Receptor modeling for source apportionment of polycyclic aromatic hydrocarbons in urban atmosphere

Kunwar P. Singh · Amrita Malik · Ranjan Kumar · Puneet Saxena · Sarita Sinha

Received: 18 September 2006 / Accepted: 23 February 2007 / Published online: 30 March 2007 © Springer Science + Business Media B.V. 2007

Abstract This study reports source apportionment of polycyclic aromatic hydrocarbons (PAHs) in particulate depositions on vegetation foliages near highway in the urban environment of Lucknow city (India) using the principal components analysis/absolute principal components scores (PCA/APCS) receptor modeling approach. The multivariate method enables identification of major PAHs sources along with their quantitative contributions with respect to individual PAH. The PCA identified three major sources of PAHs viz. combustion, vehicular emissions, and diesel based activities. The PCA/APCS receptor modeling approach revealed that the combustion sources (natural gas, wood, coal/coke, biomass) contributed 19-97% of various PAHs, vehicular emissions 0-70%, diesel based sources 0-81% and

K. P. Singh (⊠) • A. Malik • R. Kumar
Environmental Chemistry Division,
Industrial Toxicology Research Centre,
Post Box 80, Mahatma Gandhi Marg,
Lucknow 226 001, India
e-mail: kpsingh_52@yahoo.com, kunwarpsingh@gmail.com

P. Saxena AAI – DU, Allahabad, India

S. Sinha National Botanical Research Institute, Lucknow, India other miscellaneous sources 0-20% of different PAHs. The contributions of major pyrolytic and petrogenic sources to the total PAHs were 56 and 42%, respectively. Further, the combustion related sources contribute major fraction of the carcinogenic PAHs in the study area. High correlation coefficient (R^2 >0.75 for most PAHs) between the measured and predicted concentrations of PAHs suggests for the applicability of the PCA/APCS receptor modeling approach for estimation of source contribution to the PAHs in particulates.

Keywords Source apportionment · Polycyclic aromatic hydrocarbons · Particulates · Vegetation foliage · PCA/APCS · Multiple linear regression · Receptor modeling

Introduction

Polycyclic aromatic hydrocarbons (PAHs), a well documented class of contaminants found in the environment, consist of carbon and hydrogen atoms arranged in two or more fused aromatic rings. Low aqueous solubility, semi-volatility, high octanol–water partition coefficient (K_{ow}), stability and their potential or proven carcinogenicity make them of considerable ecotoxicological concern. PAHs have been included in the group of persistent toxic substances (PTS) (UNEP 2003). PAHs originate from both the natural as

well as anthropogenic sources that include thermal combustion processes, vehicular emissions, and biomass burning (Simoneit 1984; McVeety and Hites 1988). Major human activities known to produce PAHs include pyrolysis of wood to produce charcoal and carbon black, coke production, manufacturing of gas fuel, power generation from fossil fuels, combustion of fuels in internal combustion engines, incineration of industrial and domestic wastes, oil refinery and chemical engineering processes, aluminum manufacturing etc. Airborne particulates, generated from these processes and carrying PAHs, are transported globally through the atmospheric pathways. Low molecular weight PAHs (MW<252) are predominantly present in the gaseous phase, while those with higher molecular weights remain in particle-bound fractions due to their low vapour pressures and high octanol-air partition coefficient (Koa) values (Jones et al. 1992; Kaupp 1996). The particle size of particulates and atmospheric conditions mainly govern the residence time of PAHs in the atmosphere (Chamberlin and Little 1981). For small particles (~1 µm diameter), it may be a few weeks, whereas, a few days for particles with 1-10 µm diameter (Suess 1976). Petroleum derived PAHs are preferentially biodegradable as compared to combustion derived ones (Jones et al. 1986). Combustion derived PAHs emitted to the atmosphere can return back to land, water and vegetation surfaces through dry deposition of airborne particulate matter, or wet deposition by precipitation, which may subsequently interact with the aquatic and terrestrial life including the human beings causing serious health hazards.

Studies on plants-atmospheric PAHs interactions (Bakker et al. 1999; Howsam et al. 2000a) revealed that both the gaseous phase as well as particulate bound PAHs in atmosphere may remain deposited on foliage surface, retained with cuticular wax and diffuse in to the interior of the leaves. Larsson and Sahlberg (1982) demonstrated that only PAHs with low molecular weight, which are present mainly in gaseous phase in atmosphere, could be sorbed into the cuticle, whereas, high molecular weight PAHs, mainly associated with particulates would remain on the leaf surface. The higher molecular weight PAHs viz. benzo(a)anthracene, benzo(b)fluoranthene, benzo(k) fluoranthene, benzo(a)pyrene, dibenzo(a,h)anthracene, and indeno(1,2,3-cd)pyrene have been classified as carcinogens (IARC 1987). Their sources in the region and their quantitative contributions are matter of concern.

For development of appropriate strategies for adequate and effective control measures of PAHs, it is desirable to identify sources and their quantitative contributions in a region. PAHs generally occur as complex mixtures and not as single compound in environmental samples. Differences in the pattern of PAH mixtures and thus, their concentration ratios reflect their different sources and pathways, which provide a basis for identifying their sources. However, the approach based on concentration ratio of specific pairs of PAHs, which involves only a few of these is often unable to resolve sources with sufficient accuracy and may be considered incomplete for source apportionment process.

Receptor modeling approach has often been used for source apportionment of atmospheric pollutants. Among these, principal components analysis/absolute principal components scores (PCA/APCS; Thurston and Spengler 1985), factor analysis (FA; Ozeki et al. 1995), positive matrix factorization (PMF; Paatero 1997) and UNMIX (Henry 1997) are the most widely employed multivariate models. These models analyze a series of observations simultaneously to determine the number of sources, their chemical composition and quantitative contributions to each observation (Miller et al. 2002). Unlike the chemical mass balance (CMB) approach (USEPA 2000), the multivariate models do not require prior knowledge of number and nature of sources.

Principal components analysis (PCA), an exploratory data analysis method allows for identification of major contamination sources (principal components) and estimation of their spatial and temporal distributions (Tauler et al. 2004), whereas APCS method is used to quantify the contributions of all sources to each measured pollutant (Thurston and Spengler 1985). Subsequent multiple linear regression (MLR) of sample mass concentration on these APCS derives estimated mass concentration of each source. The PCA/APCS receptor model requires no input data on source characteristics, while generates quantitative information on source profile and source strength in terms of absolute concentrations. This feature enables the model hypothesis to be validated by comparison of estimated source profiles with literature data. On the other hand, like other multivariate receptor models, the PCA/APCS model requires adequate degree of freedom for an accurate statistical analysis (Thurston and Spengler 1985; Guo et al. 2004). The PCA/APCS receptor model has been applied successfully for source apportionment of contaminants in atmosphere (Thurston and Spengler 1985; Harrison et al. 1996; Bruno et al. 2001; Miller et al. 2002; Guo et al. 2004) as well as in aquatic systems (Simeonov et al. 2003; Singh et al. 2005a). However, the multivariate models may not be able to separate strongly correlated sources. Moreover, the sources of PAHs are generally overlapping, and thus, emitting several of these moieties from common sources in more or less quantities. The basic purpose of this study is to identify the major sources of various PAHs in the study region.

This study basically aims to identify sources of PAHs in urban environment of Lucknow (India) and to quantify the contributions of all major sources to each measured PAH in particulate depositions on vegetation foliages in the study area by PCA/APCS approach based receptor modeling. The results of the study would help to fill the gap in the inventory of source profile of PAHs and understanding the chemical characteristics of emissions from multiple sources in urban atmosphere. This would also help the regulatory agencies in refining the pollution inventory, formulate pollution control strategies and implement air quality standards.

Materials and methods

Sampling

Samples of different vegetation foliages (about 250 g each) along with the particulate depositions were collected carefully from the selected sites (Fig. 1) within the Lucknow city (about 3 km along the highway) up to a distance of 50 m on both the sides of the national highway-23 (NH-23) in the first week of February 2005. It is worth mentioning here that during the monsoon season (July–September), all vegetation foliages are thoroughly washed due to regular heavy rains in the region. Fresh depositions of atmospheric particulates usually begin from the month of October. Sampling of plant leaves was conducted using the transect method described elsewhere (Howsam et al. 2000b). Samples of each

species were collected from different sites on both the sides of the highway. Ten species of vegetation selected were peepal (Ficus religiosa), neem (Azadirachta indica), mango (Mangifera indica), goldmohar (Delonix regia), subabool (Leucaena leucocephela), eucalyptus (Eucalyptus spp.), sheesham (Dalbergia sisso), harda (Terminalia chebula), emli (Tamarindus indica), and pakar (Ficus rumphii) and the sampling distances from the highway were approximately 1, 5, 10, 20, 30, and 40 m. Small twigs (three or four branches) of each species were cut from at each sampling site using a long-handled pruner. The leaves were removed by hand from the twigs, and care was taken to minimize contact with the leaf surface. Leaves were transferred to solvent-washed 500 ml glass jars, which were transported to the laboratory at low (4°C) temperature and stored in deep-freezer at -20°C until processed for analysis. All the samples were processed within a week time.

The study area (Fig. 1) is characterized by surrounding households, hotels and restaurants, diesel and petrol pumps, LPG good owns, coal/coke based activities, open waste incineration, conventional biomass fuels burning, and all types of vehicles (petrol, diesel and gasoline driven) plying on the highway through out the year.

Extraction and analysis

Prior to extraction, the samples were allowed to defrost for 4-6 h at 4°C. About 100 g samples (fresh) of leaves (exactly weighed) were taken and processed (Bakker et al. 2001) for estimation of 16 PAHs, viz. naphthalene (Nap), acenaphthylene (AcN), acenaphthene (AcNp), fluorene (Fl), phenanthrene (Phen), anthracene (An), fluoranthene (Fla), pyrene (Py), benzo(a)anthracene (BaA), chrysene (Chy), benzo(b) fluoranthene (BbF), benzo(k)fluoranthene (BkF), benzo(a)pyrene (BaP), indeno(1,2,3-cd)pyrene (IP), dibenzo(a,h)anthracene (DbA), and benzo(ghi)perylene (BP). The leaves were washed for 2 min with 200 ml of EDTA solution $(3 \times 10^{-2} \text{ M}; \text{ pH 5})$ in a beaker. The EDTA solution was extracted by refluxing with 100 ml cyclohexane for 15 min. Cyclohexane layer was taken out and concentrated to 1 ml under nitrogen. Then the EDTA washed leaves were immersed in 100 ml dichloromethane (DCM) for 20 s. DCM extract was concentrated to 1 ml under





nitrogen, then 50 ml methanolic KOH was added and refluxed for 30 min. One hundred milliliter cyclohexane was added and again refluxed for 10 min. Cylohexane layer was taken out and concentrated to 1 ml under nitrogen. Both the cyclohexane extracts were pooled and 5 ml methanol was added and again it was concentrated to 2 ml. Clean up was done on column filled with 5 gm florisil. After applying the sample, column was eluted with 10 ml ethanol and 5 ml pentane and discarded. Then the column was eluted with 45 ml acetonitrile and 5 ml ethanol, which was evaporated to dryness and finally made to 1 ml in DCM for GC–MS analysis. Purified extracts (surface washings and cuticular fractions) were analyzed by Gas Chromatography–Mass Spectrometry (Perkin Elmer Turbo Mass with Auto XL GC). The mass spectrometer was operated in total ion chromatography mode (TIC) with mass range 50 to 300 and ionization was done in electron impact (EI) mode at 70 eV. The separation of 16 PAHs was carried out on PE-5^{MS} low bleed capillary column 30 m×0.25 mm ID×0.25 µm film thickness with helium as a carrier gas at 1 ml/min flow. The temperature of injector, transfer line and ion source were 285°C, 290°C and 200°C, respectively. One microlitre of sample was injected with split less mode.

The oven temperature program started at 75°C, held for 5 min, and increased to 150°C at a rate of 25°C/min and then to 265°C at a rate of 4°C/min. The final temperature was held for 10 min and then increased to 285°C at a rate of 30°C/min. Quantitative recoveries of all the PAHs were determined and final concentration values are reported after applying the corresponding recovery factor.

Quality assurance/control

All samples were analyzed in duplicate. Recoveries of the PAHs were determined by spiking a known amount (100 ng/g) of PAHs to leaf samples of known background concentration and calculating the percentage recovered after clean up. Recoveries of the 16 PAHs varied between 78 and 96 ($\pm 8-12$)%, except for benzo(k)fluoranthene and benzo(a)pyrene, which was 68 ± 12 and $52\pm 14\%$, respectively.

The limits of detection (LOD) for PAHs were determined through repeated runs (n=6) of respective standard of lowest concentration. The LOD was based as three times the observed standard deviation of six replicate analyses of the lowest standard (Notar and Leskovsek 2000). For different PAHs, the LODs ranged from 2.2 to 8.5 pg. PAH concentration values only above their corresponding LOD are reported here. Samples (60 nos.) in which all the PAHs were detected have only been included in the study. The sum of all the 16 compounds constituted ' Σ PAH.'

Principal component analysis (PCA)

Here PCA was used as an exploratory tool to identify the major sources of atmospheric PAHs. PCA is a well established tool for analyzing structure in multivariate data sets (Miller et al. 2002). In general, PCA provides information on the most meaningful parameters, which describe whole data set rendering data reduction with minimum loss of original information (Wunderlin et al. 2001; Singh et al. 2005a,b,c,d). It is a technique that attempts to explain the variance of a large set of inter-correlated variables and transforming into a smaller set of independent (uncorrelated) variables (principal components). PCA assuming a linear model to explain the observed data variance using a reduced number of components decomposes the **X** matrix of n_{obj} =rows (samples) and n_{var} =columns (variables), according to (Tauler et al. 2004):

$$x_{ij} = \sum_{f=1}^{F} a_{if} . b_{fj} + e_{ij}$$
(1)

where, a_{if} and b_{fj} are the elements of the scores and loadings matrices **A** and **B** of $(n_{obj} \times F)$ and $(n_{var} \times F)$ dimensions, respectively and e_{ij} is the error term of the element x_{ij} of the **X** data array and *F* is the number of principal components (PCs) extracted. Equation 1 in matrix form can also be written as;

$$\mathbf{X} = \mathbf{A}\mathbf{B}^{\mathrm{T}} + \mathbf{E} \tag{2}$$

where \mathbf{X} is the matrix of measurements, \mathbf{A} is the matrix of scores, \mathbf{B} is the matrix of loadings and \mathbf{E} is the error matrix containing the variance not explained by the model defined by *F* components described in \mathbf{A} and \mathbf{B} .

In Eq. 1, x_{ij} is the measured concentration of *j*th PAH in *i*th sample in data matrix **X**. The a_{if} (score of component f on row i) is the contribution of contamination source f in sample i. The b_{fi} (loading of *j*th variable on *f*th component) is the contribution of *j*th PAH in *f*th contamination source, and e_{ij} is the residual in sample *i*, variable *j* of x_{ij} not modeled by the F environmental contamination sources. This equation suggests that the measured concentrations are a weight (scores, a_{if}) sum of a reduced number (f) of main environmental contributions defined by a particular chemical composition (loadings, b_{fi}), apart from noise (multiple small unknown contributions) and experimental error defined by e_{ii} . The weights or scores a_{if} , describe the distribution of main contamination sources among the samples and the loadings b_{fi} , identify the chemical composition of these contamination sources. It may be noted that a factor does not necessarily represent a specific emission or depletion mechanism, but rather a pattern of association (Derwent et al. 1995). Thus, a component may contain more than one emission sources.

PCA was performed on experimental data set (60 samples \times 16 variables) standardized through z-scale transformation in order to remove differences in variables ranges as well as to give variables the same importance in the analysis (Simeonov et al. 2003; Singh et al. 2004). Thus, the bilinear model Eq. 1 now for z-transformed variables yields normalized PC scores. Further, the PCs were subjected to varianx

rotation generating varifactors (VFs). The varimax rotation of the matrix attempts to minimize the number of variables (PAHs) that have high loadings on a factor, thus enhances the interpretability of the factors (Buhamra et al. 1998). The number of significant components was chosen by the brokenstick criterion. The broken-stick model assumes that total variance is proportioned among the components and that the expected eigenvalue distribution should follow a broken-stick distribution. Observed eigenvalues are considered interpretable when they exceed the expected eigenvalues generated by the brokenstick model (Jackson 1993; King and Jackson 1999). Finally from the PCA performed on standardized (ztransformed) variables, we obtained the factor scores, A $(I \times F)$, loadings, B $(J \times F)$, and factor scores coefficients, **S** ($J \times F$) matrices, where I is the number of samples, J is the number of variables and F is the number of factors extracted. The factor scores matrix (A) can be obtained by multiplication of the factor scores coefficients matrix (S) and the standardized variables matrix (Z) as A=ZS (Lawley and Maxwell 1971). All the mathematical and statistical computations were made using Excel 97 and Statistica 7.0 software.

Absolute principal component scores (APCS)

The absolute principal component scores (APCS) were used to estimate the source contributions to each pollutant. Since, we performed PCA on standardized (z-transformed) variables; the yielded normalized factor scores cannot be used directly for computation of quantitative source contributions. The normalized factor scores determined subsequently in Eq. 1 were converted to un-normalized APCS. Details on the computation of APCS have been described elsewhere (Thurston and Spengler 1985; Swietlicki et al. 1996; Guo et al. 2004; Singh et al. 2005a). In brief, the preliminary steps involve standardization of all variables (PAHs) concentration as Z_{ij} ;

$$Z_{ij} = \frac{\left(x_{ij} - \overline{x}_j\right)}{\sigma_j} \tag{3}$$

where x_{ij} is the measured concentration of variable (PAHs) *j* in sample *i*. \overline{x} is the arithmetic mean concentration of variable *j*, and σ_j is the standard

deviation of variable j for all samples included in this analysis.

Now, the PCA performed on these standardized variables yielded normalized factor scores (A_z) with zero mean and unit standard deviation, an artificial sample with concentration equal to zero for all the variables (PAHs) was introduced to compute absolute zero scores for each factor (Thurston and Spengler 1985; Swietlicki et al. 1996; Guo et al. 2004; Singh et al. 2005a), so that

$$(Z_0)_j = \frac{\left(0 - \overline{x}_j\right)}{\sigma_j} = -\frac{\overline{x}_j}{\sigma_j} \tag{4}$$

Now, the absolute zero factor scores (A_0) for each sample were computed from the values of corresponding factor scores coefficients (S) obtained from PCA performed on standardized variables and the values of (Z_0) computed by Eq. 4, as

$$(A_0)_f = \sum_{j=1}^J S_{fj} \cdot (Z_0)_j \tag{5}$$

The absolute principal component scores (APCS) for each sample in each component is then estimated by subtracting the absolute zero factor scores values (A_0) of each sample from the corresponding normalized factor scores values (A_z) obtained by PCA of the standardized variables as (Thurston and Spengler 1985);

$$APCS_{f} = (A_{z})_{if} - (A_{0})_{f}$$

$$f = 1, 2, \dots, F$$
(6)

Receptor modeling

Receptor modeling approach is based on the assumption that the total concentration of each contaminant is made up of the linear sum of its contributions from each of the contamination source components collected at the receptor site, as (Thurston and Spengler 1985; Swietlicki et al. 1996; Guo et al. 2004; Singh et al. 2005a);

$$x_{ij} = \sum_{k=1}^{F} w_{jk} \cdot p_{ik} + e_{ij}$$
(7)

where, x_{ij} is the measured concentration of contaminant (variable) j in sample i. For a number of F relevant sources, w_{jk} denotes the source profile for the source k. It specifies the amounts of each of the contaminants emitted by source k. p_{ik} are the source contributions of source k to sample i and e_{ij} is the error term. From the PCA (Eq. 1) and receptor (Eq. 7) models, it is evident that the terms w_{jk} and p_{jk} in later Eq. 7 are equivalent to the factor loadings (b_{jj}) and factor scores (a_{ij}) terms obtained from Eq. 1.

Receptor modeling approach based on multiple linear regression of the absolute principal component score (APCS–MLR) is a widely employed statistical technique for source apportionment of environmental contaminants (Fung and Wong 1995; Swietlicki and Krejei 1996; Simeonov et al. 2003; Guo et al. 2004; Singh et al. 2005a). Multiple linear regression (MLR) of the measured concentration of the contaminants (PAHs) as dependent variables with the APCS as the independent variables yields estimates of the coefficients, which convert the APCS into contaminant sources mass contributions from each source for each sample. Based on statistical significance of the regression coefficients different sources (APCS) of contamination are identified. Finally, the measured PAHs concentration data as dependent variables were regressed on mass concentrations of different sources as independent variables, yielding the source contribution to C_j as (Thurston and Spengler 1985);

$$C_j = (r_0)_j + \sum_{k=1}^F r_{kj} * \operatorname{APCS}_k$$
(8)

where, $(r_0)_j$ is the constant term of multiple regression for pollutant *j*, r_{kj} is the coefficient of multiple regression of the source *k* for pollutant *j*, and APCS_k is the scaled value of the rotated factor *k* for the considered sample. The combined term, r_{kj} *APCS_k represents the contribution of source *k* to C_j . Moreover, the mean of the product r_{kj} *APCS_k on all samples represents the average contribution of the sources (*p*). Quantitative contributions from each source for individual contaminant (PAH) were compared with their measured values.

 Table 1
 Some general properties of various PAHs (ATSDR 1995)

РАН	Chemical structure	Chemical formula	Molecular weight	Vapor pressure (Pa)	Aqueous solubility (mg/L)	Log K _{ow}	IARC probable/possible carcinogenic PAHs
Naphthalene	\bigcirc	C ₁₂ H ₁₂	128.12	_	_	_	
Acenaphthylene		$C_{12}H_8$	152.20	3.87	3.93	4.07	
Fluorene	\sim	$C_{13}H_{10}$	166.2	4.27×10^{-2}	1.68-1.98	4.18	
Acenaphthene		$C_{12}H_{10}$	154.21	5.96×10^{-1}	1.93	3.98	
Phenanthrene	00	$C_{14}H_{10}$	178.2	9.07×10^{-2}	1.20	4.45	
Anthracene	()	C14H10	178.2	2.27×10^{-3}	0.076	4.45	
Fluoranthene		$C_{16}H_{10}$	202.26	6.67×10^{-4}	0.20-0.26	4.90	
Pyrene		$C_{16}H_{10}$	202.3	3.33×10 ⁻⁴	0.077	4.88	
Benzo(a)	∞	$C_{18}H_{12}$	228.29	2.93×10 ⁻⁶	0.01	5.61	+
Chrysene	$\widehat{\mathbf{m}}$	C ₁₈ H ₁₂	228.3	8.40×10^{-5}	2.8×10^{-3}	5.16	
Benzo(k)	άŭ	$C_{20}H_{12}$	252.3	8.79×10^{-9}	7.6×10^{-4}	6.06	+
fluoranthene Benzo(b) fluoranthene		$C_{20}H_{12}$	252.3	6.67×10 ⁻⁵	0.0012	6.04	+
Benzo(a)pyrene	<u> </u>	C ₂₀ H ₁₂	252.3	7.47×10^{-7}	2.3×10^{-3}	6.06	+
Dibenzo(a,h) anthracene	Sag	$C_{22}H_{14}$	278.35	1.33×10 ⁻⁸	5.0×10 ⁻⁴	6.84	+
Indeno(1,2,3- cd)pyrene	600	$C_{22}H_{12}$	276.3	1.33×10^{-9} -1.33×10 ⁻⁴	0.062	6.58	+
Benzo(ghi) perylene	88	$C_{22}H_{12}$	276.34	1.37×10 ⁻⁸	2.6×10 ⁻⁴	6.50	

PAH	*LOD (pg)	Concentration (ng/g)					
		Range	Median	1st Quartile	3rd Quartile		
Naphthalene	8.52	0.11-70.65	15.28	5.34	23.38		
Acenaphthylene	2.24	0.05-44.35	8.66	1.99	15.08		
Fluorene	2.52	0.11-25.93	9.64	1.97	15.24		
Acenaphthene	3.81	0.09-27.90	5.24	1.25	13.51		
Phenanthrene	2.54	0.06-17.00	3.41	0.60	7.93		
Anthracene	3.16	0.12-13.64	2.76	0.35	6.64		
Fluoranthene	6.22	0.05-9.08	0.95	0.06	2.88		
Pyrene	3.34	0.06-5.90	0.06	0.06	0.72		
Benzo(a)anthracene	4.21	0.09-5.07	0.25	0.09	1.04		
Chrysene	3.33	0.07-2.09	0.08	0.07	0.51		
Benzo(k)fluoranthene	6.51	0.05-1.34	0.06	0.06	0.22		
Benzo(b)fluoranthene	2.87	0.06-1.92	0.06	0.06	0.40		
Benzo(a)pyrene	4.15	0.05-7.60	0.86	0.06	2.17		
Dibenzo(a,h)anthracene	2.38	0.05-1.92	0.05	0.05	0.24		
Indeno(1,2,3-cd)pyrene	2.62	0.03-1.82	0.04	0.04	0.04		
Benzo(ghi)perylene	3.24	0.05-1.68	0.06	0.06	0.46		
\sum PAHs		0.42-204.17	54.23	11.02	87.02		

Table 2 Univariate statistics of different PAHs concentration in particulate depositions in Lucknow city, India (N=60)

*LOD Limit of detection equals three times the standard deviation of the replicate analysis of the lowest standard (n=6).

Results and discussion

Some general properties of various PAHs studied here and univariate statistics pertaining to their measured concentrations in particulate depositions on vegetation foliage are presented in Tables 1 and 2, respectively. It may be noted that low molecular weight (2–3 ring) PAHs dominated over the others. Naphthalene was observed with highest concentration among all the 16 PAHs. The composition of PAHs depends on the combustion temperature. At low to moderate temperature, as in the wood stove or as from



Fig. 2 Contribution of different groups (ring-wise) of PAHs to the total PAH burden in particulates

the combustion of coal, low molecular weight PAH compounds are abundant, whereas, at higher temperature, such as in the vehicle emissions, the higher molecular weight PAH compounds are dominant (Zhang et al. 2004). Figure 2 shows the relative contribution (mean) of different groups (2-3 ring, 4 ring, 5 ring and 6 ring) of PAHs to the total PAHs concentration. The 2-3 ring compounds, on an average, contributed 89% to the total PAHs burden. This suggests that low temperature pyrolysis (combustion) is the main source for the origin of PAHs in the study area. Moreover, the low molecular weight (2-3 ring) PAHs are dominantly present in gaseous phase in the atmosphere and during interactions with the vegetation foliages, considerable fraction of these may enter in to the cuticular wax layer on surface of the leaves. Thus, resulting in to building up of their higher concentrations (Larsson and Sahlberg 1982; Bakker et al. 1999; Howsam et al. 2000a) as observed in present study. Moreover, superficially adhered particulates (rich in higher PAHs) on the surface of leaves are washed out during heavy rains of monsoon season, and fresh particulate depositions start again from successive months, while, those (lower PAHs) diffused to the cuticular wax layers (hydrophobic) reflect cumulative levels over time.

Table 3 Varimax rotation factor loadings on various variables (PAHs)

Variable (PAH)	VF1	VF2	VF3	
Naphthalene	0.737	0.403	0.214	
Acenaphthylene	0.904	0.264	0.193	
Fluorene	0.680	0.527	0.335	
Acenaphthene	0.817	0.475	0.201	
Phenanthrene	0.612	0.553	0.402	
Anthracene	0.700	0.446	0.453	
Fluoranthene	0.656	0.392	0.500	
Pyrene	0.210	0.792	0.349	
Benzo(a)anthracene	0.839	0.228	0.062	
Chrysene	0.197	0.796	0.033	
Benzo(k)fluoranthene	0.256	0.905	-0.029	
Benzo(b)fluoranthene	0.635	0.372	0.205	
Benzo(a)pyrene	0.639	0.537	0.268	
Dibenzo(a,h)anthracene	0.435	0.708	0.124	
Indeno(1,2,3-cd)pyrene	0.084	0.056	0.935	
Benzo(ghi)perylene	0.860	0.047	-0.131	
Eigenvalue	6.39	4.46	2.00	
%Variance explained	39.92	27.87	12.51	
Source	Combustion	Vehicular	Diesel	

Bold faced values are strong loadings (>0.70).

Values in italics are moderate loadings (>0.50).

PCA and source identification

PCA performed on z-transformed data yielded three significant PCs explaining 80.3% of the data variance. The three PCs captured 64.14, 9.02 and 7.13% of variance, respectively. The varimax rotated factor loadings of normalized PAHs concentrations in deposited particulates on vegetation foliages in the study area are presented in Table 3. Varimax rotation of the PCs separated the variables (PAHs) in to identifiable source categories (factors).

Factor 1 Explaining 39.92% of the total variance, VF1 has strong (>0.70) positive loadings on naphthalene, acenaphthylene, acenaphthene, anthracene, benzo(a)anthracene, and benzo(ghi)perylene and moderate (>0.50) positive loadings on fluorene, phenanthrene, fluoranthene, benzo(b)fluoranthene, and benzo(a)pyrene. Most of these PAHs are predominant in wood and coal combustion signals (Freeman and Cattel 1990; Khalili et al. 1995; Harrison et al. 1996; Mastral et al. 1996). Fluorene is reported as dominant PAH in coke oven signature (Khalili et al. 1995). Fluorene and benzo(a)anthracene have been taken as tracer for coke oven combustion and natural

gas combustion, respectively (Simcik et al. 1999). Therefore, this factor can be selected to represent natural gas and wood/coal/coke combustion sources. The study area is known for several activities related with natural gas and wood/coal/coke combustion, such as cooking, heating during winters and small metal based workshops. Combustion of conventional fuels (wood, coal, coke, grass, biomass, cow dung cake), open waste burning/incineration etc. are common in the study region. Bhargava et al. (2004) reported that among conventional fuels (wood, cow dung cake) used in northern India, highest amounts of carcinogenic PAHs are emitted by cow dung cake burning followed by wood. In semi urban and rural areas of India, this is the most prevalent type of fuel for cooking and other heating purposes (Bhargava et al. 2004).

Factor 2 Explaining about 27.87% of variance, VF2 showed strong positive loadings on pyrene, chrysene, benzo(k)fluoranthene and dibenzo(a,h)anthracene, while moderate positive loadings on fluorene, phenanthrene and benzo(a)pyrene. All the PAHs having strong positive loadings in this factor are mainly associated with vehicular emissions (Khalili et al.





1995; Larsen and Baker 2003). The study area is in close vicinity of a national highway and a large number of both light and heavy vehicles operate throughout leading to high vehicular emissions.

Factor 3 Explaining about 12.51% of variance, VF3 showed strong positive loadings on indeno(1,2,3-cd) pyrene and moderate positive loadings on fluoranthene. Diesel and gasoline vehicle emissions have been characterized by their relatively higher concentrations (Li and Kamens 1993). Indeno(1,2,3-cd) pyrene is a tracer for diesel combustion sources (Simcik et al. 1999). Since, the number of gasoline based vehicles in the region is negligibly small; this factor has been taken as representing the diesel sources.

It may be noted that due to the overlap of the source signatures, it is difficult to differentiate between the finer sources of PAHs. However, in general, the first factor (VF1) represents the pyrolytic sources, whereas the second (VF2) and third (VF3) factors can be considered representing petrogenic sources of PAHs.

193

Table 4	Contribution	(%) (of various	sources	of PAH	s in	particulate	deposition	on vegetation	foliage
---------	--------------	-------	------------	---------	--------	------	-------------	------------	---------------	---------

S. No	РАН	Source contr	Ratio Predicted Measured	R^2			
		Misc. (S0)	Combustion (S1)	Vehicular (S2)	Diesel (S3)		
1	Naphthalene	15.63	54.60	18.12	11.65	1.06	0.75
2	Acenaphthylene	4.07	76.11	13.50	6.33	1.00	0.93
3	Fluorene	19.69	48.30	22.71	9.30	1.00	0.85
4	Acenaphthene	3.81	66.38	23.43	6.37	1.00	0.94
5	Phenanthrene	0.00	55.40	30.38	14.22	0.97	0.84
6	Anthracene	0.00	61.01	23.58	15.41	0.99	0.90
7	Fluoranthene	0.00	60.22	21.86	17.92	1.17	0.84
8	Pyrene	0.00	25.40	58.13	16.47	1.27	0.80
9	Benzo(a)anthracene	0.00	85.87	14.13	0.00	1.19	0.76
10	Chrysene	0.00	28.48	69.69	1.83	0.82	0.67
11	Benzo(k)fluoranthene	0.00	31.81	68.19	0.00	0.93	0.89
12	Benzo(b)fluoranthene	0.00	73.77	26.23	0.00	0.98	0.76
13	Benzo(a)pyrene	0.00	59.75	30.48	9.77	1.12	0.77
14	Dibenzo(a,h)anthracene	0.00	50.30	49.70	0.00	1.22	0.71
15	Indeno(1,2,3-cd)pyrene	0.00	18.72	0.00	81.28	0.85	0.89
16	Benzo(ghi)perylene	0.00	96.80	3.20	0.00	1.07	0.76
-	Mean	2.70	55.81	29.58	11.91	1.04	0.82

Further, the loadings and scores plots of first two factors are presented in Fig. 3. The loadings plot (Fig. 3a) shows groupings and relationship between the variables. It may be noted that majority of the PAHs (high loadings in first factor) are located along the VF1 axis of the plot. In the scores plot (Fig. 3b) samples belonging to the sites closer to the highway are showing relatively higher scores, thus higher contamination. The vegetation closer to the highway is relatively more accessible to receive emissions from various sources.

Source apportionment

The APCS receptor model (APCS–MLR) was applied on the three extracted factors to quantify the contributions of the sources (identified through PCA) to each measured PAH. The results are presented in Table 4. Although, in APCS, source contribution estimates are not constrained to be non-negative (Miller et al. 2002; Harrison et al. 1996), the factors chosen in this study have very low negative values. Further, the contributions of non-specified miscellaneous sources were negligibly small, except for a few cases.

As evident from the correlation coefficients (R^2) , the multiple regression exhibited good adequacy between the measured and predicted (modeled) values. The correlation coefficient (R^2) values represented the fraction of variance of measured concentrations attributable to variance in the predicted concentrations. For most of the PAHs, R^2 values were greater than 0.75 indicating a good fit between the measured and predicted concentrations. Further, the ratio (close to 1 for most PAHs) of mean predicted and measured concentration values of almost all the individual PAH suggest goodness of the receptor modeling approach (PCA/APCS) to the source apportionment of PAHs (Table 4).

From Table 4, it may be noted that the combustion processes (natural gas, wood, coal/coke, biomass) in the study area contributed between 19 and 97% of various PAHs. Contribution of vehicular emission sources in terms of different PAHs was between 0 and 70%. The third source in terms of relative contributions of PAHs was identified as diesel combustion related activities. It contributed various PAHs between 0 and 81%. Apart from these, unidentified miscellaneous sources contributed various PAHs in particulates between 0 and 20%. The mean contributions of total PAHs from pyrolytic (natural gas, wood/coal/coke, conventional fuels) and petrogenic (diesel, petroleum, gasoline) sources were about 56 and 42%, respectively (Fig. 4a).



Fig. 4 Source contributions (%) for PAHs in particulate depositions on vegetation foliages (a) \sum PAHs (b) two- and three-ring PAHs, (c) four-ring PAHs, (d) five-ring PAHs, (e) six-ring PAHs, and (f) total carcinogenic PAHs

Moreover, the contributions of various identified sources in terms of different ring-PAHs were also estimated (Fig. 4b–e). The lower PAHs (two and three-ring members) were estimated to be contributed most from combustion activities (60%), vehicular emissions (22%), diesel combustion (11%), and from miscellaneous sources (7%). The four-ring PAHs were contributed most by natural gas, wood, coal/ coke combustion (50%) and vehicular emissions (41%). The five-ring PAHs were mainly contributed by the combustion activities (54%) and vehicular emissions (44%). Contributions of the higher members (6-ring) PAHs from combustion processes and diesel based activities were 58 and 41%, respectively.

The higher (four, five and six-rings) PAHs viz. benzo(a)anthracene, benzo(b) fluoranthene, benzo(k) fluoranthene, benzo(a)pyrene, dibenzo(a,h)anthracene, and indeno(1,2,3-cd)pyrene are known carcinogens (IARC 1987) and their sources in the region and their quantitative contributions are matter of concern. Mean contributions of various sources as sum of the carcinogenic PAHs in particulate depositions in the study area are shown in Fig. 4f. It may be noted that combustion related activities contributed about 53%, vehicular about 32% and diesel related activities 15% of the total carcinogenic PAHs in the particulates. This suggests that combustion is the major activity for formation and emission of carcinogenic PAHs in the study area. Coal/coke, wood, grass, cow dung cake and other biomass based fuels are extensively used at various household/community level for cooking/heating purposes in the region (Bhargava et al. 2004) and thus, the local population may be at risk due to exposure to the PAHs.

Conclusions

A multivariate (PCA/APCS) receptor model was applied to the data set on PAHs measured in particulates deposited on vegetation foliages in the near vicinity of national highway within the urban limits of Lucknow city (India) for apportionment of sources. PCA yielded three significant components identifying combustion (natural gas, wood, coal/ coke), vehicular emissions and diesel based processes as major sources of PAHs. The contributions of pyrolytic (natural gas, wood/coal/coke, conventional fuels) and petrogenic (diesel, petroleum, gasoline) sources were about 56 and 42%, respectively. The ratio of predicted and measured concentration of different PAH between 0.82 and 1.27 (close to 1.0 in most cases) shows adequacy of the represented sources in the model. An excellent agreement between the measured and predicted concentrations $(R^2 > 0.75$ for most PAHs) of different PAHs suggests for the adequacy of the applicability of the PCA/APCS receptor modeling approach for estimation of source contribution to the PAHs in particulates. The modeling approach, thus, offers a method of source apportionment and estimation of source contributions, which would be useful in refining the emission inventories of PAHs and developing appropriate strategies towards emission control.

Acknowledgements The authors thank the Director, ITRC, Lucknow for his keen interest in this work and encouragement. Financial assistance (SRF) provided by CSIR, New Delhi to one of the authors (AM) is thankfully acknowledged. We are also grateful to Prof. V. Simeonov (Faculty of Chemistry, University of Sofia, Bulgaria) for valuable suggestions and help in data analysis.

References

- ATSDR (1995). Toxicological profile for polycyclic aromatic hydrocarbons (PAHs). Agency for Toxic Substances and Disease Registry (ATSDR). Atlanta, GA: US Department of Health and Human Services, Public Health Service.
- Bakker, M., Koerselman, J., Tolls, J., & Kolloffel, C. (2001). Localization of deposited polycyclic aromatic hydrocarbons in leaves of Plantago. *Environmental Toxicology and Chemistry*, 20, 1112–1116.
- Bakker, M., Vorenhout, M., & Kolloffel, C. (1999). Dry deposition of atmospheric polycyclic aromatic hydrocarbons in three Plantago species. *Environmental Toxicol*ogy and Chemistry, 18, 2289–2294.
- Bhargava, A., Khanna, R. N., Bhargava, S. K., & Kumar, S. (2004). Exposure risk to carcinogenic PAHs in indoor-air during biomass combustion whilst cooking in rural India. *Atmospheric Environment*, 38, 4761–4767.
- Bruno, P., Caselli, M., de Gennaro, G., & Traini, A. (2001). Source apportionment of gaseous atmospheric pollutants by means of an absolute principal component scores (APCS) receptor model. *Fresenius Journal of Analytical Chemistry*, 371, 1119–1123.
- Buhamra, S. S., Bouhamra, W. S., & Elkilani, A. S. (1998). Assessment of air quality in ninety-nine residences of Kuwait. *Environmental Technology*, 19, 357–367.
- Chamberlin, A. C., & Little, P. (1981). Plants and their atmospheric environment. In J. Grace, et al. (Eds.), (pp. 147–173). Oxford: Blackwell Scientific.
- Derwent, R. G., Middleton, D. R., Field, R. A., Goldstone, M. E., Lester, J. N., & Perry, R. (1995). Analysis and interpretation of air quality data from an urban roadside location in Central London over the period from July 1991 to July 1992. *Atmospheric Environment*, 29, 923–946.
- Freeman, D. J., & Cattel, C. R. (1990). Wood burning as a source of atmospheric polycyclic aromatic hydrocarbons. *Environmental Science and Technology*, 24, 1581–1585.
- Fung, Y. S., & Wong, L. W. Y. (1995). Apportionment of air pollution sources by receptor models in Hong Kong. *Atmospheric Environment*, 29, 2041–2048.
- Guo, H., Wang, T., & Louie, P. K. K. (2004). Source apportionment of ambient non-methane hydrocarbons in Hong Kong: Application of a principal component analysis/absolute principal component scores (PCA/ APCS) receptor model. *Environmental Pollution*, 129, 489–498.

- Harrison, R. M., Smith, D. J. T., & Luhana, L. (1996). Source apportionment of atmospheric polycyclic aromatic hydrocarbons collected from an urban location in Birmingham, U.K. *Environmental Science and Technology*, 30, 825–832.
- Henry, R. C. (1997). History and fundamentals of multivariate air quality receptor models. *Chemometrics and Intelligent Laboratory Systems*, 37, 37–42.
- Howsam, M., Jones, K. C., & Ineson, P. (2000a). PAHs associated with the leaves of three deciduous tree species. I. Concentrations and profiles. *Environmental Pollution*, 108, 413–424.
- Howsam, M., Jones, K. C., & Ineson, P. (2000b). PAHs in the soils of a mature, mixed-deciduous (Quercus–Fraxinus) woodland and the surrounding pasture. *Water, Air and Soil Pollution*, 121, 379–398.
- IARC (1987). Overall evaluation of carcinogenicity: An updating of IARC monographs. IARC monographs on the evaluation of carcinogenic risk to humans. (Vols. 1–42, Suppl. 7) Lyon, France: International Agency for Research on Cancer.
- Jackson, D. A. (1993). Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches. *Ecology*, 74, 2204–2214.
- Jones, D. M., Rowland, S. J., Douglas, A. G., & Howells, S. (1986). An examination of the Fate of Nigerian crude oil in surface sediments of the Humber estuary by gas chromatography and gas chromatography–mass spectrometry. *International Journal of Environmental and Analytical Chemistry*, 24, 227–247.
- Jones, K. C., Sanders, G., Wild, S. R., Burnett, V., & Johnston, A. E. (1992). Evidence for a decline of PCBs and PAHs in rural vegetation and air in the United Kingdom. *Nature*, 365, 137–140.
- Kaupp, H. (1996). Atmospharische eintragswege und verhalten von polychlorierten di-benzo-p-dioxinen und-furanen sowie polyzyklischen Aromaten in einem Maisbestand, Ph. D. Thesis, University of Bayreuth, Bayreuth.
- Khalili, N. R., Scheff, P. A., & Holsen, T. M. (1995). PAH source fingerprints for coke ovens, diesel and gasoline engines, highway tunnels, and wood combustion emissions. *Atmospheric Environment*, 29, 533–542.
- King, J. R., & Jackson, D. A. (1999). Variable selection in large environmental data sets using principal components analysis. *Environmetrics*, 10, 67–77.
- Larsen, R. K., & Baker, J. E. (2003). Source apportionment of polycyclic aromatic hydrocarbons in the urban atmosphere: A comparison of three methods. *Environmental Science and Technology*, 37, 1873–1881.
- Larsson, B., & Sahlberg, G. (1982). Polynuclear aromatic hydrocarbons: Physical and biological chemistry. In M. Cooke, et al. (Eds.), (pp. 417–426). New York: Springer.
- Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworths.
- Li, C. K., & Kamens, R. M. P. (1993). The use of polycyclic aromatic hydrocarbons as source signatures in receptor modeling. *Atmospheric Environment*, 27A, 523–532.
- Mastral, A. M., Callen, M., & Murillo, R. (1996). Assessment of PAH emissions as a function of coal combustion variables. *Fuel*, 75, 1533–1536.

- McVeety, B. D., & Hites, R. A. (1988). Atmospheric deposition of polycyclic aromatic hydroocarbons to water surfaces: A mass balance approach. *Atmospheric Environment*, 22, 511–536.
- Miller, S. L., Anderson, M. J., Daly, E. P., & Milford, J. B. (2002). Source apportionment of exposures to volatile organic compounds. I. Evaluation of receptor models using simulated exposure data. *Atmospheric Environment*, 36, 3629–3641.
- Notar, M., & Leskovsek, H. (2000). Determination of polycyclic aromatic hydrocarbons in marine sediments using a new ASE–SFE extraction technique. *Fresenius Journal of Analytical Chemistry*, 366, 846–850.
- Ozeki, T., Koide, K., & Kimoto, T. (1995). Evaluation of sources of acidity in rainwater using a constrained oblique rotational factor analysis. *Environmental Science and Technology*, 29, 1638–1645.
- Paatero, P. (1997). Least squares formulation of robust, nonnegative factor analysis. *Chemometrics and Intelligent Laboratory Systems*, 37, 23–35.
- Simcik, M. F., Eisenreich, S. J., & Lioy, P. J. (1999). Source apportionment and source/sink relationships of PAHs in the coastal atmosphere of Chicago and Lake Michigan. *Atmospheric Environment*, 33, 5071–5079.
- Simeonov, V., Stratis, J. A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M., & Kouimtzis, Th. (2003). Assessment of the surface water quality in Northern Greece. *Water Research*, 37, 4119–4124.
- Simoneit, B. R. T. (1984). Organic matter of the troposphere. III: Characterization and sources of petroleum and pyrogenic residues in aerosols over the western United States. *Atmospheric Environment*, 18, 51–67.
- Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variation in water quality of Gomti River (India) – A case study. *Water Research*, 38, 3980–3992.
- Singh, K. P., Malik, A., & Singh, V. K. (2005b). Chemometric analysis of hydro-chemical data of an alluvial river – A case study. *Water, Air and Soil Pollution*, 170, 383–404.
- Singh, K. P., Malik, A., Singh, V. K., Mohan, D., & Sinha, S. (2005d). Chemometric analysis of groundwater quality data of alluvial aquifer of Gangetic plain, North India. *Analytica Chimica Acta*, 550, 82–91.

- Singh, K. P., Malik, A., & Sinha, S. (2005a). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques – A case study. *Analytica Chimica Acta*, 538, 355–374.
- Singh, K. P., Malik, A., Sinha, S., Singh, V. K., & Murthy, R. C. (2005c). Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. *Water, Air and Soil Pollution, 166*, 321–341.
- Suess, M. J. (1976). The environmental load and cycle of polycyclic aromatic hydrocarbons. *Science of the Total Environment*, 6, 239–250.
- Swietlicki, E., & Krejei, R. (1996). Source characterization of the Central European atmosphere aerosol using multivariate statistical methods. *Nuclear Instrumental Methods and Physics. B*, 109/110, 519–525.
- Swietlicki, E., Puri, S., Hansson, H. C., & Edner, H. (1996). Urban air pollution source apportionment using a combination of aerosol and gas monitoring techniques. *Atmospheric Environment*, 30, 2795–2890.
- Tauler, R., Lacorte, S., Guillamon, M., Cespeded, R., Viana, P., & Barcelo, D. (2004). Chemometric modeling of main contamination sources in surface waters of Portugal. *Environmental Toxicology and Chemistry*, 23, 565–575.
- Thurston, G. D., & Spengler, J. D. (1985). A quantitative assessment of source contributions to inhalable particulate matter pollution in Metropolitan Boston. *Atmospheric Environment*, 19, 9–25.
- UNEP (2003). Global report on regionally based assessment of persistent toxic substances. Geneva, Switzerland: PNEP Chemicals.
- USEPA (2000). CMB 8.2 User's manual, Office of Air Quality Planning and Standards: Research Triangle Park, NC.
- Wunderlin, D. A., Diaz, M. P., Ame, M. V., Pesce, S. F., Hued, A. C., & Bistoni, M. A. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia River basin (Cordova – Argentina). *Water Research*, 35, 2881–2894.
- Zhang, Z. L., Hong, H. S., Zhou, J. L., & Yu, G. (2004). Phase association of polycyclic hydrocarbons in the Minjiang River Estuary, China. *Science of the Total Environment*, 323, 71–86.