is determining the optimal number of sources the model should solve for. One way to solve this problem is to run PMF using the number of sources determined by UNMIX. Another method that relies solely on PMF is described by Lee at al. (1999). Briefly, the method involves running PMF for sequentially larger numbers of sources, then plotting the maximum value of the rotation matrix

(a diagnostic output of PMF) for each run versus the number of sources solved for. A significant increase in the maximum rotation matrix value indicates the optimal number of sources, Lee et al. recommend using the number of sources just prior to the increase. In conducting this analysis we ran the model five times and generated solutions for two, three, four, five, and six sources. Figure 6 shows the maximum rotation matrix value for each model solution. A significant increase occurs between the four- and five-source solutions, so this method suggests that the four-source solution is optimal and agrees with the results of UNMIX. In the subsequent analysis we will consider only PMF's four-source solution.

In order to account for all of the mass measured at the receptor location, some investigators have introduced scaling factors into the PMF model (Dr. Tim Larson, University of Washington, Personal Communication). The use of scaling factors attempts to create the best match between source contributions determined by the model, and measurements. They are determined through a multiple linear regression, where source contributions are regressed to the total measured mass. The resulting regression slope for each source is the scaling factor for that source. The scaling factors are used as follows: the initial masses attributed to each source are multiplied by the scaling factor for that source, and the fractional compositions determined for that source are divided by the scaling factor. The use of scaling factors makes the assumption that the model has identified all the sources impacting the receptor.

We followed these steps in our analysis and determined the scaling factors for sources 1, 2, 3, and 4 to be 6.17, 0.96, 1.41, and 0.22, respectively. The r^2 of the multiple linear regression was 0.94 and the slope was 0.96, which indicates that there is a good fit between the measured masses and that the sum of the regression adjusted sources.

We summed the mass fractions for each source, after rescaling as above, and these sums were 0.22, 0.51, 2.02, and 0.73 for sources 1, 2, 3, and 4, respectively. The sum for Source 3 was greater than one, which is not physically possible. Hence, the scaling factor determined by multiple linear regression was too small for Source 3. We decided to further adjust Source 3 such that the sum of mass fractions equaled 1.0. We achieved this by using a new scaling factor of 2.83 for Source 3. Using this new scaling factor caused the r^2 for multiple linear regression to fall from 0.94 to 0.91 and the slope of the regression increased from 0.96 to 1.04.

As with UNMIX, the PMF model performance was relatively poor, likely due to low overall mass on the filters causing higher uncertainty in the species analyzed. Table 3 lists the model output for PMF's four-source solution. Listed are the regression adjusted mass fractions, the uncertainty, and the relative certainty in each species mass fraction. Despite PMF's poor performance, the relative certainty of the source profiles was generally higher for PMF than for UNMIX, and PMF does a much better job quantifying a fourth source. Figures 7a, 7b, 7c, and 7d show the adjusted fine mass attributed to sources 1, 2, 3, and 4, respectively, plotted with each year overlapping to show the seasonal cycle. Figures 8a, 8b, 8c, and 8d show the distribution of mass fractions for sources 1, 2, 3, and 4, respectively.

PMF also outputs information on the amount of variability in each species explained by each source. This explained variability (EV) can be a useful tool for qualitative source attribution. Those species that have high EV are on some level the most important in determining that source in the model. Table 4 lists the EV for each species and the variability left unexplained. Figures 9a, 9b, 9c, and 9d show the distribution of EV for sources 1, 2, 3, and 4, respectively. Figure 9e shows the amount of variability in each species that remained unexplained.

As with UNMIX, general source categories can be surmised by inspecting the relative abundance of species within a source, which species have the highest relative certainty in mass fraction, and the seasonal distribution. Additionally, the EV of each species can be evaluated.

Inspecting the distributions for Source 1 in Figures 8a and 9a, EC1, H, OC4, OP, and $(NH_4)_2SO_4$ dominate the mass fraction and EC1, H, K, OC4, and OP have the highest EV. These patterns, along with the seasonal distribution depicted in Figure 7a, are consistent with a biomass burning source.

The mass fraction distribution for Source 2 in Figure 8b shows that Al, Fe, Si, and $(NH_4)_2SO_4$ dominate the mass fraction and the distribution of EV in Figure 9b shows that Al, Ca, Fe, K, Si, and Ti have the highest EV. These species, the seasonal distribution depicted in Figure 5b, and the ratios Al/Si, Fe/Si, and Ca/Si of 0.33, 0.30, and 0.14, suggest a soil dust source.

For Source 3, Figures 8c and 9c show that S and $(NH_4)_2SO_4$ dominate both the mass fraction and EV and suggest primary or secondary industrial sources.

For Source 4, Ca, EC1, Na, NH_4NO_3, OC4, and Zn dominate the mass fraction depicted in Figure 8d. However, Ca, Cu, Pb, and Zn have the highest certainty in mass fraction and, as depicted in Figure 9d, Cu, Pb, and Zn have the highest EV. We are at present unsure how to attribute this source.

One hypothesis we considered for Source 4 was fugitive dust from the Red Dog zinc and lead mine north of Kotzebue, Alaska. The concentrated product of this mine is primarily sub-20 micrometers in size, and therefore could potentially be transported long distances. However, the mine concentrates are primarily zinc sulfide and lead sulfide and there was virtually no sulfur attributed to Source 4. Therefore, we discounted this possibility.

4.3. Comparison of PMF and UNMIX Results

The mass fraction distribution pattern for Source 1 seen in Figures 5a and 8a for UNMIX and PMF, respectively, are nearly identical. However, the magnitude of the UNMIX mass fractions are roughly twice that of PMF. Despite this difference, the total mass attributed to Source 1 by each model agrees quite well. This is seen by comparing the seasonal distributions for Source 1 represented in Figures 4a and 7a, for UNMIX and PMF, respectively. They are nearly identical in pattern and scale. Figure 10 shows a scatter plot of the masses attributed to Source 1 by UNMIX plotted against that by PMF (please note, UNMIX allows small negative values in source contributions in order to reduce bias in the results). A linear regression of the data gives a slope of 0.92 and r^2 of 0.91, with UNMIX tending to assign slightly more mass to Source 1 than PMF. The chemical composition of Source 1, the seasonal distribution of mass, and the good agreement between UNMIX and PMF give us a high confidence in identifying Source 1 as biomass burning.

There are larger differences between PMF and UNMIX for Source 2. Figure 11 shows a scatter plot of the masses attributed to Source 2 by UNMIX plotted against Source 2 for PMF. A linear correlation of the data shows a strong correlation, with an r^2 of 0.97, but the slope is 0.40. This indicates that both models track the same source almost perfectly, but UNMIX attributes more than twice the mass than PMF to this source. An analysis of the mass fraction distribution allocated by each model provides some explanation of the difference. The primary elemental constituents of soil are Al, Si, Ca, Fe, and Ti. The mass fraction ratios of Al, Fe, Ca, and Ti to Si are similar between PMF and UNMIX solutions. These ratios are: Al/Si, 0.33 and 0.45, Fe/Si, 0.30 and 0.29, Ca/Si, 0.14 and 0.12, and Ti/Si, 0.03 and 0.03, for PMF and UNMIX,

respectively. While these ratios are similar, the magnitude of the mass fractions attributed to these species differed.

IMPROVE data protocols recommend calculating a total soil component using the formula 2.20*Al + 2.49*Si + 1.63*Ca + 2.42*Fe + 1.94*Ti. The scalar factors multiplying each of these species take into account the most common form of oxides of these species found in soil. Using this formula, the mass fraction attributed to soil dust in Source 2 is 43.1% for UNMIX and 89.6% for PMF. Further analyses showed that UNMIX attributed 32.7% of the mass fraction to the sum of $(NH_4)_2SO_4$ and NH_4NO_3, and 7.2% to the sum of carbonaceous species, EC1+OC4+OP, whereas PMF attributed only 7.2% and 0.8%, respectively, to these summed sources. It is plausible that there is a source that combines soil, sulfate, nitrate, and carbonaceous species. One example is trans-Pacific transport of mixed Asian dust and industrial pollution. However, due to UNMIX's difficulty with ubiquitous sources (i.e., sulfate, nitrate, and carbonaceous species in this case), it is more plausible to explain model differences as UNMIX introducing a ubiquitous source bias while PMF demonstrates a better capacity to isolate the soil signature. Hence, we recommend using the PMF mass fractions for Source 2.

We have a high confidence in identifying Source 2 as soil dust due to the chemical composition of the mass fractions, the seasonal distribution of attributed mass, and the good correlation between UNMIX and PMF (despite the difference in absolute mass attributed).

There are large differences between PMF and UNMIX for Source 3. Figure 12 shows a scatter plot of the masses attributed to Source 3 by UNMIX plotted against Source 3 for PMF. A linear regression of the data gives a poor fit, with a r^2 of 0.50 and a slope of 0.40.

Despite this poor regression, there are some similarities between PMF and UNMIX in their mass fraction distributions. The ratio in mass fraction for $S/(NH_4)_2SO_4$ is 0.25 for both

PMF and UNMIX, and the ratio $NH_4NO_3/(NH_4)_2SO_4$ is 0.08 and 0.16 for PMF and UNMIX, respectively. While these ratios in mass factions are similar, their percent contribution is very different. The percent contribution of nitrogen and sulfur species, $(NH_4)_2SO_4+NH_4NO_3+S$, is 94.9 and 65.9 for PMF and UNMIX, respectively, the percent contribution of carbonaceous species, EC1+OC4+OP, is 1.1 and 7.8 for PMF and UNMIX, respectively, and the percent contribution of metal species is 3.9 and 9.7 for PMF and UNMIX, respectively. Hence, Source 3 for PMF is almost entirely made up of nitrogen and sulfur species, whereas for UNMIX they are only about two thirds of the mass. UNMIX attributes about seven times as much mass as PMF to carbonaceous species and over twice the mass to metal species. These complex differences between PMF and UNMIX likely contribute both to the poor correlation seen in Figure 12 and to the increased mass attributed to Source 3 by UNMIX. Based on the differences between PMF and UNMIX discussed above for Source 2, it is likely that Source 3 for UNMIX includes some bias as discussed above and perhaps some mixing of sources, whereas Source 3 for PMF is a better representation of the actual source profile.

Therefore, overall, PMF is better able to separate the sources, as well as quantify a fourth source. Having said that, the PMF result likely overestimates the contribution of some of these sources. This is because, having regressed the results of PMF to the total fine mass, we made the assumption that four sources were all that impact the receptor. This assumption neglects those sources that PMF cannot resolve, for example, sources whose composition changed significantly over time or sources that impacted the receptor too infrequently.

Figure 13 depicts the monthly average contribution to total fine mass for each PMF source. The actual contribution of these sources likely lies somewhere between that depicted in Figure 13 and the rough estimate depicted in Figure 2.

5. Discussion

As part of the Clean Air Act, the Regional Haze Rule requires States to mitigate anthropogenic sources of visibility degradation such that visibility in national parks reaches background levels within 60 years. Hence, it is important to establish the magnitude and chemical composition of both background (i.e., natural and non-Alaskan) and man made sources.

Receptor modeling in this study has quantified four sources and we have identified three. Each of the three identified sources is potentially a mixture of Alaskan man made, Alaskan natural, and non-Alaskan sources.

Impacts from biomass burning during summer months at Denali are likely from Alaskan wildfires. However, there is a significant contribution to aerosol mass from biomass burning during winter months, on average about 1 $\mu g/m^3$ (see Figure 13), which is likely not from wildfires. This wintertime biomass burning source could represents a hemispheric background, local wood stoves, or some combination of these and other sources.

The highest soil dust impacts occur in April and May, contributing on average about 0.5 $\mu g/m^3$. There are a number of potential sources for this dust, including wind generated dust from within Alaska, dust brought in from Asia through trans-Pacific transport, and dust generated from vehicular traffic on unpaved roads.

Sulfate and nitrate haze have their highest impacts in January through March, with monthly averages around 1 $\mu g/m^3$. Arctic haze potentially makes up a large portion of sulfates and nitrates measured during the winter months, but the partitioning between Arctic haze and local sources needs to be established. Additionally, during summer months, the concentration of

sulfate and nitrate aerosol approaches 0.5 $\mu g/m^3$ and cannot be attributed to Arctic haze. The source of this aerosol during summer has not been determined.

A regional haze monitoring pilot study has recently been proposed by the State of Alaska Department of Environmental Conservation. This study will compare aerosol measurements within Denali Park with those made concurrently over a longitudinal transect across Alaska. Concurrent aerosol measurements gathered over a broad area of Alaska can establish what portion of the aerosol measured in Denali Park is the result of local sources, and what portion is due to regional/synoptic scale phenomena like Arctic haze and trans-Pacific transport. However, it must be noted that the magnitude and frequency of these regional/synoptic scale phenomena likely changes from year to year. Hence, in order to accurately quantify the local and background impacts to visibility degradation, a multiyear study is recommended.

In addition to a regional impacts study, a more detailed chemical, elemental, and isotopic analysis of the aerosol data could allow 'fingerprints' of specific sources to be established and better quantified. This, in conjunction with further receptor modeling, could help in determining sources more specifically.

6. References

Henry, R. C., and G. A. Norris, EPA UNMIX 2.3 User Guide, U. S. Environmental Protection Agency, 2002.

Husar R.B., et al., The Asian Dust Events of April 1998, J. Geophys. Res., 106, 18317-18330, 2001.

Kasischke, E.S., K. P. O'Neill, N. H. F. French, and L. L. Bourgeau-Chavez, Controls on patterns of biomass burning in Alaskan boreal forests, pages 173-196 in E. S. Kasischke and B.

J. Stocks, editors, "Fire, Climate Change, and Carbon Cycling in the North American Boreal Forest", 2000, Springer-Verlag, New York.

Lee et al., Application of positive matrix factorization in source apportionment of particle pollutants in Hong Kong, Atmospheric Environment, 33, 3201-3212, 1999.

Paatero, P., and U. Tapper, Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111-126, 1994.

Paatero, P., User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, University of Helsinki, 2000.

Polissar, A. V., P. K. Hopke, P. Paatero, Y. J. Kaufmann, D. K. Hall, B. A. Bodhaine, E. G. Dutton, and J. M. Harris, The aerosol at Barrow, Alaska: long-term trends and source locations, Atmospheric Environment, 33, 2441-2458, 1999.

Table 1. Measured Aerosol Fine Mass ($PM_{2.5}$) Species from the Denali National Park IMPROVE Sampler, the Average Value of Species' Relative Measurement Uncertainty*, and Those Species Chosen For Receptor Modeling**.

Available Species for Receptor Modeling	Average Value of Relative Measurement Uncertainty	Species Used in Receptor Modeling***
Aluminum, Al	0.56	Al
Arsenic, As	0.81	
Bromine, Br	0.26	Br
Calcium, Ca	0.14	Ca
Chloride, Chl	0.80	
Chlorine, Cl	0.77	
Chromium, Cr	0.78	
Copper, Cu	0.66	Cu
Elemental Carbon Fraction 1, EC1	0.45	EC1
Elemental Carbon Fraction 2, EC2	0.76	
Elemental Carbon Fraction 3, EC3	1.03	
Iron, Fe	0.09	Fe
Hydrogen, H	0.09	H
Potassium, K	0.15	K
$PM_{2.5}$ Fine Mass, MF	0.19	MF
Magnesium, Mg	0.84	
Manganese, Mn	0.72	
Molybdenum, Mo	0.92	
Nitrite, N2	0.99	
Sodium, Na	0.51	Na
Nickel, Ni	0.88	
Ammonium Nitrate, NH_4NO_3	0.50	1.29(NO_3)
Organic Carbon Fraction 1, OC1	0.98	
Organic Carbon Fraction 2, OC2	0.86	
Organic Carbon Fraction 3, OC3	0.80	
Organic Carbon Fraction 4, OC4	0.69	1.4(OC4)
Organic Carbon Fraction from Pyrolesis, OP	0.67	1.4(OP)
Phosphorus, P	0.96	
Lead, Pb	0.41	Pb
Rubidium, Rb	0.75	
Sulfur, S	0.06	S
Selenium, Se	0.78	
Silicon, Si	0.12	Si
Ammonium Sulfate, $(NH_4)_2SO_4$	0.12	1.375(SO_4)
Strontium, Sr	0.68	Sr
Titanium, Ti	0.46	Ti
Vanadium, V	0.77	
Zinc, Zn	0.18	Zn
Zirconium, Zr	0.92	

*The relative measurement uncertainty is defined here as the ratio of the uncertainty to measured value.
**Species were chosen for receptor modeling if their relative measurement uncertainty was less than 0.7.
***Some species were multiplied by constants to account for the mass of the most prevalent form. For example, nitrate (NO_3^+) usually exists as ammonium nitrate (NH_4NO_3), and the molecular weight of NH_4NO_3 divided by the molecular weight of NO_3^+ is 1.29.

Table 2. Source Composition Results from UNMIX Receptor Modeling of IMPROVE Data from Denali National Park.

Species	Source 1 Composition			Source 2 Composition			Source 3 Composition			Source 4 Composition		
	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*
Al	0.0003	0.0005	0.3	0.0437	0.0037	5.8	-0.0011	0.0009	-0.6	0.0005	0.0615	0.0
Br	0.0001	0.0000	2.9	0.0003	0.0001	2.6	0.0008	0.0001	5.5	0.0049	0.0511	0.0
Ca	0.0007	0.0003	1.3	0.0117	0.0010	6.1	0.0085	0.0006	6.8	0.0228	0.2147	0.1
Cu	0.0000	0.0000	1.1	0.0001	0.0000	1.4	0.0003	0.0000	3.3	0.0017	0.0238	0.0
EC1	0.0957	0.0104	4.6	0.0188	0.0054	1.7	0.0400	0.0048	4.1	0.0540	0.3508	0.1
Fe	0.0010	0.0004	1.4	0.0283	0.0022	6.6	0.0025	0.0006	2.3	0.0122	0.0404	0.2
H	0.0553	0.0030	9.2	0.0297	0.0032	4.6	0.0349	0.0026	6.7	0.0734	0.4535	0.1
K	0.0045	0.0005	5.0	0.0139	0.0009	7.5	0.0053	0.0004	6.6	0.0118	0.0714	0.1
Na	0.0012	0.0012	0.5	0.0014	0.0035	0.2	0.0726	0.0081	4.5	0.0076	0.1331	0.0
NH_4NO_3	0.0146	0.0027	2.7	0.0095	0.0040	1.2	0.0769	0.0072	5.4	0.0951	0.8698	0.1
OC4	0.0983	0.0066	7.4	0.0197	0.0051	1.9	0.0251	0.0043	2.9	0.1109	0.9705	0.1
OP	0.0996	0.0133	3.8	0.0337	0.0067	2.5	0.0131	0.0053	1.2	0.0998	0.6422	0.1
Pb	0.0000	0.0000	0.8	0.0003	0.0000	3.2	0.0008	0.0001	3.3	0.0026	0.0271	0.0
S	0.0186	0.0042	2.2	0.0782	0.0100	3.9	0.1150	0.0071	8.1	0.0922	0.4520	0.1
Si	0.0031	0.0012	1.3	0.0975	0.0093	5.2	0.0044	0.0020	1.1	0.0577	0.3395	0.1
$(NH_4)_2SO_4$	0.0685	0.0162	2.1	0.3179	0.0417	3.8	0.4674	0.0304	7.7	0.2955	2.3902	0.1
Sr	0.0000	0.0000	0.3	0.0001	0.0000	2.3	0.0000	0.0000	0.2	0.0094	0.1346	0.0
Ti	0.0002	0.0001	1.4	0.0025	0.0002	7.2	0.0006	0.0001	3.3	0.0045	0.0376	0.1
Zn	0.0003	0.0000	2.7	0.0002	0.0001	1.2	0.0021	0.0004	2.8	0.0021	0.0261	0.0

*Here relative certainty is defined as the mass fraction over twice the uncertainty.

Table 3. Source Composition Results from PMF Receptor Modeling of IMPROVE Data from Denali National Park.

Species	Source 1 Composition			Source 2 Composition			Source 3 Composition			Source 4 Composition		
	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*	Mass Fraction	Uncertainty	Relative Certainty*
Al	0.00000	0.00000	0.0	0.07062	0.00041	86.7	0.00000	0.00000	0.0	0.00000	0.00005	0.0
Br	0.00010	0.00000	30.6	0.00020	0.00001	11.4	0.00050	0.00001	50.0	0.00176	0.00006	14.2
Ca	0.00020	0.00002	6.4	0.03007	0.00023	66.4	0.00465	0.00006	36.2	0.06712	0.00100	33.7
Cu	0.00001	0.00000	3.6	0.00006	0.00001	5.3	0.00008	0.00000	13.6	0.00382	0.00006	30.2
EC1	0.04827	0.00080	30.3	0.00001	0.00015	0.0	0.01117	0.00141	4.0	0.15314	0.01726	4.4
Fe	0.00050	0.00001	20.9	0.06319	0.00021	148.6	0.00251	0.00004	35.8	0.00832	0.00035	11.9
H	0.05921	0.00020	149.3	0.00269	0.00055	2.5	0.00091	0.00029	1.6	0.01236	0.00361	1.7
K	0.00303	0.00004	38.8	0.02658	0.00024	55.2	0.00289	0.00007	19.3	0.01211	0.00081	7.5
Na	0.00041	0.00006	3.2	0.00000	0.00003	0.0	0.02116	0.00031	33.9	0.11288	0.00454	12.4
NH_4NO_3	0.00498	0.00019	12.8	0.00001	0.00015	0.0	0.05735	0.00067	43.0	0.23058	0.00932	12.4
OC4	0.06000	0.00104	28.9	0.00098	0.00368	0.1	0.00001	0.00014	0.0	0.05428	0.02577	1.1
OP	0.04670	0.00087	26.8	0.00662	0.00299	1.1	0.00001	0.00020	0.0	0.00008	0.00167	0.0
Pb	0.00000	0.00000	0.0	0.00023	0.00001	11.6	0.00038	0.00000	39.2	0.00501	0.00009	28.0
S	0.01108	0.00015	37.9	0.01911	0.00064	14.8	0.17841	0.00058	154.9	0.00079	0.00440	0.1
Si	0.00180	0.00005	16.6	0.21171	0.00084	126.6	0.00667	0.00015	22.3	0.01320	0.00163	4.0
$(NH_4)_2SO_4$	0.03060	0.00060	25.4	0.07227	0.00369	9.8	0.71317	0.00227	156.8	0.00074	0.01237	0.0
Sr	0.00000	0.00000	3.8	0.00023	0.00001	16.6	0.00007	0.00000	12.8	0.00041	0.00004	5.1
Ti	0.00009	0.00001	7.1	0.00615	0.00006	49.7	0.00006	0.00002	1.9	0.00142	0.00020	3.6
Zn	0.00003	0.00000	8.1	0.00011	0.00001	5.2	0.00000	0.00000	0.0	0.04916	0.00019	126.2

*Here relative certainty is defined as the mass fraction over twice the uncertainty.

Table 4. The Fraction of Variability in the Mass of Each Species that is Explained by a Four-Source Solution Using PMF. The Remaining Unexplained Variability is Also Listed.

Species	Variability of Species Explained by Source*				
	Source 1	Source 2	Source 3	Source 4	Unexplained
Al	0.000	0.595	0.000	0.000	0.405
Br	0.198	0.050	0.380	0.064	0.307
Ca	0.025	0.362	0.231	0.154	0.228
Cu	0.031	0.040	0.169	0.322	0.437
EC1	0.583	0.000	0.066	0.042	0.309
Fe	0.074	0.650	0.151	0.027	0.098
H	0.895	0.007	0.007	0.005	0.086
K	0.322	0.314	0.142	0.030	0.192
Na	0.020	0.000	0.344	0.093	0.544
NH_4NO_3	0.102	0.000	0.455	0.089	0.354
OC4	0.684	0.002	0.000	0.014	0.300
OP	0.604	0.013	0.000	0.000	0.383
Pb	0.000	0.068	0.360	0.198	0.374
S	0.143	0.033	0.753	0.000	0.071
Si	0.071	0.655	0.113	0.012	0.149
$(NH_4)_2SO_4$	0.097	0.027	0.791	0.000	0.084
Sr	0.040	0.219	0.228	0.064	0.449
Ti	0.074	0.478	0.022	0.025	0.401
Zn	0.049	0.021	0.000	0.850	0.080

*Assumes a total variability in each species of 1.0.

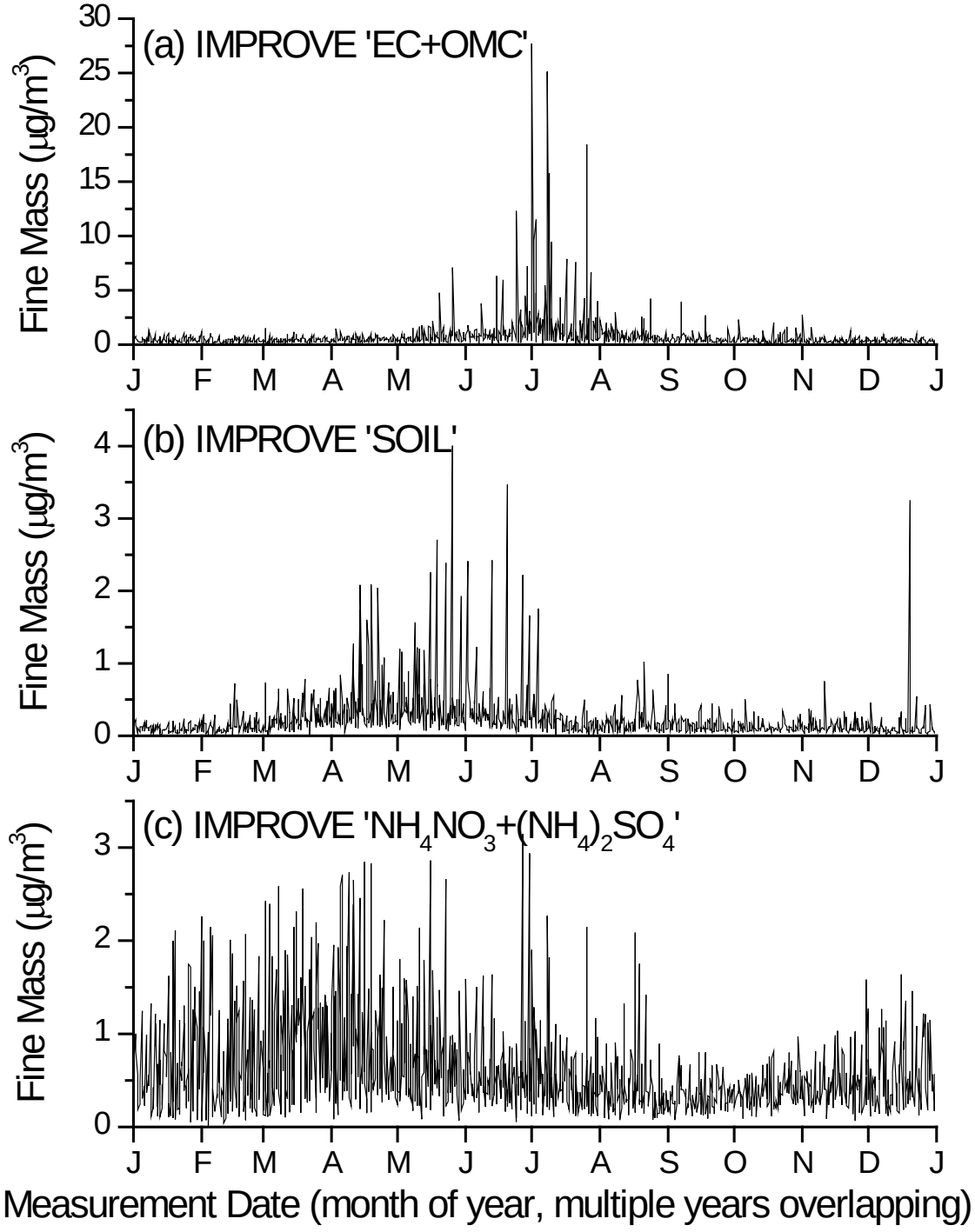

Figure 1. IMPROVE Fine Mass ($PM_{2.5}$) Data Plotted with Multiple Years Overlapping. Three Source Categories are Depicted, (a) IMPROVE 'EC'+'OMC' as an Estimate for Biomass Burning, (b) IMPROVE 'SOIL' as an Estimate for Soil Dust, and (c) IMPROVE 'NH_4NO_3'+'$(NH_4)_2SO_4$' as an Estimate for Sulfate and Nitrate Haze.

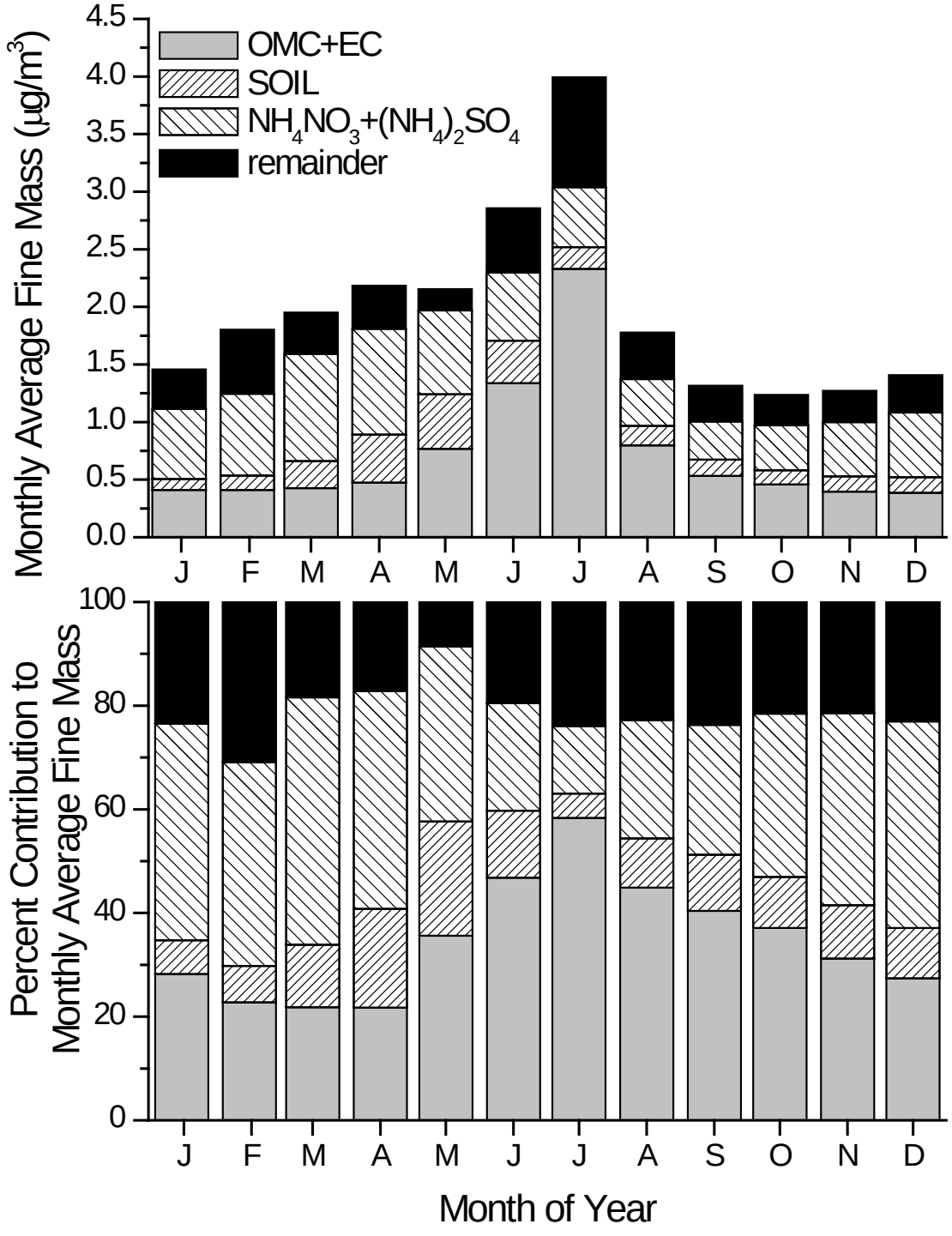

Figure 2. Monthly Averages of IMPROVE Fine Mass ($PM_{2.5}$) Data. Four Categories are Depicted, IMPROVE 'EC'+'OMC' as an Estimate for Biomass Burning, IMPROVE 'SOIL' as an Estimate for Soil Dust, IMPROVE 'NH_4NO_3'+'$(NH_4)_2SO_4$' as an Estimate for Sulfate and Nitrate Haze, and the Remaining Mass on the Filter.

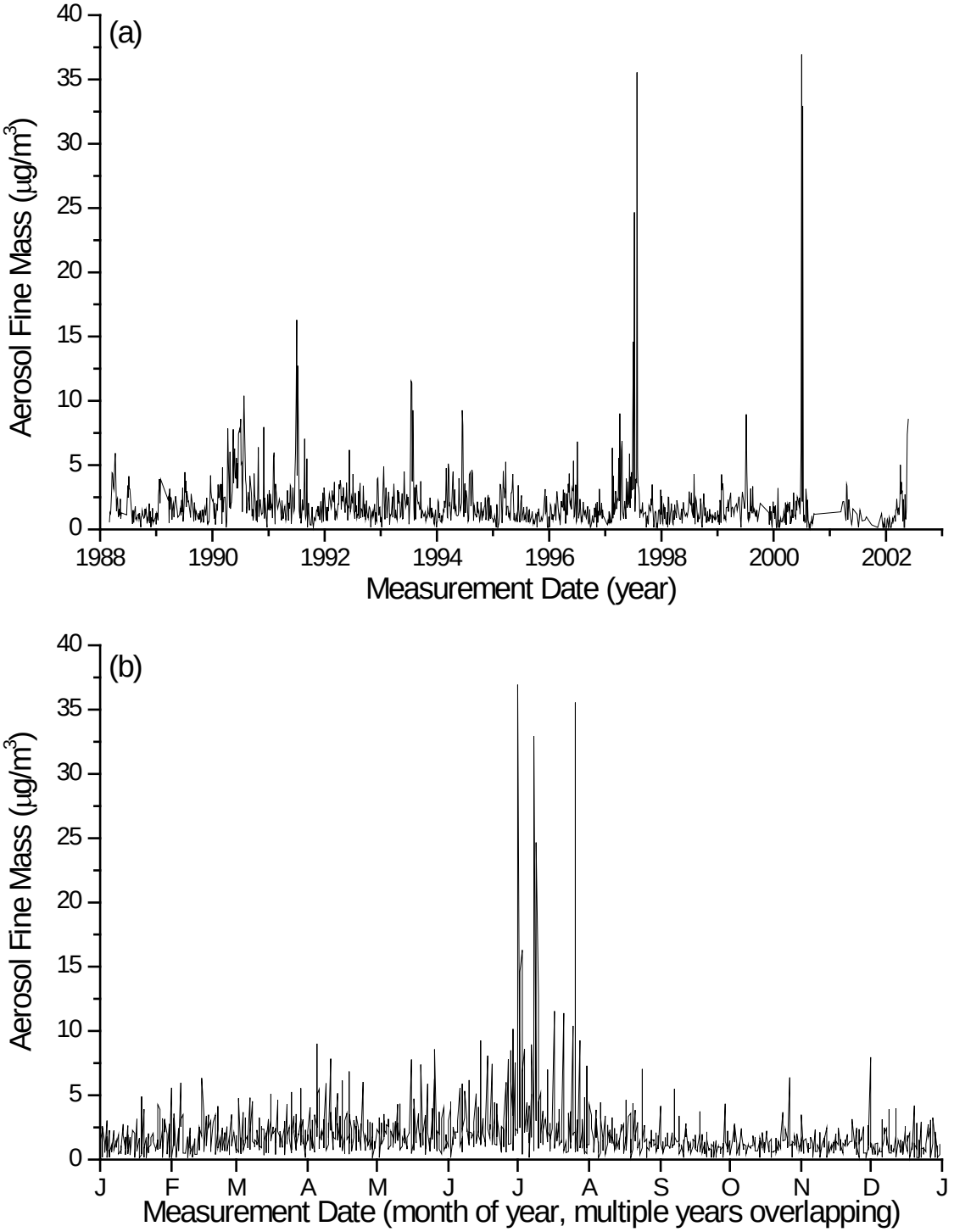

Figure 3. $PM_{2.5}$ Aerosol Fine Mass Measurements From the Denali National Park IMPROVE Sampler Plotted (a) sequentially and (b) with years overlapping.

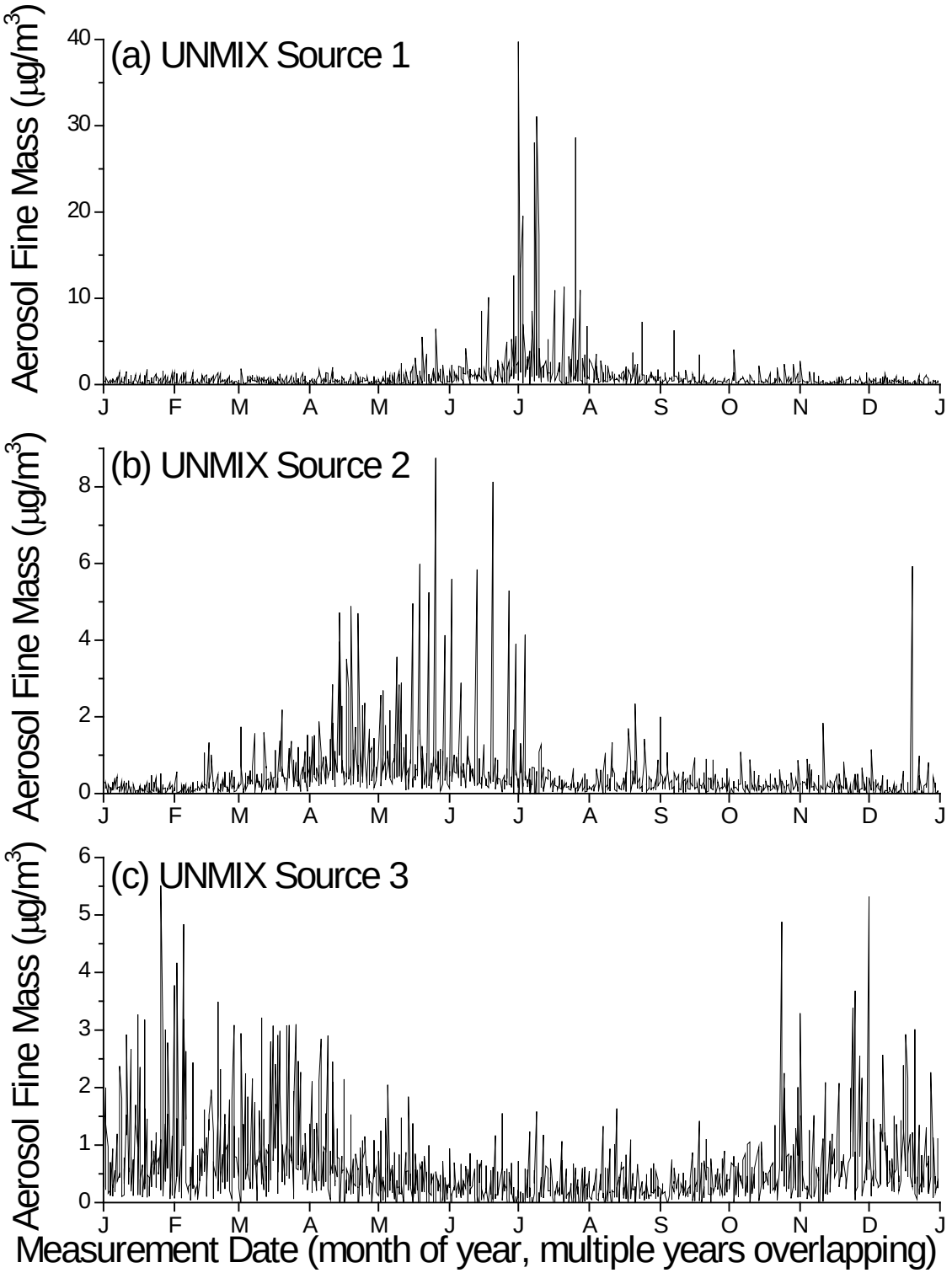

Figure 4. Seasonal Distribution of Aerosol Fine Mass ($PM_{2.5}$) Attributed to Sources 1 (a), 2 (b), and 3 (c), by the UNMIX Receptor Model.

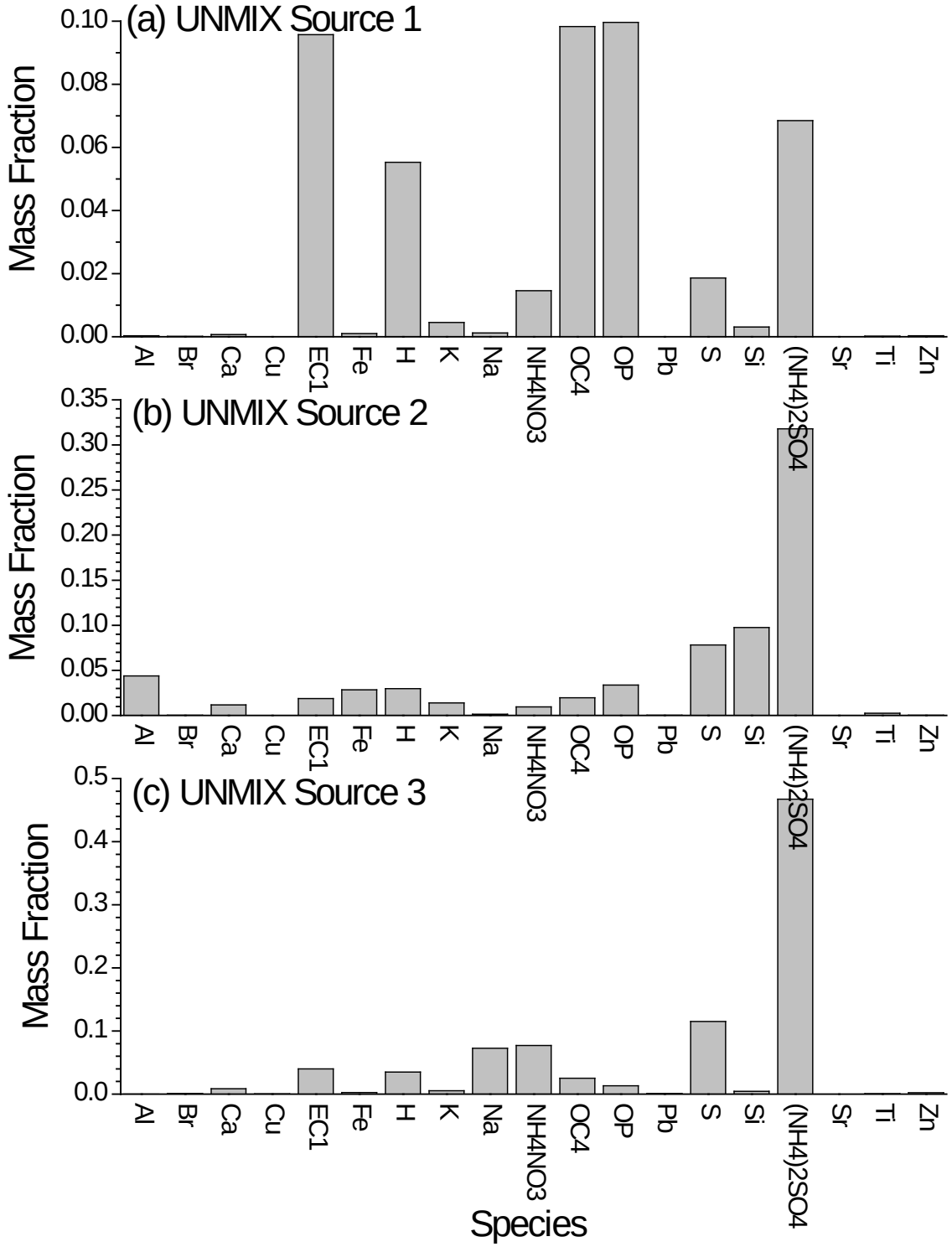

Figure 5. Source Composition Mass Fractions for Sources 1 (a), 2 (b), and 3 (c), Determined by the UNMIX Receptor Model.

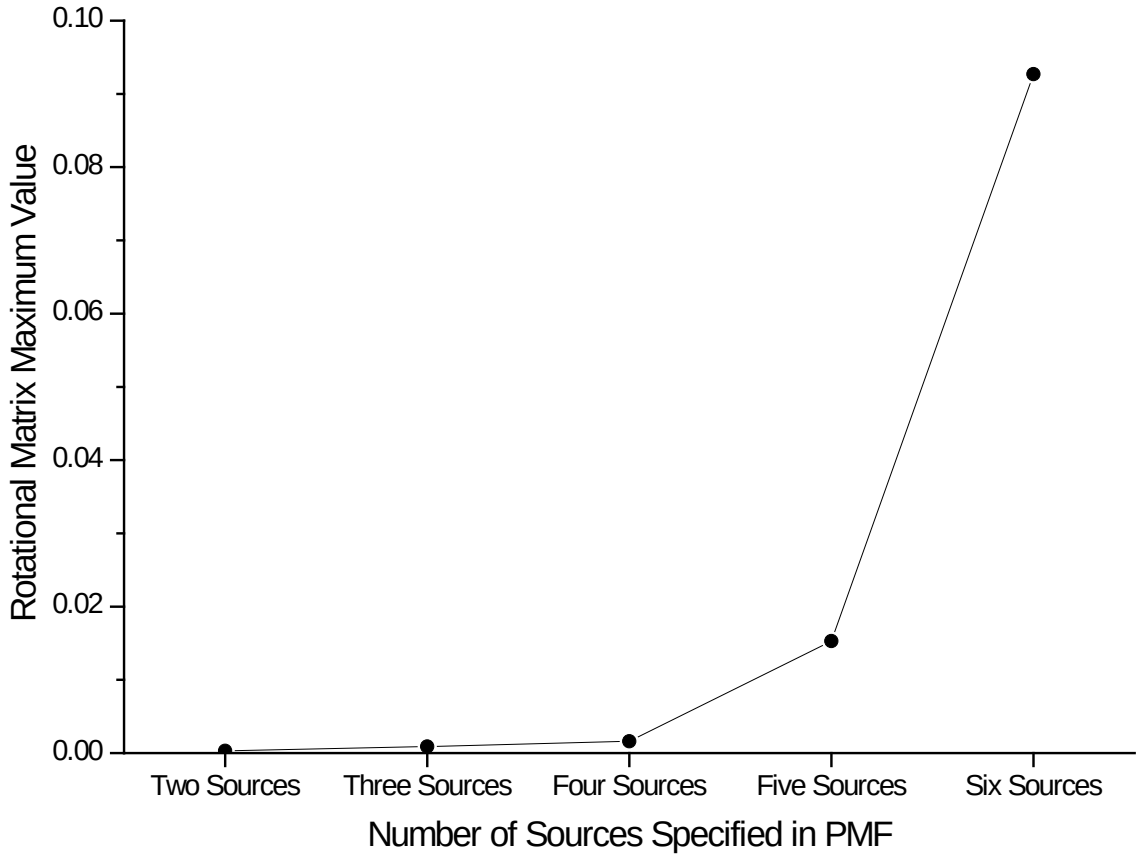

Figure 6. Rotation Matrix Maximum Value for Different Numbers of Sources in the PMF Receptor Model.

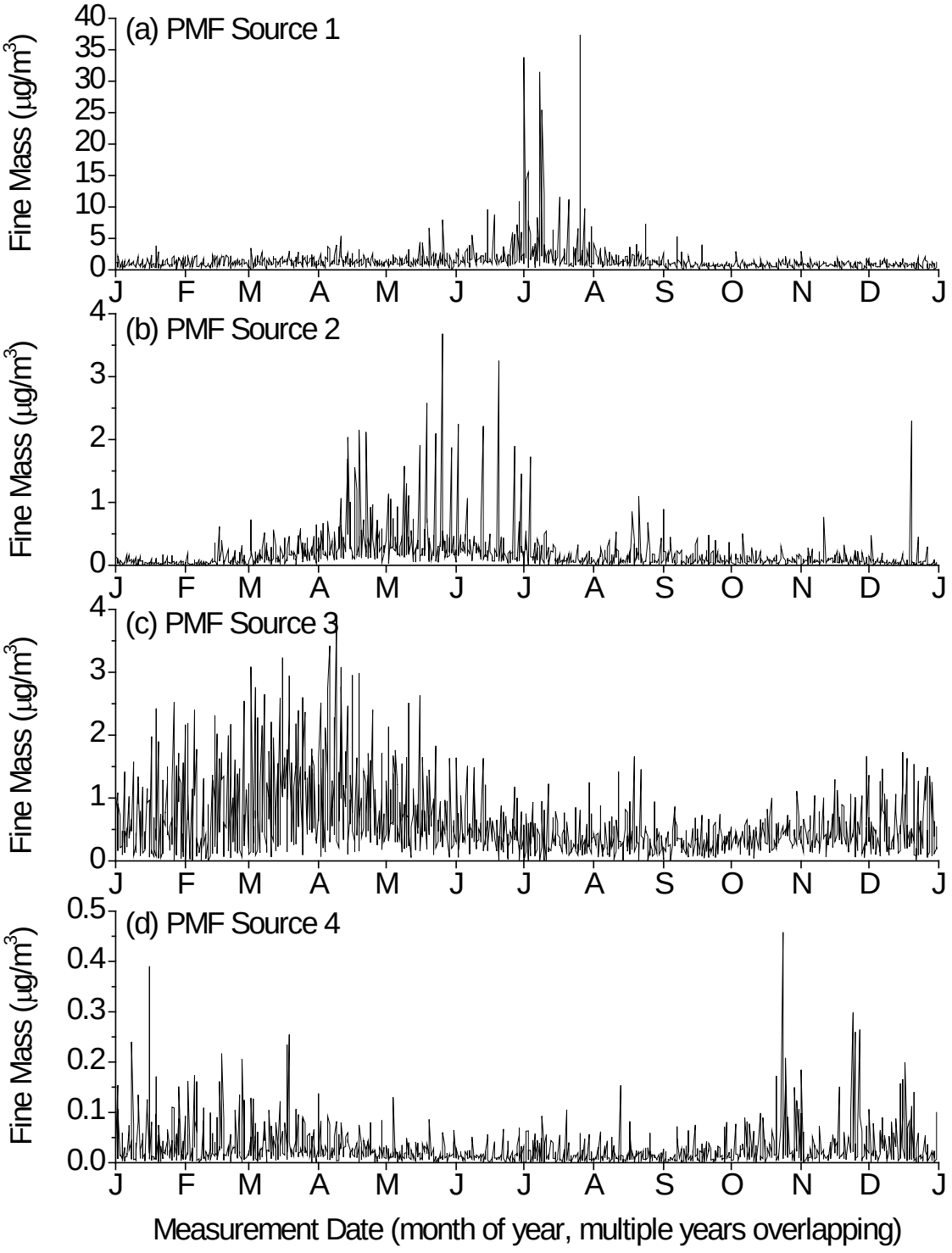

Figure 7. Seasonal Distribution of Aerosol Fine Mass ($PM_{2.5}$) Attributed to Sources 1 (a), 2 (b), 3 (c), and 4 (d) by the PMF Receptor Model.

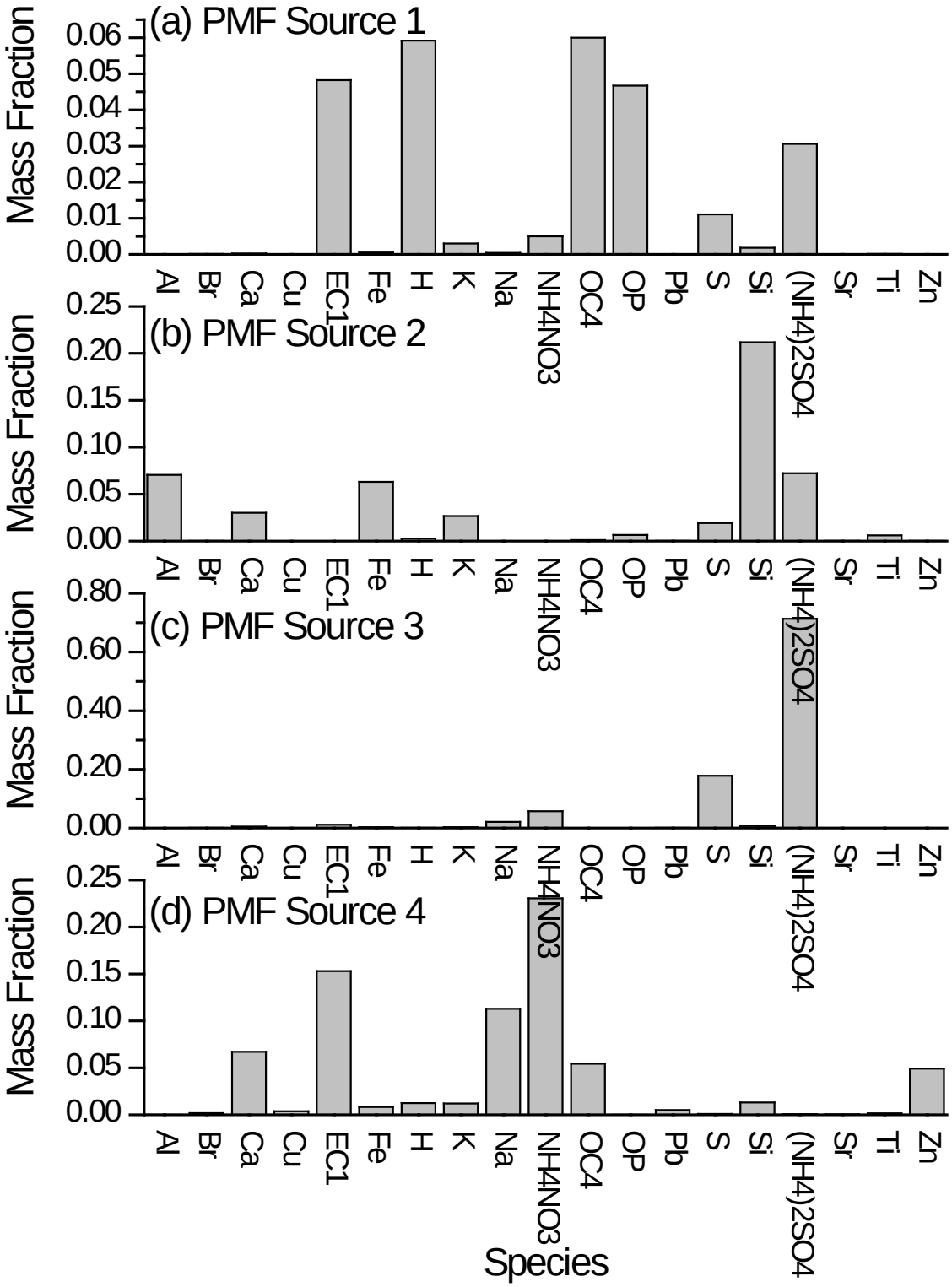

Figure 8. Source Composition Mass Fractions ($PM_{2.5}$) for Sources 1 (a), 2 (b), 3 (c), and 4 (d) Determined by the PMF Receptor Model.

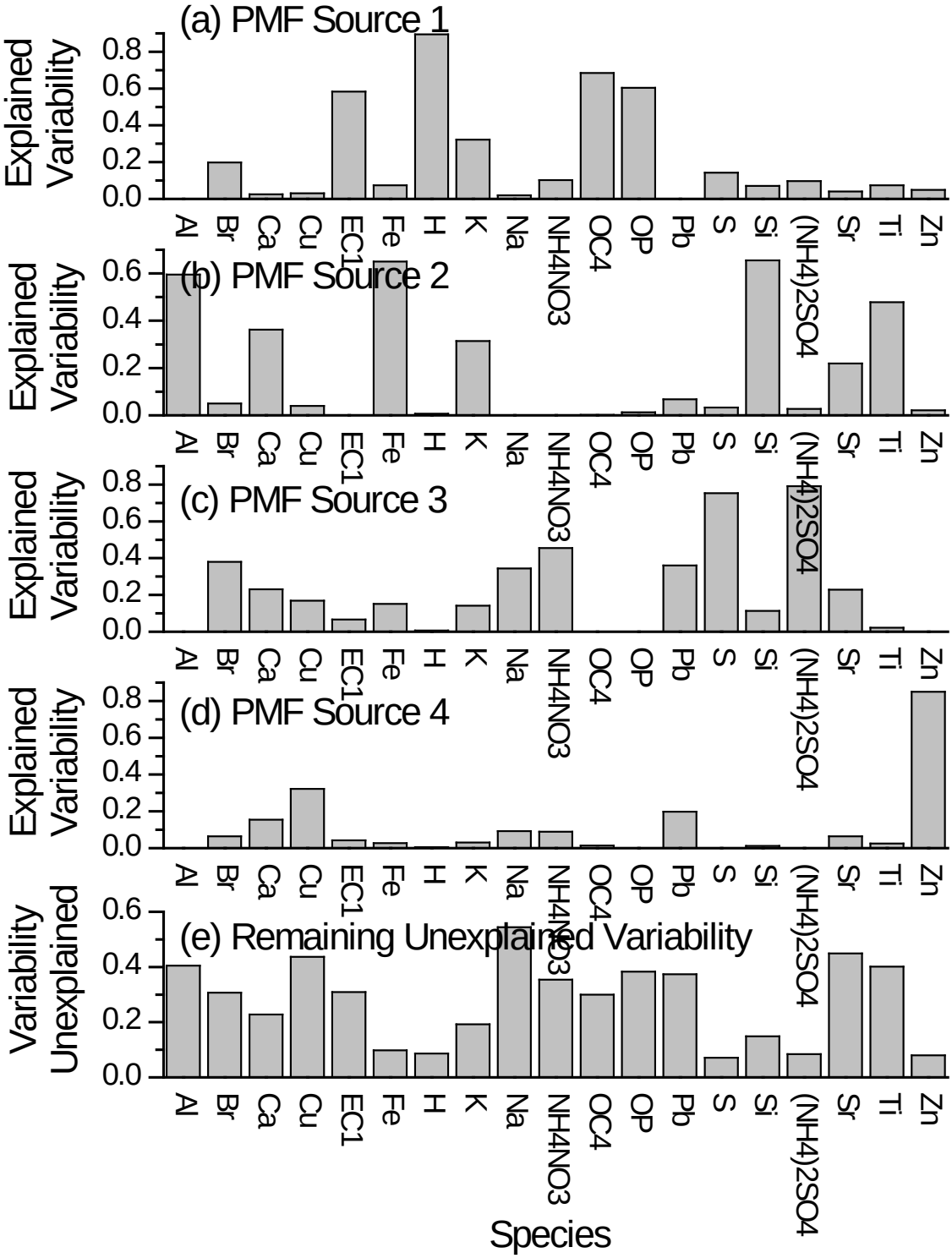

Figure 9. Explained Variability for Species in Sources 1 (a), 2 (b), 3 (c), and 4 (d) Determined by the PMF Receptor Model and the Remaining Unexplained Variability (e).

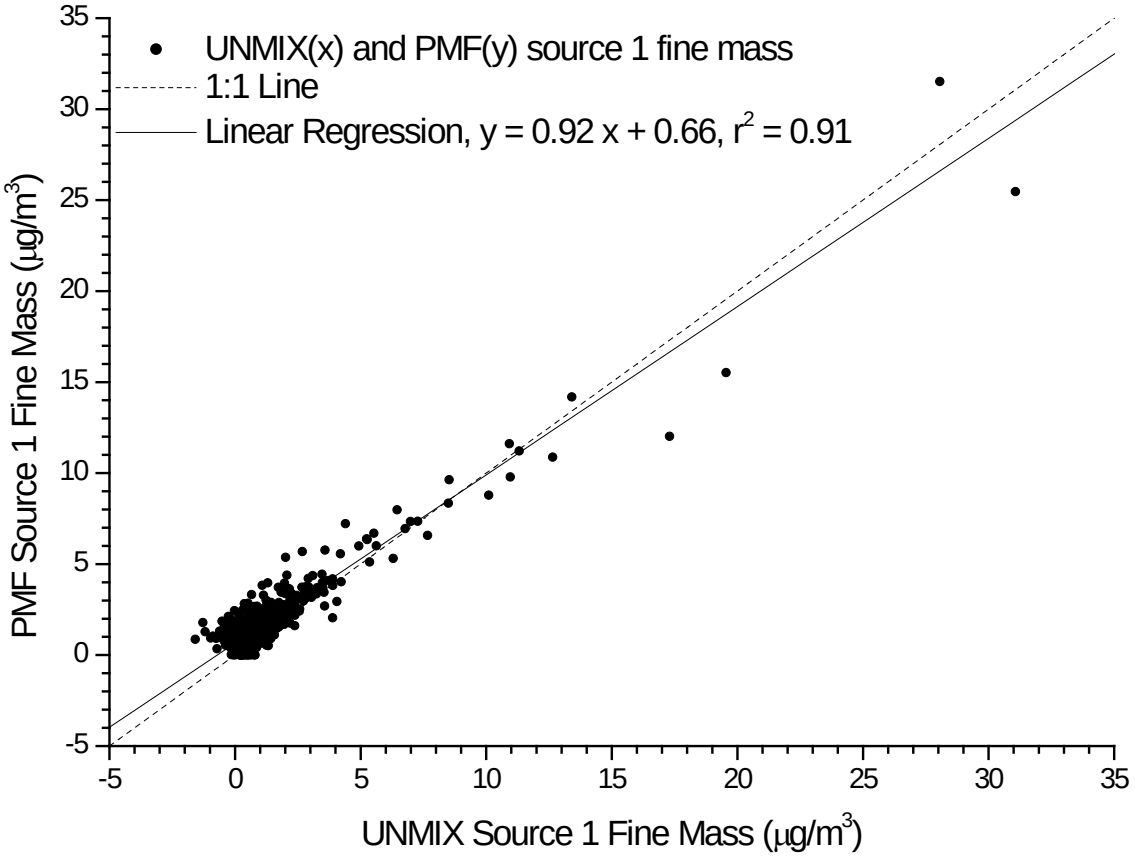

Figure 10. Aerosol Fine Mass ($PM_{2.5}$) Attributed to Source 1 by the UNMIX Receptor Model, plotted on the x axis, and by the PMF Receptor Model, plotted on the y axis.

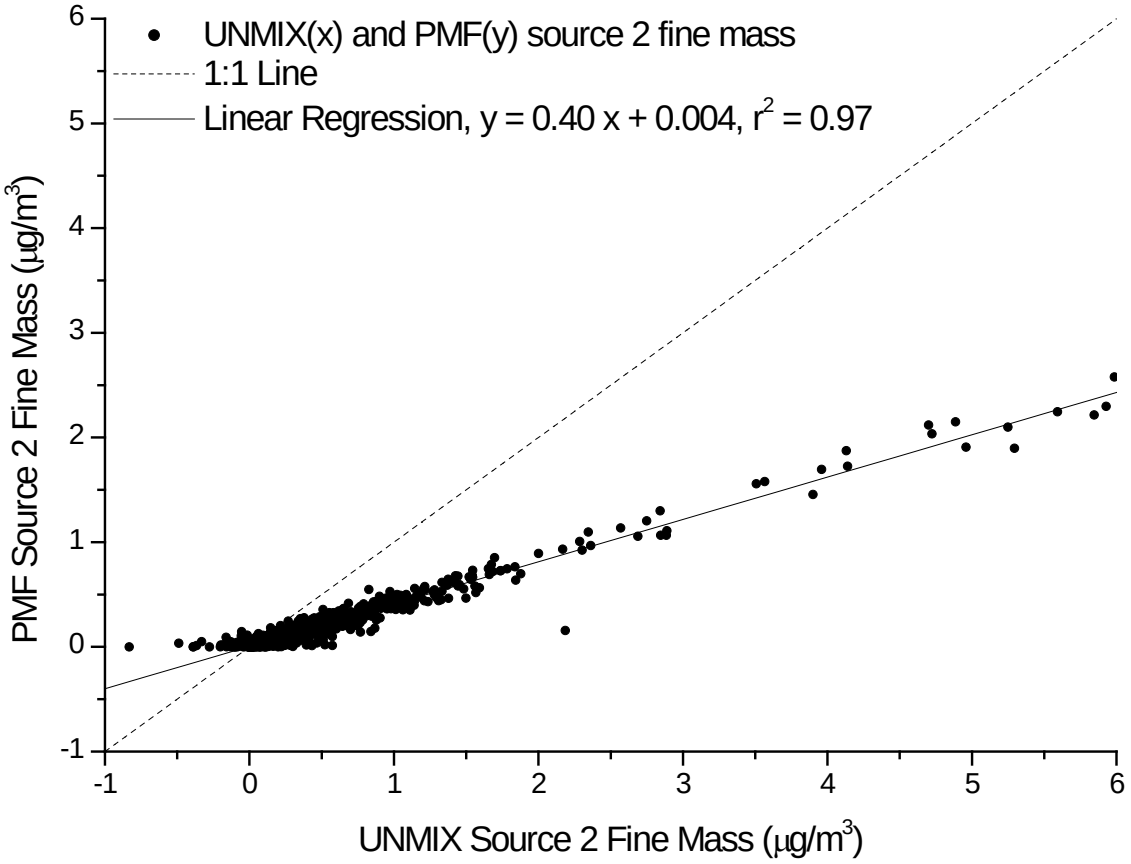

Figure 11. Aerosol Fine Mass ($PM_{2.5}$) Attributed to Source 2 by the UNMIX Receptor Model, plotted on the x axis, and by the PMF Receptor Model, plotted on the y axis.

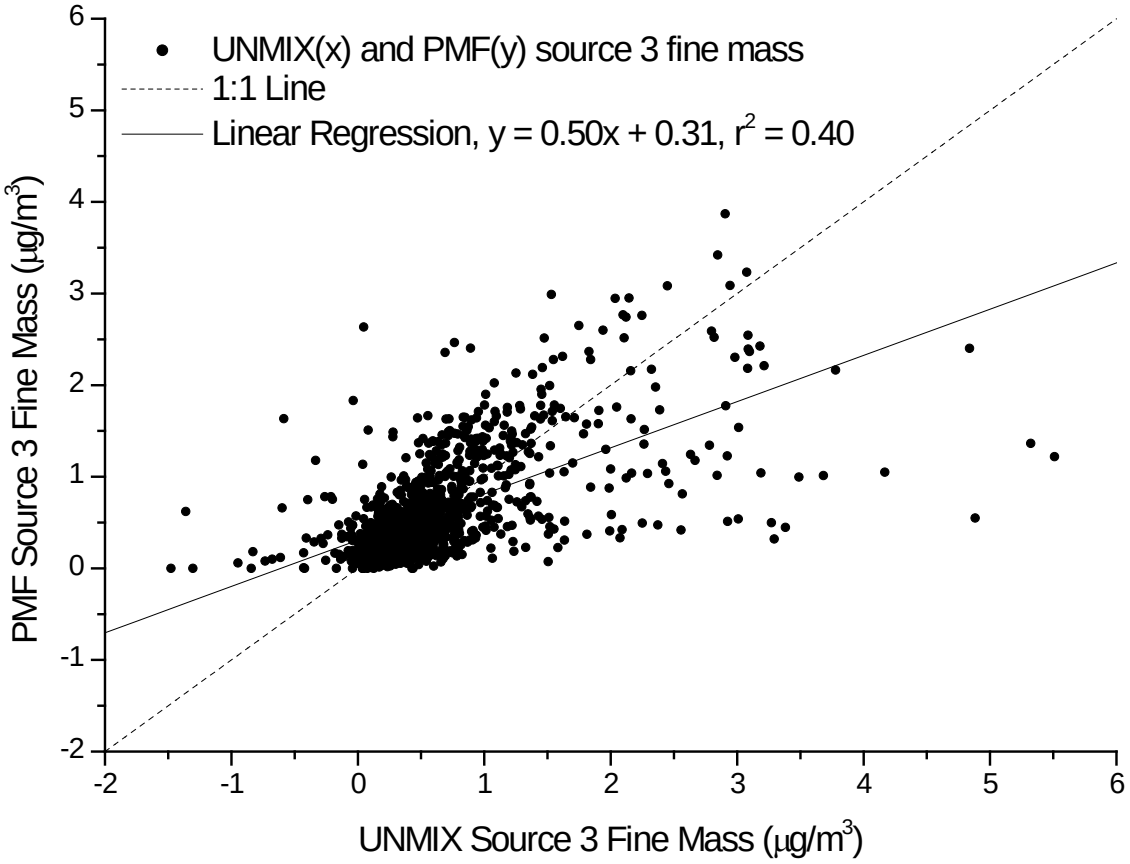

Figure 12. Aerosol Fine Mass ($PM_{2.5}$) Attributed to Source 3 by the UNMIX Receptor Model, plotted on the x axis, and by the PMF Receptor Model, plotted on the y axis.

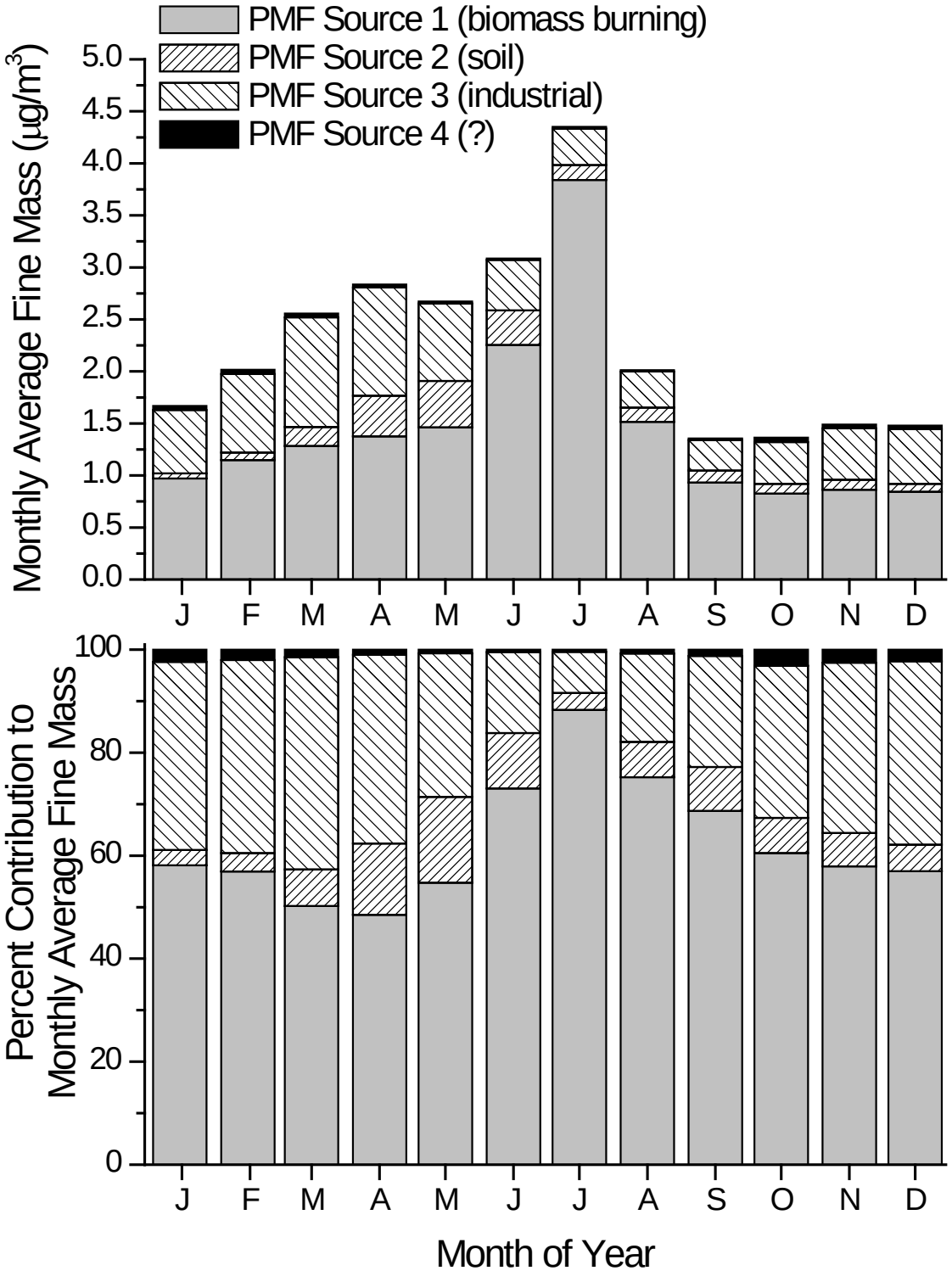

Figure 13. The Monthly Average Contribution of Sources as Determined by PMF Receptor Modeling, Depicted as Both Total Mass and Percent Contribution.