

Source Apportionment of Sediment PAHs in Lake Calumet, Chicago: Application of Factor Analysis with Nonnegative Constraints

PHILIP A. BZDUSEK AND
ERIK R. CHRISTENSEN*

Department of Civil Engineering and Mechanics and Center for Great Lakes Studies, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201

AN LI AND QIMENG ZOU

School of Public Health, University of Illinois at Chicago, 2121 West Taylor Street, MC-922, Chicago, Illinois 60612-7260

A factor analysis model with nonnegative constraints (FA) was used to apportion the sources of PAHs found in sediments of Lake Calumet and surrounding wetlands in southeast Chicago. Source profiles and contributions, with uncertainties, are determined with no prior knowledge of sources. The model includes scaling and backscaling of data with average PAH concentrations without sample normalization. This work is a follow-up to a study that used a chemical mass balance (CMB8.2) model to apportion sources to the same data set. Literature source profiles, modified based on gas/particle partitioning of individual PAHs, from eight PAH sources were considered for comparison. FA results for a two-source solution indicate coke oven (45%) and traffic (55%) are the primary PAH sources to Lake Calumet sediments. A six-source FA solution indicates that coke oven (47%) and traffic (45%) related sources are major PAH sources and wood burning–coal residential (2.3%) is a minor PAH source. From the six-source solution, two coke oven profiles are observed, a standard coke oven profile (33%), and a degraded or second coke oven profile (14%) low in phenanthrene and pyrene. Observed traffic related sources include gasoline engine (36%) exhaust and traffic tunnel air (9.3%). This work supports the previous study of Lake Calumet PAHs by CMB model. In addition, FA provides new insights since wood burning and secondary coke oven profiles were not recognized in the CMB model.

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are an environmental concern due to their carcinogenic properties (1). PAHs are compounds, containing typically two to eight aromatic rings, which are produced by high-temperature reactions, such as incomplete combustion and pyrolysis of fossil fuels and other organic material (2). PAH producing activities include the following: combustion of fuels in vehicular engines, power generation from fossil fuels, coke production, wood burning, incineration of industrial and domestic wastes, oil refinery and chemical engineering processes, etc. (3). Byproducts containing significant amounts of PAHs have been dumped on the land, in the water, or buried at subsurface sites (4). PAH-bearing airborne particulates, generated from PAH producing activities, are transported in

the atmosphere and usually find their final destination in soils and sediments of aquatic systems (4). Once deposited, some PAH compounds, such as benz[a]anthracene (BaA), may be biodegraded aerobically in the top sediment layers (5). Considering degradation and unique source profiles, PAHs can serve as tracers of pollution sources.

Factor analysis (FA) modeling has been utilized for source characterization for air pollution control (6–9). FA is able to operate with minimal knowledge about source characteristics, meteorological conditions, transport and incorporation mechanisms of pollutants into particles, etc., which are not always available (10). To assess the efficiency and accuracy of prediction, FA model results should be compared with existing information about sources, transport processes, and characteristics of receptor sites (10). A few applications of FA modeling to characterize pollution sources in aquatic sediments have been published (10, 12–14). In one case (10), time records rather than PAH source profiles were considered, and there was no backscaling. In ref 12, sample normalization was used in polytopic vector analysis on polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDD/Fs), which tends to hinder source resolution (11), and there was no comparison with CMB model results. In other cases (13, 14) factor analysis was used without nonnegative constraints making it impossible to generate actual source profiles or contributions. Chemical mass balance (CMB) modeling has been used successfully to apportion PAH sources in aquatic sediments (4, 15, 16).

The goal of this work is to determine the major PAH sources to sediments of Lake Calumet, Illinois, using a FA model with nonnegative constraints. From this point forward FA refers to our model, which includes nonnegative constraints. The work will compare results obtained from the study by Li et al. (15), which used CMB8.2 to apportion PAH sources for Lake Calumet, Illinois. To our knowledge this is the first application of an FA model for source apportionment of PAHs in sediments. The current study requires no input of source profiles and will provide uncertainties for the factor-loading and factor-score plots. New findings and conclusions will be discussed.

Factor Analysis Model with Nonnegative Constraints. A detailed discussion of the FA model can be found in Rachdawong and Christensen (17), Ozeki et al. (7), and Imamoglu (11). The fundamental equation underlying the principal component analysis based model is

$$\mathbf{D} = \mathbf{C} \times \mathbf{R} \quad (1)$$

$$(\mathbf{m} \times \mathbf{r}) \quad (\mathbf{m} \times \mathbf{n}) \quad (\mathbf{n} \times \mathbf{r})$$

where the data matrix **D** is factored into its components; the factor-loading matrix **C** and the factor-score matrix **R**, representing source compositions, and source contributions, respectively. In addition, *m*, *n*, and *r* are the number of PAHs, sources, and samples, respectively.

If the following assumptions are met, the factor loading and factor score matrices can be interpreted in terms of PAH source profiles and PAH source contributions: (1) the source profiles do not change significantly in relative input patterns from sources to receptor site; (2) the variation of individual PAH flux is proportional to its concentration; (3) the total PAH flux in a given time period is the sum of the total PAH fluxes in that time period from all considered sources; (4) source profiles and contributions do not covary; and (5) all sampling sites are affected by the same major sources (18).

Selection of the number of significant factors and FA model validation are described, in detail, by Imamoglu et al.

* Corresponding author phone: (414)229-4968; fax: (414)229-6958; e-mail: erc@uwm.edu.

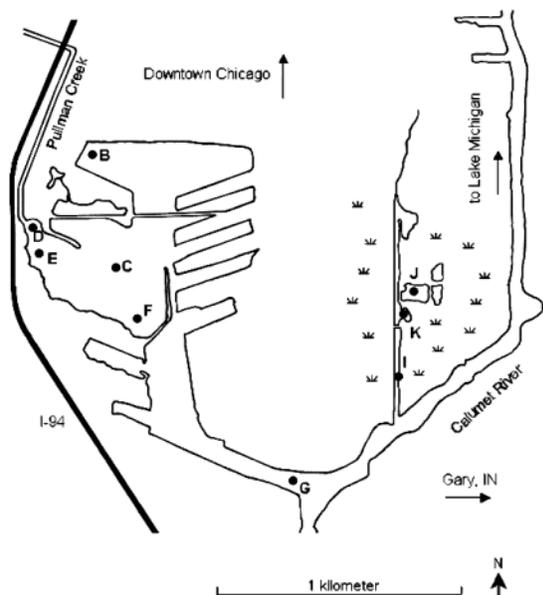


FIGURE 1. Sampling locations in Lake Calumet area.

(19). In short, the cumulative percent variance, coefficient of determination (20), Exner function (21) and convergence of the nonnegative rotations were all considered for determining the number of significant factors. Average scaling of PAHs was applied to the data matrix to reduce the effect of bias from PAHs with high concentrations (17, 22). The covariance matrix, of dimensions $m \times m$, was calculated using the scaled data in the FA model. A matrix with eigenvectors as columns, and a matrix with eigenvalues as diagonal elements were determined. Using the eigenvector and eigenvalue matrices a loading matrix is obtained. Backscaling of the loading matrix is performed by multiplying each column of the **C** matrix, with the average of each PAH. The factor-loading matrices were then normalized to 1, for presentation purposes. The inverse of this normalization was applied to the **R** matrix, thus the correct **D** matrix is obtained from multiplication of **C** and **R**.

Monte Carlo simulation was used to generate artificial data sets to test the effectiveness of the FA model (11). Successful reproduction of the factor-loading and factor-score matrices was observed (11).

Methodology

Sampling and Chemical Analysis. Information regarding the study area, sampling, sediment characterization, chemical analysis and source profiles are described in detail in refs 15 and 23. Below is a brief summary.

Lake Calumet is located 15 miles south of metropolitan Chicago. Surface water inflow is via Pullman Creek, a drainage ditch (Figure 1). Outflow is by way of the Calumet River. Since rapid industrial development began in the 1860s with the laying of railroad tracks, the size of Lake Calumet has been significantly reduced due mainly to the landfill of refuse from nearby industries and municipal wastes, including the ash and cinders from coal combustion for home heating and cooking. Industries around the lake include petroleum refineries, chemical plants, building material companies, and grain processors as well as incinerators, landfills, and illegal dumping sites. Interstate I-94 (Figure 1) opened in 1962; traffic flows to I-94 have continually increased. A coke oven plant, which was active at the time of sampling, is located less than 2 miles north of sampling location J. A north-south rail corridor runs across wetlands to the west of the lake. A number of residential communities are within a few miles of the lake.

Nine sediment cores were collected in June 1997, from the Lake Calumet area (Figure 1). Core I was collected by a push corer; all other cores were collected by a gravity corer. Cores were sectioned, and cores D, E, I, J, and K were dated using ^{210}Pb analysis. Sediment samples were Soxhlet extracted, cleaned-up, and concentrated. Instrumental analysis was completed using a HP Model 5890-II gas chromatograph equipped with DB-5 capillary column (30 m \times 0.25 mm i.d., 0.25 μm film) and a flame ionization detector. The identities of PAHs in selected samples were confirmed using a HP model 6890+/5973 GC/MS. Surrogate recovery ranged from 48% to 90% with an arithmetic mean of 72.8%. Duplicate samples yielded average relative standard deviations ranging from 10.3% to 25.6% among the 16 PAHs.

Source Profiles. PAH profiles from coal combustion and traffic related sources were reported in ref 15. Wood burning, which is a potential source in the study area, but was not investigated in the previous work (15), was included in this study. A total of 32 wood burning source profiles were collected from 12 publications. Two wood burning source profiles were generated from these source profiles using the same procedure described in ref 15. Briefly, if the original source data were reported as the sum of particulate and gaseous PAHs, they were converted to particle-only profiles by applying the P% values to the concentrations of individual PAHs. The original source profiles containing only the particulate PAHs were not modified. P% values are the percentages of particulate to total airborne PAHs and were estimated from reported gas-particle partitioning reported in more than a dozen literature sources (15). After P% corrections, PAH concentrations were normalized to that of benzo[e]pyrene (BeP). Then, the fractions were calculated, and the fractional composition is the source profile.

References for wood burning PAH profiles are given in the Supporting Information. Wood burning profiles contained the largest relative standard deviations (47%–197%) among the sources, reflecting variations in types of wood burned, combustion conditions, sampling method, and analytical procedures. Due to the large influence of P% on low molecular weight compounds, and the fact that a wood burning profile is composed of lower molecular weight compounds, two wood burning profiles are shown in Figure 2. Wood Burning I is from particulate phase only measurements, correction was not needed. Wood Burning II was derived using all the literature source profiles, with P% values applied to convert "total" to particulate PAHs. Comparison between Wood Burning I and Wood Burning II validates the P% correction approach in constructing PAH profiles. Fluoranthene (FlA) and pyrene (Py) are slightly overcorrected, while benzo[b]-fluoranthene + benzo[k]fluoranthene (Bb+kFlA), benzo[e]-pyrene (BeP), benzo[a]pyrene (BaP), and indeno[123-cd]-pyrene (IP) are somewhat undercorrected.

In addition to the phase speciation of the PAHs, source profiles may also be affected by chemical, photochemical, and biological degradation of PAHs before they enter the sediment, as discussed in ref 15. The stability during atmospheric transport of PAHs was not considered due to the proximity of the sampling sites to sources. Partitioning between water and settling particles in the lake may also affect the abundance of PAH species before they reach bottom sediments. Due to the strong affinity of PAHs to particulate matter, their loss to the lake water may not be significant. In this study, possible PAH degradation may be observed through modified source profiles.

FA Modeling. The inputs to the FA model are the number of sources and the data matrix. The initial data matrix consisted of 17 PAHs and 75 samples, with units, of $\mu\text{g}/\text{kg}$. Naphthalene was not included in the modeling, because of the high uncertainties in its source profiles and possible evaporative losses during chemical analysis of the sediment

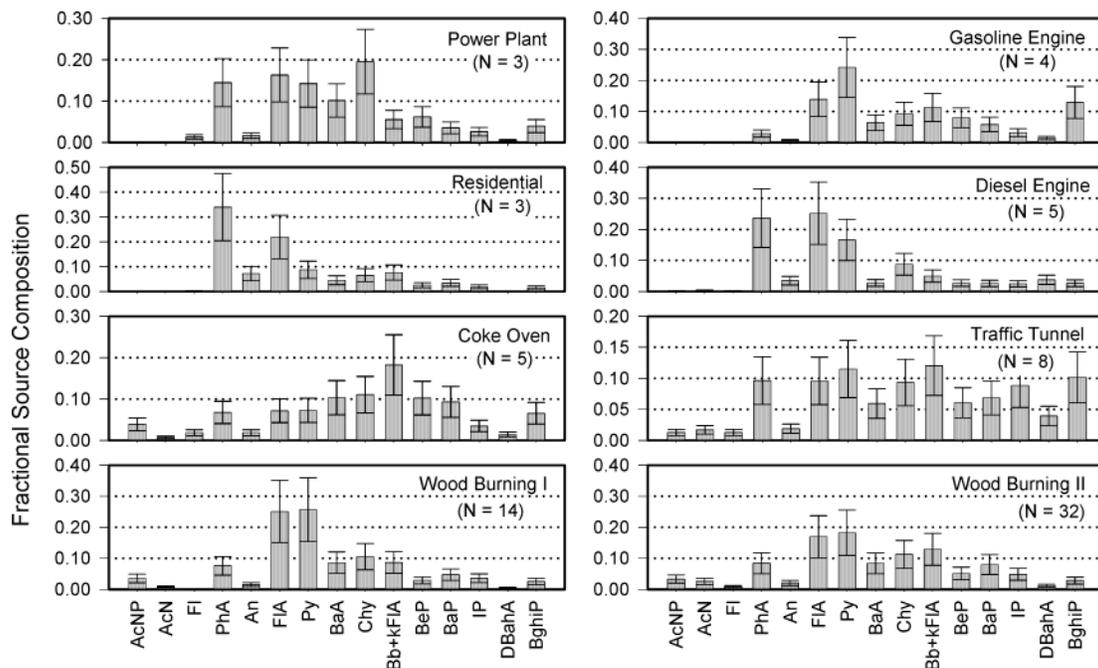


FIGURE 2. Source profiles. Wood Burning I contains only particulate phase literature sources. Wood Burning II contains both particulate- and gaseous-phase literature sources. Error bars represent a constant relative error of 40%.

samples (15). As a result of difficulties associated with GC separation of BbFIA and BkFIA the sum of the two isomers was used as the abbreviation Bb+kFIA (15). Also, samples with concentrations less than 2 $\mu\text{g/g}$ were not considered (15). Thus, the resulting data matrix was 15 \times 49. FA was applied for a two-, three-, four-, five-, and six-source solution; however, the FA model only yielded convergent solutions for the two- and six-source solutions.

In addition to the 15 \times 49 data set, a 15 \times 43 data set was modeled. This data set eliminated core K, where all six samples had significantly higher IP concentrations than the average IP concentration for the data set. Two- and five-source FA solutions were obtained; three-, four-, and six-source applications did not yield convergent solutions. PAH sources were determined, qualitatively, by visually comparing patterns between literature source profiles and factor loadings, and quantitatively, by the sum of squares difference between normalized literature PAH source profiles, and normalized FA model factor loadings.

Monte Carlo Simulation. To determine the uncertainty of the FA model results Monte Carlo simulation was utilized. The governing equation follows

$$D_{ij} = A_{ij} + C_{ij}A_{ij}[\sqrt{2}\text{erf}^{-1}(2R_{ij} - 1)] \quad (2)$$

where D_{ij} is the generated PAH concentration from PAH i and sample j , A_{ij} is the starting concentration of PAH i from sample j , C_{ij} is the coefficient of variation of PAH i from sample j , erf^{-1} is the inverse Gaussian error function and R_{ij} is a random number between 0 and 1.

Using Monte Carlo simulation nine artificial data matrices **D** were generated and modeled by FA to yield nine factor-loading and nine factor-score matrices. The standard deviation of the mean for each entry in the factor-loading and factor-score matrices was calculated to be the uncertainty.

A_{ij} was taken to be the modified 15 \times 49 data set. The coefficient of variation for each sample was obtained from the standard deviations of replicate experimental measurements (15). If the coefficient of variation for a given sample was less than 20% the value was used in the Monte Carlo simulation, otherwise 20% was used as the coefficient of variation value. If the coefficient of variations were much

larger than 20% the FA model would converge for less than 50% of the model runs, demonstrating the limits of the robustness of the FA model. Nonconvergent solutions were not included in the uncertainty analysis; new random data sets were generated until 9 convergent solutions were obtained.

Results and Discussion

FA Model Performance. Table 1 shows the results of diagnostic tools, including the coefficient of determination (COD), cumulative percent variance, and Exner function. Each element in Table 1 is determined by running the FA model without nonnegative rotations but with scaling and backscaling. Table 1 also defines PAH abbreviations.

The Exner function is below 0.10, and the cumulative percent variance is greater than 90% for all but the one factor solution, representing excellent correlation. COD values for the two-factor solution are generally greater than 0.90, but less than 0.95, and for the six-factor solution fluorene (Fl) is the only PAH with a COD less than 0.90. Notice that IP was not well determined until the sixth source was added. In general, the diagnostic tools indicate an excellent fit between the modeled and actual data.

The two-source factor-loading solution presented in Figure 3 displays two PAH source profiles. By visually comparing PAH patterns (Figure 2), and from the sum of squares of differences between modeled and literature PAH profiles (Table 2), loading 1 of 2 appears to be coke oven emissions, and loading 2 of 2 traffic tunnel airborne particles. A disagreement was observed for IP in the coke oven profile (loading 1 of 2), with the modeled values being much larger than the literature value. One explanation for the enriched IP values is the high IP concentrations measured in core K. The influence of IP on the FA model is seen in the six-source solution (Figure 4) as a separate source (loading 6 of 6). All other PAHs have a reasonable fit to the model. For this reason the value of IP was reduced to 30% of its modeled value during the sum of squares calculation (Table 2), so the poor fit of IP would not overshadow the excellent fit of the other PAHs. The traffic tunnel profile fits the literature profile (Figure 2, Table 2) well. Fluoranthene (FIA) and pyrene (Py)

TABLE 1. Results of Diagnostic Tools Application for the Determination of the Number of Significant Factors for Lake Calumet (IL)

PAHs	coefficient of determination factors					
	1	2	3	4	5	6
acenaphthylene (AcNP)	0.31	0.45	0.76	0.95	0.99	1.00
acenaphthene (AcN)	0.35	0.37	0.48	0.70	0.99	0.99
fluorene (FI)	0.30	0.61	0.62	0.66	0.78	0.79
phenanthrene (PhA)	0.79	0.83	0.83	0.90	0.93	0.95
anthracene (An)	0.65	0.92	0.93	0.93	0.93	0.94
fluoranthene (FIA)	0.77	0.87	0.88	0.96	0.97	0.97
pyrene (Py)	0.66	0.85	0.85	0.98	0.98	0.98
benzo[a]anthracene (BaA)	0.91	0.95	0.95	0.97	0.97	0.97
chrysene (Chy)	0.96	0.96	0.96	0.97	0.98	0.98
benzo[b]+[k]fluoranthene (Bb+kFIA)	0.93	0.93	0.94	0.94	0.95	0.96
benzo[e]pyrene (BeP)	0.92	0.93	0.96	0.96	0.97	0.98
benzo[a]pyrene (BaP)	0.79	0.82	0.90	0.93	0.96	0.97
indeno[123-cd]pyrene (IP)	0.63	0.86	0.87	0.87	0.89	1.00
dibenz[a,h]anthracene (DBahA)	0.31	0.76	0.95	0.95	0.99	1.00
benzo[ghi]perylene (BghiP)	0.79	0.88	0.94	0.95	0.96	0.98
cumulative %						
variance	88.4	93.3	95.7	97.3	98.6	99.2
Exner function	0.13	0.07	0.06	0.04	0.03	0.02

TABLE 2. Sum of Squares for All PAH Compounds^a

factor loadings (Figures 3 and 4)	literature PAH profiles (Figure 2)						
	power plant	coal residential	coke oven	wood burning I	gasoline engine	diesel engine	traffic tunnel
1 of 2 ^b	0.061	0.130	0.007	0.093	0.046	0.107	0.013
2 of 2	0.025	0.089	0.019	0.028	0.021	0.053	0.015
1 of 6	0.087	0.167	0.016	0.119	0.063	0.143	0.024
2 of 6 ^c	0.033	0.092	0.063	0.018	0.018	0.042	0.038
3 of 6	0.077	0.123	0.036	0.090	0.051	0.103	0.027
4 of 6	0.128	0.220	0.024	0.168	0.100	0.202	0.055
5 of 6 ^d	0.042	0.045	0.085	0.049	0.057	0.037	0.062

^a Bold faced type indicates probable PAH source profile as discussed in the text. ^b IP concentration reduced to 30% of the value modeled. ^c BghiP concentration changed to the amount present in the literature gasoline engine profile. ^d PhA and FIA were not considered in the analysis because of high uncertainties in the literature source profiles for wood burning.

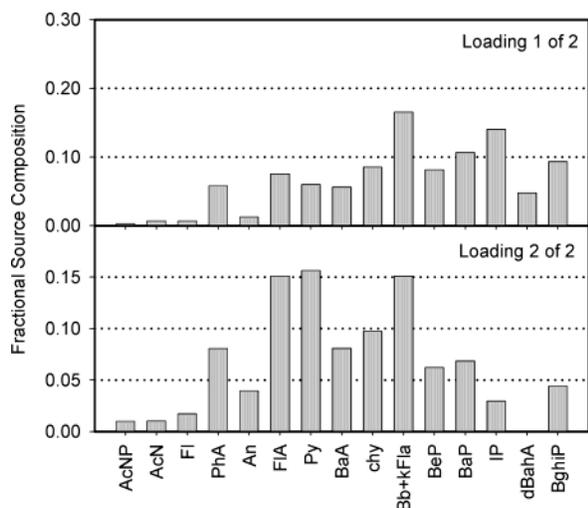


FIGURE 3. Factor loading plots for Lake Calumet two-source factor analysis solutions.

are overestimated, while benzo[ghi]perylene (BghiP) is underestimated. The percent contributions from the coke oven emissions and traffic tunnel were calculated to be 45% and 55%, respectively, based on the average of the four individual percent contributions from cores D, E, I, and K. This agrees well with the overall results of 52% from coal-related sources and 48% from traffic tunnel obtained from CMB modeling Operation #5 (15) and the results of 51% from coke oven and 42% from both gasoline and diesel engine

emissions by CMB Operation #2 (15). Operations #2 and #5 were particularly significant statistically.

Note that the source profiles derived from two-factor solution of FA (Figure 3) match coke oven and traffic tunnel sources, not the "Coal Average" and "Traffic Average" shown in Figure 2 of ref 15. In fact, these two averages show considerable similarities, which resulted in poor CMB modeling results (high percentage of source contribution estimates were "inestimable"). Averaging source profiles across categories diminishes the difference, resulting in unsuccessful model runs.

The agreements between the literature (Figure 2) and the model-derived source profiles and between the relative source contributions resulting from the CMB and FA models are remarkable. The CMB model relies on the availability and the adequacy of the source profiles of all major sources, from which the source contributions were computed using statistical techniques, such as, the effective variance weighted solution used in EPA's CMB8.2. In contrast, no *a priori* knowledge about the source emission is needed to run the FA model. Compared with previous applications of factor analysis for atmospheric apportionment (9, 24), our FA modeling with nonnegative constrains has the advantage of detailed comparison with literature source profiles and with results of CMB modeling.

The six-source factor-loading solution is presented in Figure 4. By comparing with the literature source profiles shown in Figure 2, and sum of squares calculations in Table 2, the six sources include two coke oven sources, a gasoline engine source, a traffic tunnel source, a wood burning/residential coal source, and a loading dominated by IP.

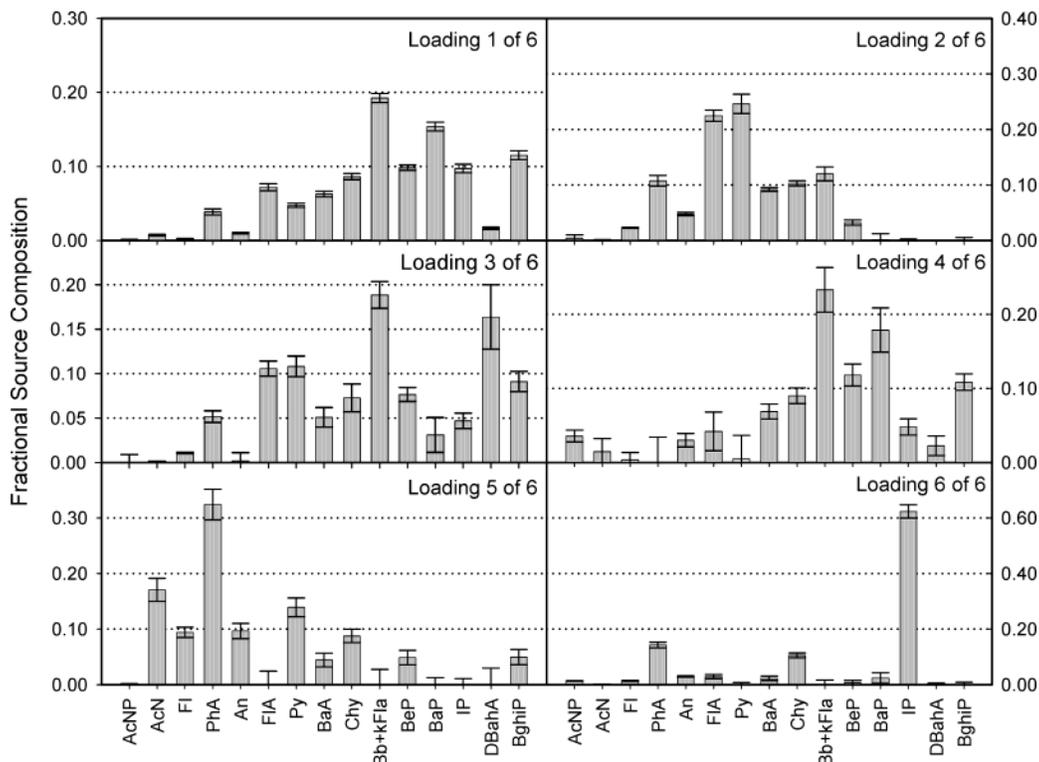


FIGURE 4. Factor loading plot for Lake Calumet six-source factor analysis solution with error bars representing the standard deviation of the mean for nine FA model runs using data sets created by Monte Carlo simulation.

Loading 1 of 6 is a coke oven profile. The model is not able to reproduce BaP and IP really well, i.e., the modeled profiles do not match the literature profiles. The model reproduces most other PAHs accurately, as supported by the sum of squares calculation (Table 2) for coke oven, (loading 1 of 6) which is the lowest.

Loading 2 of 6 represents a gasoline engine profile. A distinct pattern is observed between the literature and modeled profiles for the PAHs FIA, Py, BaA, Chy, Bb+kFla, BeP, and BaP. This pattern is not present in other literature source profiles, such as wood burning and power plant. The model has difficulties reproducing BaP and BghiP to the fractional source compositions from the literature. Uncertainty analysis does indicate BaP and BghiP are present in some simulations. For the sum of squares calculation the concentration of BghiP was increased to that of the gasoline engine literature profile to determine if the remaining PAHs were modeled well. Loading 3 of 6 is a traffic tunnel profile. Low molecular weight PAHs are reproduced fairly accurately by the model; however, some higher molecular weight PAHs such as BaP, IP, and dibenz[*a,h*]anthracene (dBahA) vary from the literature profiles. The variance of IP and dBahA could result from the poor separations between IP and dBahA (15), and also the uncertainty for dBahA is the largest observed. The uncertainty for BaP is also significant. Sum of squares calculations (Table 2) indicate traffic tunnel is the best fit to model data.

An interesting point from the above discussion is that there are two traffic sources generated by the model. Considering the primary traffic source to be I-94, located just west of Lake Calumet (Figure 1), one may wonder how this is possible. We believe the separate sources are a result of varying inputs from traffic sources. For example, I-94 may have primary gasoline engine traffic during daytime hours and more of a mixture of traffic tunnel sources during the nighttime. The split in traffic sources could produce two sources that are distinguished by the FA model.

Loading 4 of 6 is another coke oven profile. This profile is low in PhA and Py; however, it is enriched with BaP and

An. Uncertainty analysis does indicate PhA and Py are present and quite uncertain. This profile could result from a separate coke oven facility, possibly in Gary, Indiana. This would indicate the first coke oven source is from the coke plant located two miles north of site J. Another explanation for the second coke oven source is that it represents a degraded coke oven profile, with losses of PhA and Py observed. It is also possible that the same coke oven emits two different signals based on fuel source, installation of controls, and so forth, although the prevalence of the secondary profile (Loading 4 of 6) in core I and the primary profile (Loading 1 of 6) in core D (Figure 5) suggest two different coke oven sources.

Loading 5 of 6 has properties of a wood burning profile or a mix of a wood burning and residential coal profile. Most of the literature sources for wood burning do not measure acenaphthene (AcN); however, it is present in loading 5 of 6 and is much more abundant than indicated by the literature sources. Two outlier values of AcN in the data set, J-3 and I-5, may be influencing the FA model. PhA is very abundant in loading 5 of 6, possibly a result of combined residential coal and wood burning source or an artifact of the model. FIA is abundant in wood burning profiles but only seen in the uncertainty analysis for loading 5 of 6. The wood burning literature profiles are uncertain as seen with Wood Burning I and Wood Burning II (Figure 2), thus variance is expected between loading 5 of 6 and a literature wood burning profile. The sum of squares calculation (Table 2) for loading 5 of 6 eliminates PhA and FIA because of high uncertainties in the literature source profiles for wood burning. As a result of the small percent contribution (2.3%), from loading 5 of 6, it is possible that other PAH sources such as power plant are influencing the source profile. Source contributions from loading 5 of 6 will be discussed below and provide evidence for a wood burning or combined wood burning-residential coal source. Loading 6 of 6 is dominated by IP and believed to be a result of the high IP concentrations in core K.

To examine if high IP values in core K would account for a factor of six, a FA model run was made where core K was

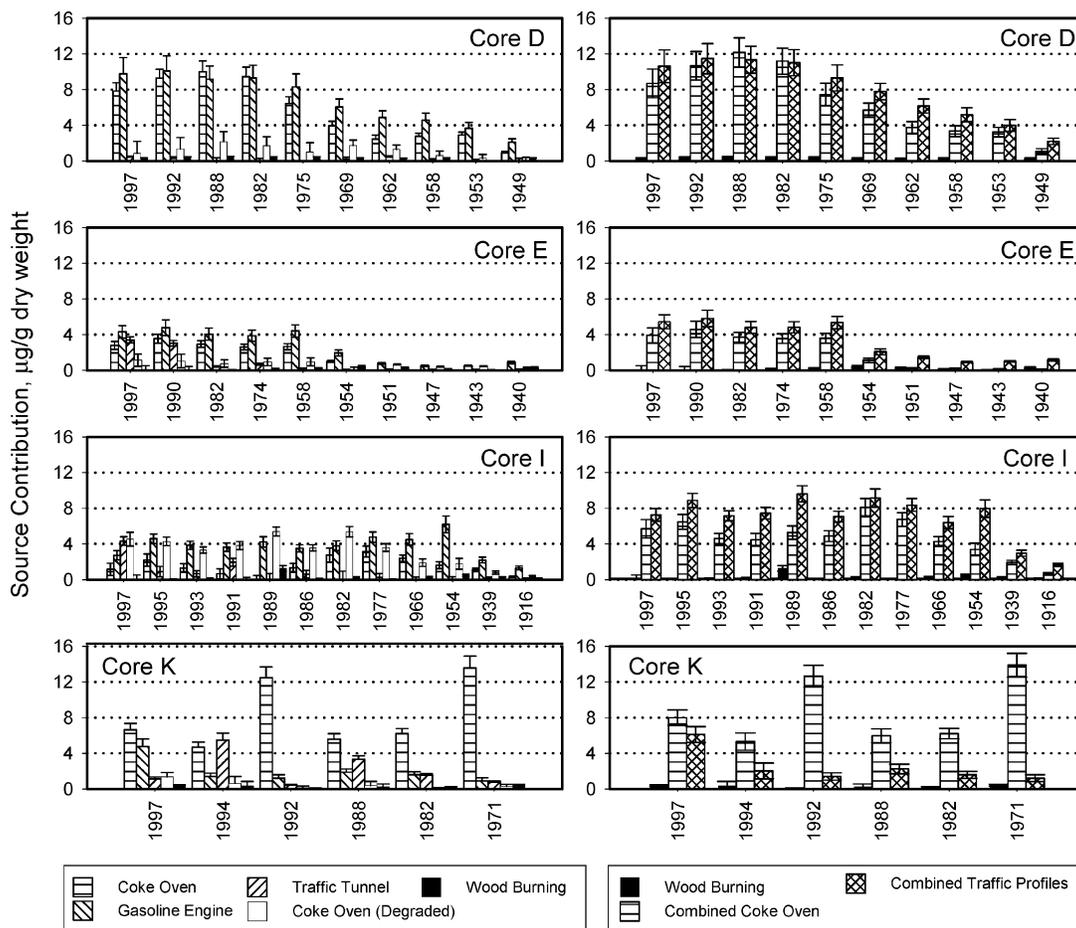


FIGURE 5. Left: Source contributions for five of six factors. Right: Source contributions for combination of coke oven and traffic sources, and wood burning—residential coal source. Error bars represent the standard deviation of the mean for nine data sets created by Monte Carlo simulation.

eliminated from the data set. Two- and five-factor solutions were convergent (three, four, and six were not convergent), and the resulting loadings were similar to the two- and six-factor solutions without loading 6 of 6. Thus, factor 6 represents high IP values in core K. Loading 1 of 2 (Figure 3) did see a reduction in IP also. Core K was left in the study to include as much information as possible.

Overall Source Contributions. Source contributions for the two-source solution (not shown) are in relative agreement with the six-source solution. Therefore, only the six-source source contributions are shown in Figure 5 and discussed below. The left column of Figure 5 shows the individual contributions from all five PAH sources, coke oven, gasoline engine, traffic tunnel, coke oven, and wood burning, with the sixth factor containing mostly IP omitted. The percent contribution from each source was calculated based on the average of the four individual percent contributions from cores D, E, I, and K. The percent contribution from each source is ~33%, 36%, 9.3%, 14%, and 2.3%, respectively, the IP factor contributes ~6.0%. The right column shows the wood burning and the sum of coke oven and traffic sources, with the sixth factor omitted. In general, all cores are dominated by the first coke oven source and a traffic source. This finding is in agreement with ref 15.

The individual source contribution profiles from the right columns of Figure 3, from ref 15 (Figure 6) and Figure 5 (current study) will now be compared and analyzed. Only cores (D, E, I, and K) containing more than two samples are plotted. Core D is comparable in both figures, with combined traffic having the larger contribution. A larger traffic contribution is expected due to core D's location near I-94. Traffic

contributions peak in the late 1980s and early 1990s. A continual increase in traffic contributions is seen before 1988. The wood burning—residential coal profiles were found to be much less significant by both studies, although Figure 6 shows a fluctuating profile, while Figure 5 displays more constant contributions that have a minimum in 1953.

Core E displays higher coke oven and traffic contributions from 1958 to 1997 in both Figures 5 and 6. After 1958 the contributions increase significantly. The increase in traffic contributions after 1958 could be a result of I-94 opening in 1962. The contributions from coke oven, traffic, and wood burning—coal residential are all in general agreement. No distinct wood burning minimum is observed.

For core I source contributions from coke oven and combined traffic vary between ref 15 (Figure 6) and Figure 5. For example, in Figure 6 coke oven is a larger contributor of PAHs before 1989, but traffic remains the primary PAH source in Figure 5. However, the contribution magnitudes, in $\mu\text{g/g}$, for both Figures are similar. In Figure 5 the traffic contribution is fluctuating, but a sharp increase is observed after 1954. The wood burning—coal residential profiles do have an interesting feature, for 1989 both models show a higher contribution, possibly attributed to a data set outlier I-5, which contains a high AcN concentration. The wood burning and coal residential profiles do show a minimum in the early 1980s and 1970s in agreement with U.S. wood consumption figures (18, 25, 26).

Core K is dominated by traffic tunnel in Figure 6, but combined coke oven dominates in Figure 5. Core K is ~2 miles south of a coke oven plant that was active during sampling, thus high coke oven contributions can be expected.

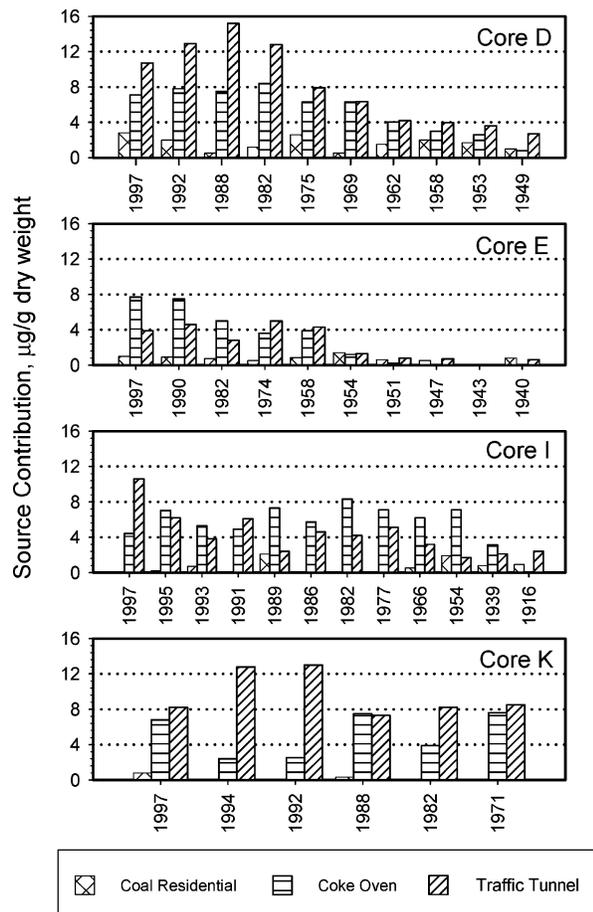


FIGURE 6. Source contributions from CMB model analysis, by Li et al. 2003 (15), for Operation #5.

With the exception of 1997, traffic profiles are fairly constant in Figure 5; however, a small peak in 1988 is observed. Wood burning and coal residential have small contributions in both models. However, in Figure 5 there appears to be a minimum between 1982 and 1992, indicating that a wood burning source may be more likely than a coal residential source (18). The variance between Figure 6 and Figure 5 may be attributed to several outlier IP values in core K, with Figure 5 probably being more accurate since IP here is represented by a separate factor.

Overall, combined coke oven (47%) and combined traffic (45%) have the largest PAH contribution to Lake Calumet. Traffic contributions increase significantly after the 1960s, and there seems to be a consistent maximum contribution in the late 1980s and early 1990s. Coke oven contributions are highest in core K, which is located near the coke oven facility. Wood burning is a minor PAH source. Loading 5 of 6 (Figure 5) is taken as the mix of residential coal combustions and wood burning rather than only residential coal combustion. The reason is that this source contribution is observed to continue after the 1950s, while coal was not used as domestic heating fuel much after the 1950s, and cores D and I have minimum wood burning values in the 1950s and 1960s. From the literature (18, 25, 26), a wood burning minimum in U.S. consumption is seen in the 1960s.

Compared with previous FA models for atmospheric PAHs, we show here that coke oven emissions is the most important coal-related PAH source, and we provide uncertainties of the estimated source profiles. In addition, time records of wood consumption are used to demonstrate that the low-molecular weight dominated PAH source profile is likely to be from wood burning rather than residential coal.

Acknowledgments

This work was supported by the U.S. National Science Foundation grant number BES-0107402. Student support is also provided by the Pilot Grant Program of the Illinois Education and Research Center, National Institute of Occupational Safety & Health (NIOSH) Award #: T42/CCT510424-09.

Supporting Information Available

Additional references, a summary of the wood type and phase measured (Table S1), and a summary of literature PAH source compositions normalized to benzo[*e*]pyrene (Table S2) for wood burning I and II. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Literature Cited

- Neff, J. M. *Polycyclic Aromatic Hydrocarbons in the Aquatic Environment: Sources, Fates, and Biological Effects*; Applied Science Publishers: London, 1979.
- Harrison, R. M.; Smith, D. J. T.; Luhana, L. *Environ. Sci. Technol.* **1996**, *30*, 825–832.
- Baek, S. O.; Field, R. A.; Goldstone, M. E.; Kirk, P. W.; Lester, J. N.; Perry, R. *Water, Air, Soil Pollut.* **1991**, *60*, 279–300.
- Christensen, E. R.; Li, A.; Ab Razak, I. A.; Rachdawong, P.; Karls, J. F. *J. Great Lakes Res.* **1997**, *23*(1), 61–73.
- Su, M. C. Ph.D. Dissertation, University of Wisconsin–Milwaukee, Milwaukee, WI, 1997.
- Hopke, P. K. *Receptor Modeling in Environmental Chemistry*; Wiley Science: New York, 1985.
- Ozeki, T.; Koide, K.; Kimoto, T. *Environ. Sci. Technol.* **1995**, *29*, 1638–1645.
- Sjogren, M.; Li, H.; Rannug, U.; Westerholm, R. *Environ. Sci. Technol.* **1996**, *30*, 38–49.
- Simcik, M. F.; Eisenreich, S. J.; Lioy, P. J. *Atmos. Environ.* **1999**, *33*, 5071–5079.
- Rachdawong, P.; Christensen, E. R.; Karls, J. F. *Water. Res.* **1998**, *32*(8), 2422–2430.
- Imamoglu, I. Ph.D. Dissertation, University of Wisconsin–Milwaukee, Milwaukee, WI, 2001.
- Ehrlich, R.; Wennig, R. J.; Johnson, G. W.; Su, S. H.; Paustenbach, D. J.; *Arch. Environ. Contam. Toxicol.* **1994**, *27*, 486–500.
- Kennicutt, M. C. II; Wade, T. L.; Presley, B. J.; Requejo, A. G.; Brooks, J. M.; Denoux, G. J. *Environ. Sci. Technol.* **1994**, *28*, 1–15.
- Yunker, M. B.; Snowdon, L. R.; MacDonald, R. W.; Smith, J. N.; Fowler, M. G.; Skibo, D. N.; McLaughlin, F. A.; Danyushevskaya, A. I.; Petrova, V. I.; Ivanov, G. I. *Environ. Sci. Technol.* **1996**, *30*, 1310–1320.
- Li, A.; Jang, J. K.; Scheff, P. A. *Environ. Sci. Technol.* **2003**, *37*(13), 2958–2965.
- Singh, A. K.; Gin, M. F.; Ni, F.; Christensen, E. R. *Water Sci. Technol.* **1993**, *28*, 91–102.
- Rachdawong, P.; Christensen, E. R. *Environ. Sci. Technol.* **1997**, *31*, 2686–2691.
- Rachdawong, P. Ph.D. Dissertation, University of Wisconsin–Milwaukee, Milwaukee, WI, 1997.
- Imamoglu, I.; Li, K.; Christensen, E. R. *Environ. Toxicol. Chem.* **2002**, *21*(11), 2283–2291.
- Johnson, G. W.; Jarman, W. M.; Bacon, C. E.; Davis, J. A.; Ehrlich, R.; Risebrough, R. W. *Environ. Sci. Technol.* **2000**, *34*, 552–559.
- Malinowski, E. R. *Factor Analysis in Chemistry*, 2nd ed.; Wiley-Interscience: New York, 1991.
- Imamoglu, I.; Christensen, E. R. *Water. Res.* **2002**, *36*, 3449–3462.
- Jang, J. K.; Li, A. *Chemosphere* **2001**, *44*(6), 1439–1445.
- Larson, R. K.; Baker, J. E. *Environ. Sci. Technol.* **2003**, *37*, 1873–1881.
- Energy Information Administration (EIA). *Estimation of wood consumption, 1949–1981*; Washington, DC, 1981.
- Energy Information Administration (EIA). *Estimation of wood consumption, 1979–1983*; Washington, DC, 1983.

Received for review July 29, 2003. Revised manuscript received September 29, 2003. Accepted October 8, 2003.

ES034842K