

GOVERNMENT OF INDIA  
ATOMIC ENERGY COMMISSION

**SOURCE APPORTIONMENT OF TOXIC CHEMICAL POLLUTANTS  
AT TROMBAY REGION**

by

S.K. Sahu, G.G. Pandit and V.D. Puranik  
Environmental Assessment Division

BHABHA ATOMIC RESEARCH CENTRE  
MUMBAI, INDIA  
2007

**BIBLIOGRAPHIC DESCRIPTION SHEET FOR TECHNICAL REPORT**  
(as per IS : 9400 - 1980)

01	<i>Security classification :</i>	Unclassified
02	<i>Distribution :</i>	External
03	<i>Report status :</i>	New
04	<i>Series :</i>	BARC External
05	<i>Report type :</i>	Technical Report
06	<i>Report No. :</i>	BARC/2007/E/007
07	<i>Part No. or Volume No. :</i>	
08	<i>Contract No. :</i>	
10	<i>Title and subtitle :</i>	Source apportionment of toxic chemical pollutants at Trombay Region
11	<i>Collation :</i>	52 p., 3 figs., 3 tabs.
13	<i>Project No. :</i>	
20	<i>Personal author(s) :</i>	S.K. Sahu; G.G. Pandit; V.D. Puranik
21	<i>Affiliation of author(s) :</i>	Environmental Assessment Division, Bhabha Atomic Research Centre, Mumbai
22	<i>Corporate author(s) :</i>	Bhabha Atomic Research Centre, Mumbai-400 085
23	<i>Originating unit :</i>	Environmental Assessment Division, BARC, Mumbai
24	<i>Sponsor(s) Name :</i>	Department of Atomic Energy
	<i>Type :</i>	Government

Contd...

30	<i>Date of submission :</i>	April 2007
31	<i>Publication/Issue date :</i>	May 2007
40	<i>Publisher/Distributor :</i>	Associate Director, Knowledge Management Group and Head, Scientific Information Resource Division, Bhabha Atomic Research Centre, Mumbai
42	<i>Form of distribution :</i>	Hard copy
50	<i>Language of text :</i>	English
51	<i>Language of summary :</i>	English, Hindi
52	<i>No. of references :</i>	49 refs.
53	<i>Gives data on :</i>	
60	<i>Abstract :</i>	Anthropogenic activities like industrial production and transportation, a wide range of chemical pollutants such as trace and toxic metals, pesticides, polycyclic aromatic hydrocarbons etc. eventually find their way into various environmental compartments. One of the main issues of environmental pollution is the chemical composition of aerosols and their sources. In spite of all the efforts a considerable part of the atmospheric aerosol mass is still not accounted for. This report describes some of the activities of Environmental Assessment Division which are having direct relevance to the public health and regulatory bodies. Extensive studies were carried out in our laboratories for the Trombay site, over the years; on the organic as well as inorganic pollution in the environment to understand inter compartmental behaviour of these chemical pollutants. In this report an attempt has been made to collect different size fractionated ambient aerosols and to quantify the percentage contribution of each size fraction to the total aerosol mass. Subsequently, an effort has been made for chemical characterization (inorganic, organic and carbon content) of these particulate matter using different analytical techniques. The comprehensive data set on chemical characterization of particulate matter thus generated is being used with receptor modeling techniques to identify the possible sources contributing to the observed concentrations of the measured pollutants. The use of this comprehensive data set in receptor modeling has been helpful in distinguishing the source types in a better way. Receptor modeling techniques are powerful tools that can be used to locate sources of pollutants to the atmosphere. The major advantage of the receptor models is that actual ambient data are used to apportion source contributions, negating the need for dispersion calculations. Pollution sources affecting the sampling site were statistically identified using varimax rotated factor analysis of the concentration data. Factor analysis, a form of receptor modeling technique was used for source apportionment of different fractions of atmospheric aerosols. This study will provide information feedback on the actual effects of anthropogenic activities on the local environment.
70	<i>Keywords/Descriptors :</i>	NUCLEAR FACILITIES; AIR POLLUTION MONITORING; CHEMICAL EFFLUENTS; ECOSYSTEMS; ENVIRONMENT; ENVIRONMENTAL EFFECTS; PUBLIC HEALTH; TOXIC MATERIALS; TOXICITY; PESTICIDES; AEROSOLS; MULTI-ELEMENT ANALYSIS
71	<i>INIS Subject Category :</i>	S54
99	<i>Supplementary elements :</i>	

## सारांश

मानवीय गतिविधियों, औद्योगिक उत्पादन एवं वाहन गतिविधियों के फलस्वरूप विभिन्न प्रकार के प्रदुषक जैसे कि हानिकारक धातु, कीटनाशक, पोलिसायक्लिक अरोमैटिक हायड्रोकार्बन आदि उत्पन्न होते हैं, जो कि पर्यावरण के अलग अलग भागों (वायुमंडल, जलियस्त्रोत, मिट्टी आदि) में प्रवेश कर जाते हैं। वातावरण में पाए जाने वाले सुक्ष्म कण और उनकी संरचना प्रदूषित वातावरण का एक विचारणीय विषय है। इस प्रतिवेदन में पर्यावरण मूल्यांकन प्रभाग द्वारा की गयी कुछ गतिविधियों, जिनका मानव स्वास्थ्य और नियामक सकार्यों से सीधा संबंध है, उल्लेख किया गया है। हमारी ट्राम्बे स्थित प्रयोगशालाओं में कई वर्षों से कार्बनिक एवं अकार्बनिक प्रदूषकों का व्यवहार एवं प्रभाव जानने के लिए विश्लेषण एवं मूल्यांकन किया गया है। विभिन्न माप के कणों को एकत्र करके, विश्लेषण एवं मूल्यांकन द्वारा उनका कुल कणीय भार में योगदान जानने की चेष्टा की गयी है। साथ ही विभिन्न तकनीकों द्वारा कार्बनिक पदार्थों (अकार्बनिक, कार्बनिक) अंश का भी मूल्यांकन किया गया है। Receptor modeling तकनीक प्रदुषण के स्रोत का पता लगाने के लिए एक उपयुक्त तरीका है। Receptor modeling तकनीक में कणों का स्रोत एवं प्रत्येक स्रोत का योगदान ज्ञात करने के लिए विश्लेषण द्वारा प्राप्त आंकड़ों का उपयोग किया जाता है। नमूने एकत्र करने के स्थान पर प्रदूषित कणों का स्रोत Varimax Related Factor विश्लेषण द्वारा ज्ञात किया गया। विभिन्न कणों का स्रोत जानने के लिए खंड विश्लेषण (Factor analysis) का प्रयोग किया गया। इस प्रकार प्राप्त आंकड़ों से मानवीय गतिविधियों द्वारा वातावरण पर होने वाले प्रभावों का आंकलन किया जा सकेगा।

**Source Apportionment of Toxic Chemical Pollutants at  
Trombay Region**

**S.K. Sahu, G.G. Pandit and V.D. Puranik**

**Environmental Assessment Division  
Bhabha Atomic Research Centre  
Mumbai-400 085**

# **Contents**

## **Chapter-1: Background**

### **1.1 Introduction**

### **1.2 Brief History**

### **1.3 General Principle**

## **Chapter-2: Overview of Receptor Oriented techniques**

### **2.1 Introduction**

### **2.2 Chemical Mass Balance**

### **2.3 Multivariate Analysis**

#### **2.3.1 Factor Analysis**

#### **2.3.2 Target Transformation Factor Analysis**

#### **2.3.3 Positive matrix Factorization**

### **2.4 Back Trajectories Methods**

#### **2.4.1 Residence Time Analysis**

#### **2.4.2 Quantitative Transport Bias Analysis**

#### **2.4.3 Potential Source Contribution Function Analysis**

#### **2.4.4 Residence Time Weighted Concentrations Analysis**

### **2.5 Comparison of Different Techniques**

## **Chapter-3: Works done on Source Apportionment of Air Pollutants at Trombay**

### **3.1 Collection of particulate matter samples**

#### **3.1.1 Results on Particulate Matter Levels**

### **3.2 Chemical Characterization:**

**3.2.1 Trace and Toxic Elements in Ambient Air**

**3.2.2 Organic Pollutants in Ambient Air**

**3.2.3 Carbon Analysis**

**3.3 Application of Receptor Modeling Techniques**

**3.3.1 Source Apportionment of Polycyclic Aromatic Hydrocarbons**

**3.3.2 Factor Analysis and Source Identification of Trace Elements  
in SPM**

**3.3.3 Factor Analysis and Source Identification of Trace Elements  
in PM<sub>2.5</sub>**

**3.3.4 Source Apportionment of RSPM**

**3.3.5 Source Apportionment of Fine Suspended Particulate Matter**

**Chapter-4: References**

## **List of Tables:**

**Table-1: Summary of the Receptor Modeling Techniques**

**Table 3.1 Concentrations of different size fractionated particulate matter and their ratios**

**Table-3.2: Factor Loading Matrix After Varimax Rotation**

**Table 3.3: Factor Loading Matrix of SPM after Varimax Rotation**

**Table 3.4: Factor Loading Matrix of PM<sub>2.5</sub> after Varimax Rotation**

**Table 3.5: Varimax rotated Factor analysis of PM<sub>10</sub> data**

**Table 3.6: Varimax rotated Factor analysis of PM<sub>2.5</sub> data**

## **List of Figures:**

**Fig. 3.1 Correlation between the PM<sub>10</sub> and PM<sub>2.5</sub> mass concentrations**

**Fig. 3.2 Correlation between the PM<sub>10</sub> and TSPM mass concentrations**

**Fig 3.3. Percentage Contribution of Different Sources to Total PAHs**

**Fig 3.4 Percentage Contribution from Different Sources to PM<sub>10</sub> Mass**

**Fig 3.5 Percentage Contribution from Different Sources to PM<sub>2.5</sub> Mass**

## Summary

Anthropogenic activities like industrial production and transportation, a wide range of chemical pollutants such as trace and toxic metals, pesticides, polycyclic aromatic hydrocarbons etc. eventually find their way into various environmental compartments. One of the main issues of environmental pollution is the chemical composition of aerosols and their sources. In spite of all the efforts a considerable part of the atmospheric aerosol mass is still not accounted for. This report describes some of the activities of Environmental Assessment Division which are having direct relevance to the public health and regulatory bodies. Extensive studies were carried out in our laboratories for the Trombay site, over the years; on the organic as well as inorganic pollution in the environment to understand inter compartmental behaviour of these chemical pollutants. In this report an attempt has been made to collect different size fractionated ambient aerosols and to quantify the percentage contribution of each size fraction to the total aerosol mass. Subsequently, an effort has been made for chemical characterization (inorganic, organic and carbon content) of these particulate matter using different analytical techniques. The comprehensive data set on chemical characterization of particulate matter thus generated is being used with receptor modeling techniques to identify the possible sources contributing to the observed concentrations of the measured pollutants. The use of this comprehensive data set in receptor modeling has been helpful in distinguishing the source types in a better way. Receptor modeling techniques are powerful tools that can be used to locate sources of pollutants to the atmosphere. The major advantage of the receptor models is that actual ambient data are used to apportion source contributions, negating the need for dispersion calculations. Pollution sources affecting the sampling site were statistically identified using varimax rotated factor analysis of the concentration data. Factor analysis, a form of receptor modeling technique was used for source apportionment of different fractions of atmospheric aerosols. This study will provide information feedback on the actual effects of anthropogenic activities on the local environment.

# Chapter-1

## Background

### 1.1 Introduction:

The atmosphere is a very complex system. Significant improvements have been made over past few decades in mathematical modeling of pollutant dispersion in the atmosphere. However, there are still many instances when these models are insufficient in providing the information required. Also, the performance of these models is greatly limited by the accuracy of input data on source emission characteristics and meteorological parameters. Hence alternative methods were developed to assist in the identification of sources, and apportionment of the observed pollutant concentrations to those sources. These methods are called receptor-oriented or receptor models since they focus on the behavior of the ambient environment at the point of impact. This is opposed to source-oriented dispersion models that focus on transport, dilution, and transformations of pollutants occurring at the source and follow the pollutants to the sampling or receptor site.

### 1.2 Brief History:

Several independent factors lead to the development of receptor-oriented models in the early 1970s. Most important of course was the interest in the environmental problems in general, and air quality in particular that started with the book "Silent Spring" by Rachel Carson (Henry,1999). In the United States the environmental movement of the 1960s led to environmental regulations, most notably the Clean Air Act of 1969 and its Amendment of 1977. The regulations spawned by these laws spurred the launching of monitoring networks and many special air quality studies which, for the first time, started producing large amounts of air quality data. Statistical analysis of this data and application of receptor oriented techniques was made possible by the introduction of relatively low-priced minicomputers in the 1970s. The paper by Bilfford and Meeker (1967) is the first report in the literature to present multivariate analysis of air quality data. The idea behind

the paper was to characterize air quality in general rather than to attempt a quantitative source apportionment. However, Peterson (1970) had the goal of predicting concentrations of sulfur dioxide in St. Louis by multivariate analysis of the spatial variability; an idea that was not pursued further until the 1990s (Henry et al., 1991a).

A major breakthrough came independently in the work of Winchester and Nifong (1971) and Miller, Friedlander and Hidy (1972) in which, they used trace elemental concentrations with conservation of mass. The result was the Chemical Mass Balance (CMB) model which was quickly applied to source apportionment of airborne particles in many studies (J. Bogen, 1973; R. Heindryckx et al., 1974; S.K. Friedlander, 1973; G. Gartrell et al., 1975; D.F. Gatz, 1975; G.S. Kowalczyk et al., 1978). CMB was applied earlier to volatile organic gas composition data by Mayrshon and Crabtree (1976). The original CMB model uses only simple multiple linear regression to fit mass balance equations to data from one sample at a time. The first application of a receptor model using factor analysis was done by Hopke et al. (1976). Henry and Hidy (1979) found physically meaningful air quality patterns in the first multivariate analysis of an air quality data set that combined pollutant gas concentrations, particulate composition, and meteorological variables.

Earlier trajectory models were used for meteorology, to investigate air mass flow around mountains (Steinacker, 1984) and for climatology, to identify the pathways of water vapor transport (D'Abreton et al., 1996) or desert dust (Chiapello et al., 1997). Stohl (1996) first applied the trajectory model to the environmental science to establish source-receptor relationships of air pollutants. Thus, over a period of few decades the combination of interest driven by regulations with advances in instrumental analytical analysis and the wide availability of computers resulted in the new field of air quality receptor modeling.

### **1.3 General Principle:**

The fundamental principle of receptor modeling is that mass conservation and mass balance analysis can be used to identify and apportion sources of contaminants in the atmosphere. The approach to obtaining a data set for receptor modeling is to determine a large number of chemical constituents in a number of samples. The mass balance

equation can then be extended to account for all  $m$  elements in the  $n$  samples as contributions from  $p$  independent sources

$$x_{ij} = \sum_{k=1}^p a_{ik} f_{kj} \quad i = 1, m; j = 1, n \quad \Lambda \Lambda \quad (1.1)$$

where  $x_{ij}$  is the  $i$ th elemental concentration measured in the  $j$ th sample,  $a_{ik}$  is the gravimetric concentration of the  $i$ th element in material from the  $k$ th source, and  $f_{kj}$  is the airborne mass concentration of material from the  $k$ th source contributing to the  $j$ th sample.

Equation (1.1) can be solved in a variety of ways depending on the amount of information that is initially known about the system. In all cases, the ambient concentrations of the species of interest have to be measured for a series of samples. Although there are some methods that can resolve the sources for a single sample, one sample does not characterize the air quality in the area.

There are a set of natural physical constraints on the system that must be considered in developing any model for identifying and apportioning the sources of air pollutants (Henry, 1991b), such as:

1. The original data must be reproduced by the model; the model must explain the observations.
2. The predicted source compositions must be non-negative; a source cannot have a negative percentage of an element.
3. The predicted source contributions must all be non-negative; a source cannot emit negative mass.
4. The sum of the predicted elemental mass contributions for each source must be less than or equal to the total measured mass for each element; the whole is greater than or equal to the sum of its parts.

## Chapter-2

### Overview of Receptor Oriented techniques

#### 2.1 Introduction:

Receptor Oriented models use the pollutant compositions measured at a specific monitoring site to apportion the contribution of individual sources. The application of receptor models usually requires source characterization measurements. While some receptor oriented techniques, such as multivariate methods that divide pollutant contributions into generic group, do not require explicit source measurements, a priori knowledge of emission characteristics is needed to assign sources to a specific category. The following assumptions generally apply to this type of model:

1. Source emissions have a constant composition;
2. Emission components are conservative not removed preferentially;
3. Emission components do not undergo chemical conversion.

It is critical that the number and type of constituents measured and analytical procedures used be consistent, whether they are within the samples at a single receptor site, among the samples at different locations or within the source sample measurements.

The major advantage of the receptor models is that actual ambient data are used to apportion source contributions, negating the need for dispersion calculations. Receptor algorithms only require relative emission composition information, not quantitative emission factors. Multivariate methods offer an additional advantage because fugitive emission sources may be identified through the application of the model.

Trajectory models, which describe the paths air parcel take, have been used to establish source-receptor relationships of pollutants for several years now. Earlier trajectories were only used to investigate the transport processes associated with individual air pollution events, but now sophisticated methods based on large sets of trajectories exist, to study also the air pollution climatology of a site. The trajectory based receptor models are useful to locate the geographical positions of the sources in a given area.

Unfortunately, receptor oriented models have several disadvantages. Because results cannot be confirmed directly, they are difficult to validate. Often it is necessary to

compare the findings between receptor and source models in order to check the validity of calculated values. Receptor oriented algorithms are likely to produce erroneous results whenever any of the basic assumptions are violated.

## 2.2 Chemical Mass Balance (CMB):

The basic assumption of this approach is that particles do not undergo any physical or chemical transformation during transport from the release source to receptor. The observed concentration pattern would be a linear combination of the patterns of pollutants from the sources each weighted by a source strength term. If the number of elements measured equaled or exceeded the number of source types, one could obtain the source strength term ( $M_j$ ) by solving a set of simultaneous equations given by:

$$C_i = \sum_j M_j X_{ij} \text{-----}(2.1)$$

where  $C_i$  is the concentration of pollutant  $i$  in the sample and  $X_{ij}$  is the concentration of the pollutant  $i$  from the  $j$  source. It has been assumed that all sources with a potential for significantly contributing to the receptor have been identified and have had their emissions characterized. Underestimating the number of sources had little effect on the calculated source contributions if the prominent species contributed by the missing sources were excluded from the solution. When the prominent species of the omitted sources were included in the calculation of source contributions, the contributions of sources with properties in common with the omitted sources were overestimated. When the source-types actually present were excluded from the solution, the sum of the source contributions was much less than the total measured mass. The low calculated to measured concentration ratios indicated which source compositions should be included.

In CMB it has been assumed that the source composition are linearly independent of each other. But in actual practice, often two sources have nearly identical compositions, or one source composition is almost identical to linear combination of other source compositions. This problem, “multicollinearity” is a potentially serious problem for CMB receptor models. One approach to the multicollinearity problem is to examine the source signatures and determine groups of sources whose compositions are chemically similar. One such group, for example, could include sources in the crustal category, such as soil

and road dust. Then a signature could be selected that represents the entire group. Subsequently, CMB could be applied for the selected groups of sources. This source selection approach reduces the multicollinearity problem and the potentially large errors that often are caused by it. However, grouping sources reduces the resolution of the source apportionment. It may also be possible to eliminate some sources on the basis of engineering judgment. A source should be eliminated on this basis only if there is strong physical evidence (such as wind direction data for the sampling period in question) that it does not contribute significantly to the monitored concentrations. Deleting sources may eliminate some of the multicollinearity problems (Henry, 1992).

The accuracy of CMB model depends upon the correct selection of number and nature of sources in the study region. One of the major short coming of CMB is, it requires experience and knowledge about the source composition in the study region. Therefore, detection of any unknown source is very difficult.

## **2.3 Multivariate Analysis**

### **2.3.1 Factor Analysis (FA):**

The use of Factor Analysis in the interpretation of aerosol data in terms of number and the nature of aerosol sources is important in pollution research. The inherent advantages of factor analysis are that no assumptions are made about the number of sources in advance and variables like particle sizes, meteorological parameters, concentrations of gaseous and particulate entities measured simultaneously can be included in the data set. The factor analysis is useful only when large data sets containing many variables are available. Henry et al. (1984) suggested that the minimum number of samples ( $n$ ) required for factor analysis should be such that  $n > 30 + (V + 3)/2$ , where  $V$  represents the number of variables.

The simplest form of factor analysis is the principal component analysis (PCA). PCA is a statistical technique that replaces a large number of inter-correlated variables with a smaller number of independent variables. In other words, a new set of variables are found as linear combinations of the measured variables so that the observed variations in the system can be reproduced by smaller number of these casual factors. It was originally developed and used in psychology to provide mathematical models of psychological

theories of human ability and behavior (Herman, 1976). The PCA applied in the field of air pollution is expressed as:

$$C_{ik} = \sum_{j=1}^N L_{ij} S_{jk} \quad \text{-----}(2.2)$$

where  $C_{ik}$  is the normalized value of concentration of the  $i$ th species for the  $k$ th sample.  $N$  is the total number of sources.  $S_{jk}$  is the factor score of the  $j$ th common factor ( $j$ th source) for the  $k$ th sample.  $L_{ij}$  is the factor loading of the  $i$ th species of the  $j$ th source.

The eigenvector analysis compresses the information content of the data set into as few eigenvectors as possible. In considering the number of factor used to describe the system, it is essential to carefully examine the problem of reconstructing both the variability within the data, and actual data itself. A large relative decrease in the magnitude of the eigenvalues is one indicator of the correct number of factors. It can often be useful to plot the eigenvalues as a function of factor number and look for sharp breaks in the slope of the line (Cattell, 1966). If the eigenvalue is a measure of the information content of the corresponding eigenvector, then only sufficiently large eigenvalues need to be retained in order to reproduce the variation initially present in the data. One of the most commonly used criteria for selecting the number of factors to keep is retaining only those eigenvalues greater than one. Kaiser and Hunka (1973) make a strong argument that although an eigenvalue greater than one does set a lower limit on the number of factors to be retained, it does not set a simultaneous upper bound (Hopke, 1999). There must be at least as many factors as there are eigenvalues greater than one; but there can be more than that number that are important to the understanding of the system's behavior.

The short comings of the Factor Analysis are:

1. It requires large datasets containing many variables.
2. It analyses the variations from a mean rather than absolute concentrations and where the variations from sample to sample are not much, it is not effective.
3. It cannot resolve aerosol sources that have very similar composition pattern.
4. Factor analysis can not give relative concentration of pollutants within the sources. The output indicates only the fraction of the variation of the concentration that is explained by the factors. To determine the contributions of

each source to the samples, a statistical method known as target transformation factor analysis (TTFA) is used

5. Unusual events or errors in sampling and analysis results in very high or very low values (outliers) of one or more variables. These outliers have to be removed before factor analysis to avoid error propagation.

### **2.3.2 Target Transformation Factor Analysis (TTFA):**

Target Transformation Factor Analysis (TTFA) is a hybrid approach as it combines some of the best features of chemical mass balance and factor analysis. The objective of TTFA is to extract a maximum information on the number and nature of sources with no or very limited a priori information other than the elemental compositional data. One operational difference between TTFA as applied by Hopke and co-worker (Hopke et al., 1980; Alpert and Hopke, 1980) and ordinary factor analysis is the use of correlation about the origin rather than the mean, and between samples rather than between the elemental species.

In order to calculate source contributions, it is necessary to rescale the source vectors determined by TTFA so that they are comparable with actual source composition vectors. The necessary scaling factors are determined by multiple regression of total mass against the (unscaled) factor loadings for each sample period. Finally, ordinary weighted least squares CMB is used to calculate source contributions from the rescaled TTFA source vectors.

### **2.3.3 Positive Matrix Factorization Analysis:**

In order to overcome the problem of the negative loadings, to minimize the rotation ambiguity and to avoid the insufficient information derived only from the covariance matrix in factor analysis, an advanced algorithm called positive matrix factorization (PMF) in receptor modeling was used. PMF gives strictly non-negative results in every loading matrix (Paatero and Tapper, 1994). The major difference between the principal component analysis (PCA) and PMF is that non-negativity of factors (both loadings and scores) is built-in for PMF model. In addition, PMF does not rely on information from

the correlation matrix but utilizes a point by point least square minimization scheme. Therefore, the profile produced can be directly compared to the input matrix without transformation. This is the distinctive advantage of PMF over FA or PCA.

In PMF any matrix X of data of dimension n rows and m columns, where n and m are number of samples and the number of species respectively, can be factorized into two matrices, namely Z (n x p) and C (p x m), and the unexplained part of X, E, where p represents the number of factor extracted.

$$X = Z * C + E \text{ -----(2.3)}$$

The product of Z and C can be explain the systematic variation in X. The above expression is similar to that of PCA. However, PCA only concentrates on explaining the unweighted sum of the squares of residuals. High concentration samples (from certain times and/or sites) tend to dominate such analysis while data of low concentrations are virtually ignored. In contrast, the PMF model emphasizes information from all samples. This is achieved by weighting the square of residuals with the reciprocals of the squares of the standard deviation of the data values. Usually species with higher concentrations have larger absolute standard deviations, hence their weight coefficients in the fit are smaller than those in unweighted models such as PCA.

The objective of PMF analysis is to minimize Q, defined as

$$Q = \sum_{i=1}^m \sum_{j=1}^n \frac{e_{ij}^2}{s_{ij}^2} \text{ subject to } z_{ik} \geq 0 \text{ and } c_{kj} \geq 0 \text{ -----(2.4)}$$

where  $z_{ik}$  and  $c_{kj}$  are elements of Z and C, respectively. The residual  $e_{ij}$  are defined by

$$e_{ij} = x_{ij} - \sum_{k=1}^p z_{ik} c_{kj} \text{ -----(2.5)}$$

and  $s_{ij}$  is the standard deviation of  $x_{ij}$ . The solution for Q is obtained by an iterative minimization algorithm “PMF2”.

The robust mode of using PMF has been suggested for analyzing environmental data (Paatero, 1996). It can avoid excessively large values in the data set, which can disproportionately affect the results. This can be achieved by introducing a filter function  $h_{ij}$  in the least square minimization of Q (Paatero,1997).

$$Q = \sum_{i=1}^m \sum_{j=1}^n \frac{e_{ij}}{h_{ij} s_{ij}} \text{ -----(2.6)}$$

$$\text{where, } h_{ij} = 1 \text{ if } \left| \frac{e_{ij}}{s_{ij}} \right| \leq \alpha$$

$$\text{and } h_{ij} = \left| \frac{e_{ij}}{s_{ij}} \right| / \alpha \text{ otherwise.}$$

The parameter  $\alpha$  is called the outlier threshold distance. Data value laying more than  $\alpha$  standard deviation above and below the fitted value will be treated as outliers and receive decreased weight in PMF 2.

Both CMB and PMF provide quantitative estimates of the source contributions. In the CMB analysis, source profiles are provided whereas in PMF, the source profiles are estimated. If some of the source profiles are known, they can be used in PMF to constrain the extracted source profiles and thereby reduce the rotational indeterminacy. With the PMF it is not possible to precisely assign errors to the source profiles and contributions. In a CMB analysis, it is possible to assign error estimates to each source contribution value. However, there are no diagnostics provided in the CMB model that would alert the user to misspecification of the source profiles. Also since the CMB analysis is done on a sample-by-sample basis, there can be errors in the estimated source contributions because of the variations that can occur in the source profiles. PMF uses all of the data and thus, estimates the average source profile over the time interval during which samples were acquired.

#### **2.4 Source-Receptor Relationships Incorporating Back Trajectories Methods:**

The transportation of the pollutants from the source to the sampling location can be described with the help of dispersion models. However, using an analogous model of atmospheric transport, a trajectory model calculates the position of the air being sampled backward in time from the receptor site from various starting times throughout the sampling interval. There exit two different ways to view air motions, namely the Eulerian and the Lagrangian perspectives. The first one focuses on points fixed in space through which the air flows, the second one focuses on individual air parcels as they move through time and space. The paths of these air parcels are known as trajectories. If the initial position  $X_0$  at time  $t_0$  of the air parcel is known then it may be followed either

forward (forward trajectories) or backward (backward trajectories) in time. The spatial coordinates  $X_0$  and  $t_0$  provide a means of identifying each air parcel for all time. The initial coordinates are called material or Lagrangian coordinates. The trajectories are then used in various types of analysis. Residence Time Analysis, Quantitative Bias trajectory Analysis, Potential Source Contribution Function Analysis and Time Weighted Concentrations Analysis are the most popular methods available in the literature which incorporate back trajectories.

The major advantages of back trajectories analysis are:

1. The geographical location of the sources can be identified.
2. Source parameter measurements not required.
3. It does not require any source profiles.
4. Problems that could be arises by the use of different sampling equipments or analytical methods at the source versus the receptors, or by the use of measurements widely separated in time are avoided.
5. Fugitive emissions, which are difficult to measure by the conventional methods, can often be observed.

#### **2.4.1 Residence Time Analysis:**

An estimate of the effect anthropogenic sources have on pollutant concentrations at a sensitive receptor requires a knowledge of the pathways that air parcels travel as they move from the source to the receptor. It is important to know how long the air parcel remains in a given region as it moves towards a receptor. Air parcel which travel quickly through a pollutant source region have less time to accumulate pollutants than do air parcels which remain in the source region for an extended time. Knowledge of spatial and temporal dynamics of air parcel movement can be gained by modeling air trajectories from a site backward in time by some suitable technique.

In residence time analysis gridded array is created around the sampling location. Suppose the previous positions, specified at some time increment, of an air parcel which has certain pollutant characteristics and which arrived at a point of interest at some time  $t$  can be identified. These positions (longitude and latitude coordinates) can be used to define the air parcel's trajectory. Trajectories are a sequence of segments, each of which

represents a fixed amount of time. Thus, each endpoint can be considered to be an indication that the air parcel has spent a given time within that gridded cell. The total residence time that air spends in the given cells would be the total number of endpoints that fall into that cell. These values can be plotted over a map. The residence time values associated with high concentration can be plotted to examine likely directions from which contaminated air is transported to the sampling site.

The problem with this method is that all of the trajectories begin at the receptor site and thus the residence time is maximum in cells surrounding the sampling location. To overcome this problem Poirot et al. (1986) considered an annular ring of width equal to that of a grid cell at a distance  $D_{ij}$  from the receptor site. The area of this annulus can then be compared to that of a grid cell.

$$\frac{\text{Annular area}}{\text{Single grid cell area}} = \frac{\pi[(D_{ij} + L)^2 - (D_{ij} - L)^2]}{L^2} \text{-----}(2.7)$$

where L is the half-width of a grid cell. This ratio increases as the distance increases so that its inverse will represent the decrease in the probability that a grid cell will be traversed by a trajectory compared with a centrally located cell. Thus, we can consider the right hand side of the equation (2.7) as a geometric adjustment factor  $F_{ij}$ . The residence time of any specific grid cell is then multiplied by  $F_{ij}$  to yield the geometrically corrected residence time.

**2.4.2 Quantitative Transport Bias Analysis:**

The Quantitative Transport Bias Analysis (QTBA) explicitly incorporates deposition along the back trajectory. The transport potential for a given sampling time includes the mean transport computed by using the trajectory model plus the horizontal spread imposed by atmospheric dispersion and uncertainty in transport estimates. The probability of a reactive, depositing tracer arriving at a point  $x$  at a time  $t$  can be expressed by Cass (1981) as

$$A(x,t) = \int_{t-\tau}^t \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x,t/x',t') dx' dt' \text{-----}(2.8)$$

Where  $T(x,t/x',t')$  is the potential mass transfer function in two dimensions and is defined by,

$$T(x,t/x',t') = Q(x,t/x',t') R(t/t') D(x',t') \Lambda(x',t') \text{ -----(2.9)}$$

In which  $Q(x,t/x',t')$  is the transition probability density function of an air parcel located at  $x'$  at time  $t'$  arriving at a receptor site  $x$  at time  $t$ ,  $R(t/t')$  is the probability of the tracer not being reacted to another species in the interval between  $t'$  and  $t$ .  $D(x',t')$  is the probability that the tracer was not dry deposited at  $(x',t')$  and  $\Lambda(x',t')$  is the probability that the material is not lost by wet deposition at  $(x',t')$ . The integration was conducted over a time period  $\tau$ , the length of back trajectory calculation.

The probabilities for deposition are poorly known and various formulations of these probabilities functions could be used. The removal by wet and dry deposition is assumed to be linearly proportional to the concentration of the species of interest with the rate constant for dry ( $k_d$ ) and wet ( $k_w$ ) deposition to be defined as

$$k_d = v_d/h \text{ -----(2.10)}$$

$$k_w = \theta_i \rho/h \text{ -----(2.11)}$$

where  $v_d$  is the dry deposition velocity,  $\theta_i$  is the washout ratio for the  $i$ th species,  $\rho$  is the precipitation rate (mm/hr) and  $h$  being the mixing height.

The transition probability density function,  $Q(x,t/x',t')$  is approximated as a normal distribution about the trajectory center line with a standard deviation that increases linearly with time in the upwind direction, and can be expressed as

$$Q(x,t/x',t') = (1/2\pi\sigma_x\sigma_y) \exp [(-1/2)\{(x''/\sigma_x) + (y''/\sigma_y)\}^2] \text{ -----(2.12)}$$

where  $x'' = x + x'(t')$  and  $y'' = y + y'(t')$  with  $(x,y)$  being the coordinates of the grid and  $x'(t')$  and  $y'(t')$  being the coordinates of the center line of the trajectory. It is assumed that the standard deviations grow with time and can be approximated by

$$\sigma_x(t') = \sigma_y(t') = at' \text{ -----(2.13)}$$

with a dispersion speed,  $a = 5.4$  Kms/hrs.

The potential mass transfer field for a given trajectory,  $T(x,t/x',t')$ , was integrated over the upwind period,  $\tau$ , of each trajectory to produce a 2-D probability of transport field.

The integrated potential mass transfer potential for a given trajectory,  $k$ , arriving at time  $t$  is then calculated as

$$\tilde{T}_k(x/x') = \frac{\int_{t-\tau}^t T(x,t/x',t') dt'}{\int_{t-\tau}^t dt'} \quad \Lambda \Lambda \quad (2.14)$$

This probability field predicts the possibility of contribution from any upwind area to a specific receptor site without regard to the actual location of the emission sources.

The measured concentrations of the species can then be used to calculate the transport bias. Using the potential mass transfer fields calculated for each trajectory, k, weighted by the corresponding concentration measured during that sampling interval results in a concentration-weighted mass transfer potential defined

$$\tilde{T}(x/x') = \frac{\left[ \sum_1^k \tilde{T}_k(x/x') \chi_k(X) \right]}{\sum_1^k \chi_k(X)} \quad \text{-----}(2.15)$$

where  $\tilde{T}_k(x/x')$  is the integrated potential mass transfer and  $\chi_k(X)$  is the concentration of the species of interest measured at x when the trajectory k arrive. The difference between the weighted and the unweighted field represents the degree of bias associated with the transport of the species of interest to the receptor site.

The concentration-weighted mass transfer potential is then divided by the unweighted mass transfer potential, resulting in a dimensionless ratio. The ratio is then multiplied by the mean concentration of the species of interest  $\bar{X}$  yielding,

$$QTBA = P(x/x') = \frac{\tilde{T}(x/x')}{\tilde{T}_k(x/x')} \bar{X} \quad \text{-----}(2.16)$$

From an individual receptor the  $\tilde{T}(x/x')$  field indicates the direction and preferred path associated with above average concentration. However this does not address the distance from the receptor to the contributing sources. The results of QTBA for a single site would suggest that the source was somewhere upwind along the corridor of the highest probability for delivering the above average concentrations. This shortcoming can be address through the use of concurrent measurements at multiple stations. By over laying

the QTBA field for each receptor, systematic patterns of transport of higher concentration from particular source area to multiple receptors are identified.

### 2.4.3 Potential Source Contribution Function (PSCF) Analysis:

One receptor modeling method that simultaneously process chemical, geophysical and meteorological information is Potential Source Contribution Function (PSCF) analysis. A PSCF analysis combines the ambient concentration of a single chemical species with meteorological data in the form of backward air parcel trajectories and is able to yield a gridded map identifying the possible geographic locations of emission sources for this particular species. To calculate the PSCF, the whole geographic region covered by the trajectories is divided into an array of grid cells whose size is dependent on the geographical scale of the problem so that the PSCF will be a function of locations as defined by the cell indices  $i$  and  $j$ .

Let  $n_{ij}$  be the number of trajectory segment end points residing in the  $ij$ th cell during a time interval  $T$ , and  $N$  the total number of end points computed for the time interval, then the probability of this event  $A_{ij}$ , is given by

$$P[A_{ij}] = n_{ij}/N \text{ -----(2.17)}$$

Where  $P[A_{ij}]$  is the probability of residence time of a randomly selected air parcel in the  $ij$ th cell relative to the total time interval  $T$ .

Suppose  $m_{ij}$  be the number of segment endpoints in the same  $ij$ th cell for those trajectories which arrive at the receptor when the pollutant concentration is greater than the cutoff value. The probability of this high concentration event,  $B_{ij}$  is given by  $P[B_{ij}]$

$$P[B_{ij}] = m_{ij}/N \text{ -----(2.18)}$$

Like  $P[A_{ij}]$ , this subset probability is related to the residence time of air parcel in the  $ij$ th cell but the probability  $B$  is for the contaminated air.

The Potential Source Contribution Function (PSCF) is defined as the conditional probability that an air parcel which passed through the  $ij$ th cell had a high concentration upon arrival at the trajectory endpoint. Regions with a high PSCF value are those that have a high potential to adversely affect air pollution concentrations at the receptor. The conditional probability is found by,

$$PSCF = P[B_{ij}/A_{ij}] = P[B_{ij} \cap A_{ij}]/P[A_{ij}] \text{ -----(2.19)}$$

Since the set of events  $B_{ij}$  is a subset of the set  $\{A_{ij}\}$ , the intersection of  $B_{ij}$  and  $A_{ij}$  is simply  $B_{ij}$ . Thus ,

$$PSCF = P[B_{ij}/A_{ij}] = P[B_{ij}]/P[A_{ij}] = m_{ij}/n_{ij} \text{ -----(2.20)}$$

It is important to note that the high potential contribution to air pollution represented by a high conditional probability is realized only if an air parcel actually passes over the region. Cells containing emission sources would be identified with conditional probabilities close to one of the trajectories that have crossed the cells effectively transport the emitted contaminant to the receptor site. The PSCF model thus provides a mean to map the source potentials of geographical area.

If only a small number of trajectory endpoints are associated with a grid cell, there is a higher degree of uncertainty associated with that cell's PCSF value (i.e., it is possible to have a PCSF value of one when only one trajectory passed over that grid space). Zeng and Hopke (1989) suggested that such PSCF values should not be given the same weight as a value resulting from a large number of trajectory endpoints. The arbitrary weight function is as follows:

$$W(n_{ij}) = \begin{cases} 1.00 & \text{if } n_{ij} > 4 \\ 0.85 & \text{if } n_{ij} = 3 \\ 0.65 & \text{if } n_{ij} = 2 \\ 0.50 & \text{if } n_{ij} = 1 \end{cases}$$

Thus, these weights have been imposed to reduce the values in cell with low numbers of end points.

#### **2.4.4 Residence Time Weighted Concentrations Analysis:**

Stohl (1996) has developed a method of combining chemical concentration data with air parcel back trajectories. In this method, the trajectories are weighted by the measured concentrations at one or more receptor sites in order to develop a concentration contribution field. This field can be used to estimate the relative contributions for the various grid cells to the observed receptor site concentrations. However, it cannot be extended to yield emission estimates without taking wet and dry deposition and dispersion into account.

The analysis begins with the calculation of a trajectory-weighted mean concentration for each grid cell,

$$\ln(\bar{C}_{ij}) = \frac{1}{\sum_{k=1}^N \tau_{ijk}} \sum_{k=1}^N \ln(C_k) \cdot \tau_{ijk} \text{ -----(2.21)}$$

where i, j are the indices of the grid cell, k is the index of the trajectory segment, N is the total number of trajectories used for the computation,  $C_k$  is the pollutant concentration measured upon arrival of trajectory k at the receptor site, and  $\tau_{ijk}$  is the time spent by the trajectory k in the grid cell (i,j).

This approach weights each grid cell along the trajectory with the same value, while in reality the pollutant emissions just take place in some specific segments. Therefore this method underestimate the spatial gradients of the true source fields. If two trajectories generally pass over the same cells except that one includes the pollutant emission grid cell, one will be associated with a low concentration while the other with a high value. To account for this, an iterative scheme was developed by Stohl (1996), which redistributed the measured concentrations along the trajectories according to the estimated concentration field from the previous iteration.  $\bar{C}_{ij}$  was calculated from the above equation and serve as first guess field. For simplicity, let us consider only single trajectory k, which is split into  $M_k$  segments with 1 hrs length each.  $C_k$  is again the pollutant concentration measured upon arrival of trajectory k. Let  $X_{mk}$  be the mean concentration of the grid cells which are hit by the segments  $m=1, \dots, M_k$  of the trajectory k ( $X_{mk} = \exp \bar{C}_{ij}$ ), and  $\bar{X}_k$  is the average concentration of the grid cells hit by the  $M_k$  segments of the trajectory k. Then the redistributed concentrations along the trajectory k can be computed as follows:

$$C_{mk} = C_k \frac{X_{mk}}{\bar{X}_k} \quad m = 1, \dots, M_k. \text{ -----(2.22)}$$

This equation can now be applied to all individual trajectories.

## **2.5. Comparison of different receptor modeling techniques:**

Table 2.1 summarizes different receptor modeling techniques available in the literature. Chemical mass balance (CMB) is routinely used in atmospheric source apportionment of ambient air pollutants. But it requires a priori knowledge of the number of sources in the study region and also the location-specific source profiles. In the absence of location specific source profile the source profiles available in the literature can be used, but the error resulting from using source profiles from other regions to represent local sources are not known. A study conducted by Sharma and Patil (1994) suggested that US EPA source profile were not suitable to use in chemical mass balance models for India. The major advantage of CMB is that the sample size required is very small. The inherent advantages of the multivariate analysis are that no assumptions are made about the number of sources in advance and variables like particle sizes, meteorological parameters, concentrations of gaseous and particulate entities measured simultaneously can be included in the data set. But it requires large datasets containing many variables for better resolution of sources. The major advantage of back trajectory analysis is that the possible geographical location of the sources can be identified. It neither requires any source profiles nor the source parameter measurements for identifying the possible source locations. The major disadvantage of back trajectory based receptor models is that a large number of samples are required for accurate prediction of source locations. With the help of multivariate and back trajectory based receptor-modeling techniques unknown or fugitive sources can be identified.

Receptor models, both the mathematical (Multivariate Analysis, CMB) and trajectory (RTA, QTBA, and PSCF) types promise to be helpful tools for source attribution for gaseous air pollutants, suspended particulate and regional haze. However, each of these techniques also has limitations, and none of them should be considered a “stand alone” technique (Polissar et al., 2001). Both the mathematical and trajectory approaches are limited by the detail of the input (species concentrations or meteorological) data. Both techniques are sensitive to systematic errors or biases in the input data, including those introduced by the modelers, such as in treatment of variables, which are missing, or below detection limits. The mathematical techniques can identify sources of influence on the data, but a good deal of “subjective” judgment is inevitably required in the

interpretation of what these identified sources actually represent. The ensemble trajectory techniques produce only qualitative indications of predominant transport patterns. In the absence of very costly chemical tracer experiments, there are virtually no methods to evaluate the performance of any of these mathematical or trajectory receptor models (Poirot et al., 2001). Inter-comparison of multiple receptor techniques with experimental measurements will be a useful future approach for improving the understanding of source-receptor relationships and for improving the confidence in the individual model results.

**Table-2.1: Summary of the Receptor Modeling Techniques**

<b>Model</b>	<b>Requires Source Information</b>	<b>Quantitative Apportionment</b>	<b>Source Location Information</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>Chemical Mass Balance</b>	Yes	Yes	No	Easy to use, Single sample is sufficient. includes errors in source components	Can misspecify sources; does not include variability in source compositions; sensitive to collinearity Detection of unknown sources not possible.
<b>Principal Component Analysis</b>	No	Yes	No	Easy to use, identification of source types; identification of unusual samples Detection of fugitive sources possible.	Depends on correlation that can be driven by meteorology or collocation. Large number of data set required.
<b>Positive Matrix Factorization</b>	No	Yes	No	Treats missing and detection limit data well; estimates uncertainties Detection of fugitive sources possible	Assumes that all sources have been identified. Large number of data set required.
<b>Residence Time Analysis</b>	No	No	Yes	Provides likely source area Detection of fugitive sources possible	Needs sufficient number of samples
<b>Area Influence Analysis</b>	No	No	Yes	Provides likely source area Detection of fugitive sources possible	Needs sufficient number of samples
<b>Quantitative trajectory Bias Analysis</b>	No	Yes	Yes	Incorporate dispersion, wet and dry deposition; combines measurements from different sites Detection of fugitive sources possible	Requires detailed information of precipitation along trajectory
<b>Potential Source Contribution Function</b>	No	No	Yes	Provides likely source area Detection of fugitive sources possible	Needs sufficient number of samples
<b>Residence Time Weighted Concentration</b>	No	Possibly	Yes	Provides likely source area and estimates of their relative contribution Detection of fugitive sources possible	Needs sufficient number of samples

## Chapter-3

### Work done on Source Apportionment of Air Pollutants at Trombay

#### 3.1 Collection of particulate matter samples:

Monitoring of the particulate matter was started from November 2003, and two samples per week were collected. All the samples were collected for 24 hours from the sampling site. The samples were collected at a height of 8 meters from the ground level. Pre-weighed filter papers were used to collect the samples. A strict regime of quality control and quality assurance was followed. Field blanks were collected to quantify possible contamination of the samples during collection, transport, storage.

Thus, 116 samples each of  $PM_{10}$  (particulate matter having aerodynamic diameter  $10\ \mu\text{m}$ ) and particulate matter having aerodynamic diameter greater than  $10\ \mu\text{m}$  ( $>PM_{10}$ ) have so far collected using a high volume respirable dust sampler (Envirotech, APM 460). The sampler uses the cyclone impaction principle to separate the coarser particles ( $>PM_{10}$ ) from the air stream before filtering it on the filter paper. The  $>PM_{10}$  fraction can also be assessed by collection of the dust retained in the cyclone cup.

Besides this 89  $PM_{2.5}$  (particulate matter with aerodynamic diameter less than  $2.5\ \mu\text{m}$ ) and  $PM_{2.5-10}$  (particulate matter with aerodynamic diameter between  $2.5\ \mu\text{m}$  and  $10\ \mu\text{m}$ ) samples were collected using a fine particulate sampler (Envirotech, APM 550). The fine particulate sampler uses a set of impactor standardized and documented by USEPA to separate coarse particulates from air stream. It has a circular symmetry air-inlet, so that air entry is unaffected by wind direction and is designed to keep out very large particles. The inlet section immediately leads to an impactor stage designed to trap particles with an aerodynamic diameter larger than  $10\ \mu\text{m}$ . Thus the air stream in the down tube consists of only medium and fine particulates. The streamline air flow in down tube is accelerated through the nozzle of the well shaped (WINS) impactor designed to trap medium size particulates with an aerodynamic diameter between  $2.5$  and  $10\ \mu\text{m}$ . To avoid sampling errors due to the tendency of small particles to bounce off the impaction surface a  $37\ \text{mm}$  diameter filter paper immersed in silicone oil is used as an impaction surface. The air stream leaving the WINS impactor consists of only fine particulates with an aerodynamic diameter smaller than  $2.5\ \mu\text{m}$ . These fine particles are collected on a filter paper of

47 mm diameter. The impactor system is designed to operate at an air-flow rate of 1 m<sup>3</sup>/hr or 16.7 LPM. The fine particulate sampler uses an oil-free rotary pump to produce the suction pressure and a critical flow orifice for maintaining a constant sampling rate.

### **3.1.1 Results on Particulate Matter Levels:**

Simultaneous collection of PM<sub>2.5</sub>, PM<sub>2.5-10</sub>, PM<sub>10</sub>, >PM<sub>10</sub> and total suspended particulate matter (TSPM = PM<sub>10</sub> + >PM<sub>10</sub>) has been done at the sampling site. The concentration of different fraction of the particulate matter is given in Table 3.1. The PM<sub>2.5</sub> fraction was found to have mass concentration ranging between 32.7 µg/m<sup>3</sup> and 64.0 µg/m<sup>3</sup> with a mean concentration of 46.7 µg/m<sup>3</sup>. The mass concentration of the PM<sub>10</sub> was found to be within the range of 53.5 µg/m<sup>3</sup> to 138.4 µg/m<sup>3</sup>. At the sampling site the mean concentration of PM<sub>10</sub> was found to be 98.3 µg/m<sup>3</sup>. A good correlation has been observed between PM<sub>2.5</sub> and PM<sub>10</sub>. The result has been shown in Fig 3.1. The result indicates more than 80% of the variance in the PM<sub>10</sub> mass has been explained by the variation in the PM<sub>2.5</sub> mass. This is because PM<sub>2.5</sub> comprises a large fraction of PM<sub>10</sub>. At the sampling site the ratios of PM<sub>2.5</sub> to PM<sub>10</sub> varies from 0.41 to 0.64 with an average value of 0.55. The high PM<sub>2.5</sub> to PM<sub>10</sub> mass ratio at the sampling site may be due to the combined contribution from the vehicular as well as industrial emission. The total suspended particulate matter at the sampling site was found to be in the range of 98.1 µg/m<sup>3</sup> to 206 µg/m<sup>3</sup> with a mean concentration of 160.4 µg/m<sup>3</sup>. A good positive correlation has been observed between PM<sub>10</sub> and TSPM mass concentrations (Fig. 3.2). At the sampling site the ratios of PM<sub>10</sub> to TSPM varies from 0.34 to 0.78 while that between PM<sub>2.5</sub> and TSPM was found to be in the range of 0.13 to 0.38.

**Table 3.1 Concentrations of different size fractionated particulate matter and their ratios**

	Minimum	Maximum	Mean $\pm$ S.D.
PM <sub>2.5</sub> (89)*	32.7 $\mu\text{g}/\text{m}^3$	64.0 $\mu\text{g}/\text{m}^3$	46.7 $\pm$ 16.8 $\mu\text{g}/\text{m}^3$
PM <sub>2.5-10</sub> (89)	20.5 $\mu\text{g}/\text{m}^3$	78.3 $\mu\text{g}/\text{m}^3$	38.9 $\pm$ 20.5 $\mu\text{g}/\text{m}^3$
PM <sub>10</sub> (116)	53.5 $\mu\text{g}/\text{m}^3$	138.4 $\mu\text{g}/\text{m}^3$	98.3 $\pm$ 32.8 $\mu\text{g}/\text{m}^3$
>PM <sub>10</sub> (116)	42.4 $\mu\text{g}/\text{m}^3$	102.3 $\mu\text{g}/\text{m}^3$	63.7 $\pm$ 28.4 $\mu\text{g}/\text{m}^3$
TSPM (116)	98.1 $\mu\text{g}/\text{m}^3$	206 $\mu\text{g}/\text{m}^3$	160.4 $\pm$ 69.3 $\mu\text{g}/\text{m}^3$
PM <sub>2.5</sub> / PM <sub>10</sub>	0.41	0.64	0.55 $\pm$ 0.24
PM <sub>2.5</sub> / TSPM	0.13	0.38	0.22 $\pm$ 0.13
PM <sub>10</sub> / TSPM	0.34	0.78	0.65 $\pm$ 0.37

\* Bracketed value indicates number of samples

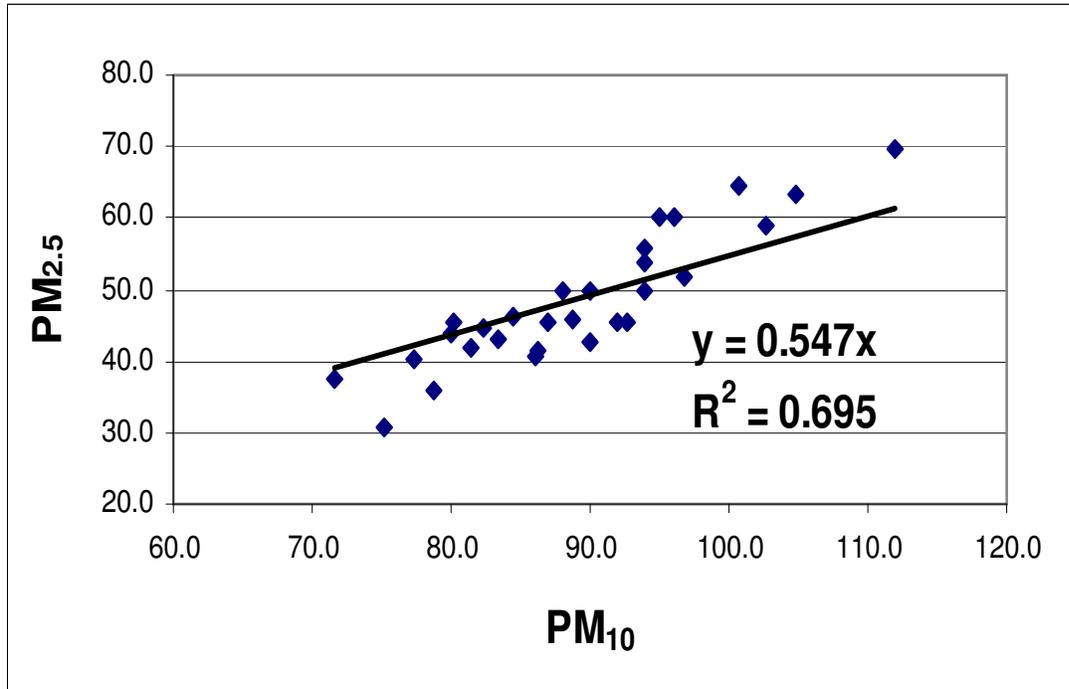


Fig. 3.1 Correlation between the PM<sub>10</sub> and PM<sub>2.5</sub> mass concentrations

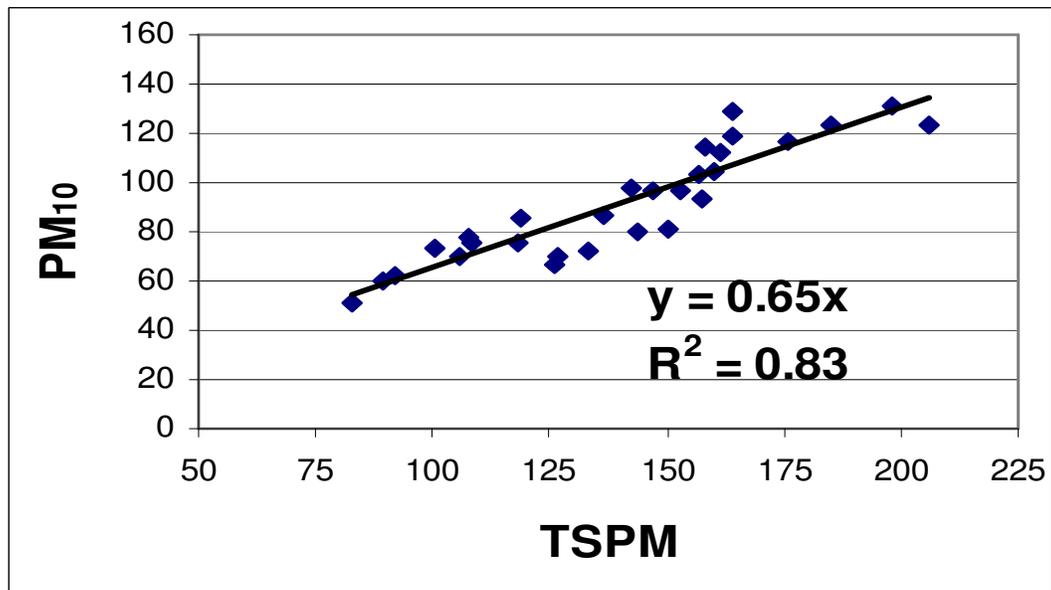


Fig. 3.2 Correlation between the PM<sub>10</sub> and TSPM mass concentrations

## **3.2 Chemical Characterization:**

### **3.2.1 Trace and Toxic Elements in Ambient Air:**

The inorganic component of the particulate matter mainly consists of trace elements. Emissions of the trace elements to the atmosphere are normally considered the most important pathways, as the atmosphere is probably the main avenue for the transport of these pollutants. These toxic pollutants are contributed to the atmosphere from a wide variety of natural and anthropogenic sources. Natural sources include dust raised by winds, volcanic activity, forest fires, vegetation and sea salt sprays. The major anthropogenic sources that contribute to the elemental concentration in the atmosphere are stationary fuel combustion, internal combustion engines, non-ferrous metal manufacturing, iron, steel and ferro alloy plants, waste incineration and cement production (Pacyna, 1986). Traditionally, the elemental tracers in particulate matter have been used as a tool for source apportionment because they are source specific and maintain their integrity in the atmosphere. Thus, it is commonly accepted that Al in aerosols is a tracer of crustal material, while Pb comes mainly from leaded gasoline and smelter, and V is from the combustion of heavy fuel oil rich in V-porphyrins (Nriagu and Pacyna, 1988). It has thus been demonstrated that trace element tracers are effective in providing source information on crustal, marine, and anthropogenic sources.

Particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>) samples were analysed for the characterization and quantification elemental constituents. The filter paper samples were acid digested using a mixture of nitric acid and perchloric acid in the ratio of 3:1. After gentle heating for nearly 8 hrs, a colourless solution was obtained. Atomic absorption spectrometer (AAS), was used for the analysis of acid soluble fraction of PM<sub>10</sub> and PM<sub>2.5</sub> aerosols.

### **3.2.2 Organic Pollutants in Ambient Air:**

Organic compounds are important components of particulate matter, whether in urban, rural or remote areas. Recent work has emphasized the characterization of individual compounds that might serve as tracer of specific source categories of organic aerosol. Identification of organic compounds in particulate matter is currently an active area of research, especially in regards to

understanding the chemical composition of polar organic fractions, to identifying source category tracers for use in source apportionment, and for determining which organic compounds affects human health. The organic matter in aerosols is derived from two major sources: biogenic detritus (plant, wax, microbes, etc.) and anthropogenic emissions (Hildemann et al., 1991). Thus, the organic markers can provide additional information on the emissions from combustion, microbial activities, and biogenic contribution from higher plant wax.

Among the organic compounds of anthropogenic origin, the polycyclic aromatic hydrocarbons (PAHs) make up a significant group. PAHs are the compounds containing two or more aromatic (benzene) rings including only carbon and hydrogen in their structure fused together in different possible arrangements (Dyremark, 1994). PAHs are released into the atmospheric air through incomplete combustion of fossil fuels. The urban sources are coal-fired power plants, organic flares in petrochemical industries and automobile exhausts. Refuse incineration is also a potential source. Releases from domestic cooking stoves and cigarette smoke are two indoor sources of these compounds. PAHs are the main families of solvent extractable organic compounds in aerosols and each possesses characteristic source information.

The characterization and quantification of different PAH was carried out using a Shimadzu HPLC system (LC-10 AD) with UV-visible detector. Before analysis the filter paper samples were extracted ultrasonically for forty minutes with 50 ml of hexane as solvent. Sixteen different PAH compounds were analysed with the help of HPLC using isocratic method.

### **3.2.3 Carbon Analysis:**

Two classes of carbon are commonly measured in ambient aerosol samples: organic or non-light absorbing carbon and elemental or light absorbing carbon. Several analytical methods for the speciation of organic and elemental carbon in ambient and source particulate samples have been evaluated. Chow and Watson (1998) summarize different carbon analysis methods along with their measurement principles. For source apportionment by receptor models, several consistent but distinct fractions of carbon in both source and receptor samples are desired, regardless of their chemical properties. Differences in ratios of the carbon concentrations in these fractions

form part of the source profile that distinguishes the contribution of one source from the contribution of other sources (Watson et al., 1994).

In the current study the mass concentration of EC, OC and total carbon (TC = EC + OC) were measured in the particulate matters (PM<sub>10</sub> and PM<sub>2.5</sub>) using total organic carbon analyzer (Shimadzu TOC analyzer, Model SSM 5000A). The total carbon and organic carbon were quantified by thermo-graphic method. The filters were burnt in oxygen atmosphere and the release of CO<sub>2</sub> was measured by a non-dispersive infrared (NDIR) detector. For measurement of OC the filter was heated up to 620 °C and for TC up to 900 °C. The EC was calculated by subtracting OC from TC.

### **3.3 Application of Receptor Modeling Techniques at Trombay:**

Pollution sources affecting the sampling site were statistically identified using varimax rotated factor analysis of the concentration data. Factor analysis is a statistical technique which can be applied to a set of variables in order to reduce their dimensionality. That is to replace a large set of intercorrelated variables with a smaller number of independent variables. These new variables (components) are derived from the original variables and are simply linear combinations of those variables. It is important to note that, for multivariate analysis such as proposed in this work, sufficient degrees of freedom should be available in the model. Thus, the dataset employed must have many more observation than variables if stable results are to be derived.

#### **3.3.1 Source Apportionment of Polycyclic Aromatic Hydrocarbons**

Factor analysis with varimax rotation was applied to the concentration data to find out various possible PAHs sources. Five factors were obtained based on the criteria of eigen value greater than 1 (Table-3.2).

Factor 1: It explains about 35% of the variance with a high loading for Benzo(k)Fluranthene, Benzo(b)Fluranthene, Phenantherene and Perylene. According to Hwang et al. (2003) the presence of PAHs with 3 and 4 aromatic rings, such as Benzo(k)Fluranthene, Benzo(b)Fluranthene and Phenantherene, indicates the dominance of diesel combustion. Therefore, the presence of these compounds can be attributed to diesel exhausts.

Factor 2: It explains 18% of the variance with high loads in Benzo(a)Anthracene, Benzo(a)Pyrene and Dibenzo(a,h)Anthracene and a lower load for Perylene. Kulkarni and Venkataraman (2000) have shown that Benzo(a)Pyrene can be a good tracer for wood burning while Freeman and Cattell (1990) showed that this compound may be associated with combustion of vegetation. Masclet et. al. (1986) has shown that Dibenzo(a,h)Anthracene can be a typical tracer for incineration. Therefore it is very difficult to assign any specific source for this factor. However, it is more appropriate to attribute this factor to biomass burning and incineration.

Factor 3: This factor is responsible for 14% of the total variance. This factor is predominantly weighted in Anthracene and Benzo(a)Anthracene. Khalili et al. (1995) identified that Anthracene, Phenanthrene and Benzo(a)Anthracene were tracer for coal combustion. Therefore coal combustion can be characterized as the source for this factor.

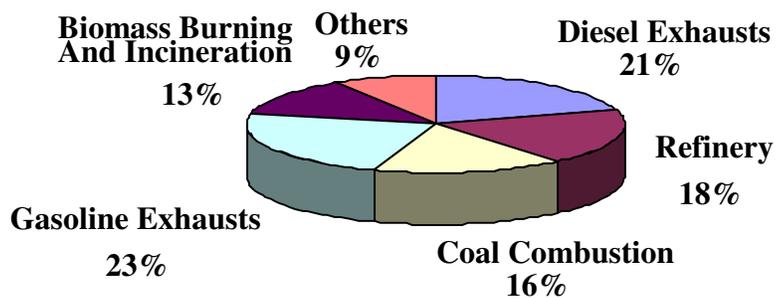
Factor 4: The fourth factor is responsible for 12% of the total variance. This factor is weighted in Benzo(ghi)Perylene and Fluranthene. Miguel et al. (1998) have shown that Benzo(ghi)Perylene is characteristic of vehicular emissions from gasoline engines. Therefore this factor can be attributed to the vehicular emissions for gasoline exhausts.

Factor 5: It explains 10% of the variance and has a high load for Fluorene and Pyrene. Masclet et al. (1986) has shown that light PAHs (3 and 2 aromatic rings) are predominant in emissions from petrol refinery and are characterized by high concentration of Fluorene and Pyrene. Thus in this work this factor is associated with refinery emissions.

The multiple linear regression was performed to determine the percent contribution of different PAH sources (Fig. 3.3). The mean percent contribution from the diesel exhausts was found to be 20.6 % while that from gasoline exhausts was 23.2 %. So, the vehicular exhaust contributed almost 44 % of the total PAHs at the sampling site. 18.4 % of the total PAHs were coming from refinery sources. The coal combustion contributed 16.3 % to the total PAHs while the biomass burning and incineration contributed 12.5 %. The remaining 9 % is contributed by other sources.

**Table-3.2: Factor Loading Matrix After Varimax Rotation**

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Fluorene	0.089	-0.135	0.057	-0.005	0.818
Phenanthrene	0.853	0.152	-0.135	0.175	-0.285
Anthracene	-0.077	-0.056	0.921	0.139	0.011
Fluoranthene	0.003	0.003	0.231	0.919	0.099
Pyrene	0.454	-0.384	0.003	-0.102	-0.607
Benzo(a)anthracene	0.156	0.632	0.671	0.0135	0.160
Perylene	0.793	0.530	-0.026	-0.068	0.102
Benzo(b)fluoranthene	0.673	-0.032	-0.118	0.164	0.155
Benzo(k)fluoranthene	0.876	0.151	0.337	-0.082	-0.076
Benzo(a)pyrene	-0.023	0.741	0.138	0.126	-0.193
Benzo(ghi)perylene	0.284	0.466	-0.143	0.650	-0.136
Dibenzo(ah)anthracene	0.360	0.768	-0.130	0.063	0.130
	Diesel Exhausts	Biomass Burning and Incineration	Coal Combustion	Gasoline Exhausts	Refinery Emissions

**Fig 3.3. Percentage Contribution of Different Sources to Total PAHs**

### 3.3.2 Factor Analysis and Source Identification of Trace Elements in SPM

Varimax rotated Factor Analysis was carried out using SPSS software (SPSS, 1988) with SPM, PM<sub>2.5</sub> and measured trace metals as variables. Four factors were selected, based on the criteria of cumulative percentage variance~ 80% and eigen value > 1.0. Higher factor loadings of particular elements can help in identifying the possible sources. Varimax rotated Factor Analysis showed four possible groups or factors (based on factor loading > 0.7) indicating four different contributing sources for the trace metals and anions at the sampling location. Results of Varimax rotated factor Analysis for the sampling location are depicted in Table 29 and 30.

From Table 3.3 it is evident that total percentage variance explained by the four factors was more than 70%. Major portion of percentage variance (59%) was explained by factor 1. The factor loading in factor one comprises of SPM, Cd, Pb, Cu, Mn, Fe, Ca, Mg, K. The association of SPM with factor 1 indicates that the variation in SPM is mainly related with the variation in the species associated with this factor. Looking at the receptor domain, since a major highway is located at about 300 metres away from the sampling location, it may be inferred that the factor 1 may represent contribution from wind blown soil or resuspended dust. Pb is not currently being used as an anti-knocking agent; however, the Pb which is chemically bound to the soil originating from the emission of earlier years can be reason for the association with factor 1. However association of Fe and Mn with Factor 1 may be stated that Factor 1 can be identified as soil dust clubbed together by the contribution of metal industries.

Factor 2 explained percentage variance of 12 with high loading of Cr and Zn. It may represent source like refuse burning, as Zn is associated with refuse burning. Cr is also associated with this group, which is slightly questionable. There are several citations which has attributed high chromium in road dust and near by vegetation to vehicular emissions.

Na was separated as factor 3 with a percentage variance of 6.4. Factor 3 can represent contribution from sea salts. Mumbai is an island city in the Arabian Sea on the western margin of Deccan trap country and a significant amount of SPM in the Mumbai urban atmosphere is apportioned to marine aerosols.

The factor 4 showed a heavy loading of Co with percentage variance of 1.5. In absence of additional tracers it is difficult to identify the exact source of Co.

Source apportionment of the 24 hr. average SPM using Factor Analysis-Multiple Regression is shown in Table 31 for the sampling location. For the 24 hr. average SPM concentration, 63% is contributed by soil dust and metal industry.17% is contributed by refuse burning.6.9% marine aerosols, 1.6% from an isolated source (for Co) and 11.48% remains as unexplained contribution.

**Table 3.3: Factor Loading Matrix of SPM after Varimax Rotation**

	Factor 1	Factor 2	Factor 3	Factor 4
Mass	<b>0.86</b>	0.31	0.31	-0.07
Cd	<b>0.66</b>	0.26	0.31	0.33
Pb	<b>0.84</b>	0.21	0.37	-0.08
Cu	<b>0.85</b>	0.32	0.16	-0.07
Mn	<b>0.79</b>	0.44	0.11	-0.01
Co	-0.11	-0.18	0.09	<b>0.92</b>
Cr	0.39	<b>0.72</b>	-0.06	-0.07
Fe	<b>0.83</b>	0.33	-0.05	-0.16
Zn	0.22	<b>0.82</b>	0.08	-0.14
Ca	<b>0.85</b>	0.26	0.23	-0.10
Mg	<b>0.82</b>	0.14	0.03	0.03
K	<b>0.69</b>	-0.01	0.54	-0.21
Na	0.18	0.01	<b>0.92</b>	0.17
<i>Eigen value</i>	7.63	1.55	0.83	0.65
<b>Percent of Variance</b>	58.7	12	6.4	5
<b>Cumulative Percentage</b>	58.7	70.7	77	82

### 3.3.3 Factor Analysis and Source Identification of Trace Elements in PM<sub>2.5</sub>

According to Table 3.4, the total percentage variance explained by the four factors was more than 70%. In this case, factor with eigen value  $< 1$  is also selected as one factor to compare with four sources of SPM. Factor 1 with a percentage variance of 50, comprises of PM<sub>2.5</sub>, Cd, Pb, Mn, Cu, Fe and K. Pb, Cu, Mn, etc. are in less than 2.5  $\mu\text{m}$  size in high temperature combustion aerosols. Hence, factor 1 can be assumed to be coming from high temperature combustion aerosol. However, the association of Fe and K with Factor 1 can be explained in the same way as in SPM. Factor 2 with a percentage variance of 12, is clustered together with Ca and Mg. One probability is that factor 2 represents contribution from crustal soil sources, mainly wind blown. It may represent the contribution from cement as few civil construction activities were going on at a few meters away from the sampling site. Factor 3 with a percentage variance of 8, comprises of Co and Cr. Factor 3 can be explained as contribution from any specific metal industry. But association of Cr with this factor can be stated as explained in SPM. The factor 4 with a percentage variance of 6, comprises of Na, which can be explained as sea-salt as a source; since sampling site is located near Mumbai harbour. The negative factor loading is an indication of opposite location of this as compared to other three sources with respect to wind direction i.e., when wind is coming from the direction of three sources Na is not reaching the sampling site and vice versa.

Source apportionment to the 72 hr. average PM<sub>2.5</sub> is shown in Table 32 for the sampling location. For the 72 hr. average PM<sub>2.5</sub> concentration, 30% is contributed by high temperature combustion. 57% is contributed by soil or cement dust. 3.4% metal industry, 1.7% from marine aerosols and 7.9% remains as unexplained contribution.

**Table 3.4: Factor Loading Matrix of PM<sub>2.5</sub> after Varimax Rotation**

	Factor 1	Factor 2	Factor 3	Factor 4
Mass	<b>0.68</b>	0.32	0.42	0.31
Cd	0.35	0.43	0.32	<b>0.62</b>
Pb	<b>0.59</b>	0.35	0.42	0.41
Cu	<b>0.62</b>	0.41	0.36	0.39
Mn	<b>0.76</b>	0.14	0.46	0.11
Co	-0.27	0.21	<b>-0.70</b>	-0.01
Cr	0.07	0.42	<b>0.71</b>	0.23
Fe	<b>0.89</b>	0.04	0.08	-0.05
Zn	0.56	0.29	0.49	0.35
Ca	0.25	<b>0.80</b>	-0.02	-0.32
Mg	0.11	<b>0.82</b>	0.13	0.12
K	<b>0.57</b>	0.50	-0.11	0.27
Na	-0.09	0.17	-0.08	<b>-0.89</b>
<i>Eigen Value</i>	1.5	0.61	0.07	0.81
<b>Percent of Variance</b>	50	12.4	8.3	6.2
<b>Cumulative Percentage</b>	50	62.4	70.7	76.9

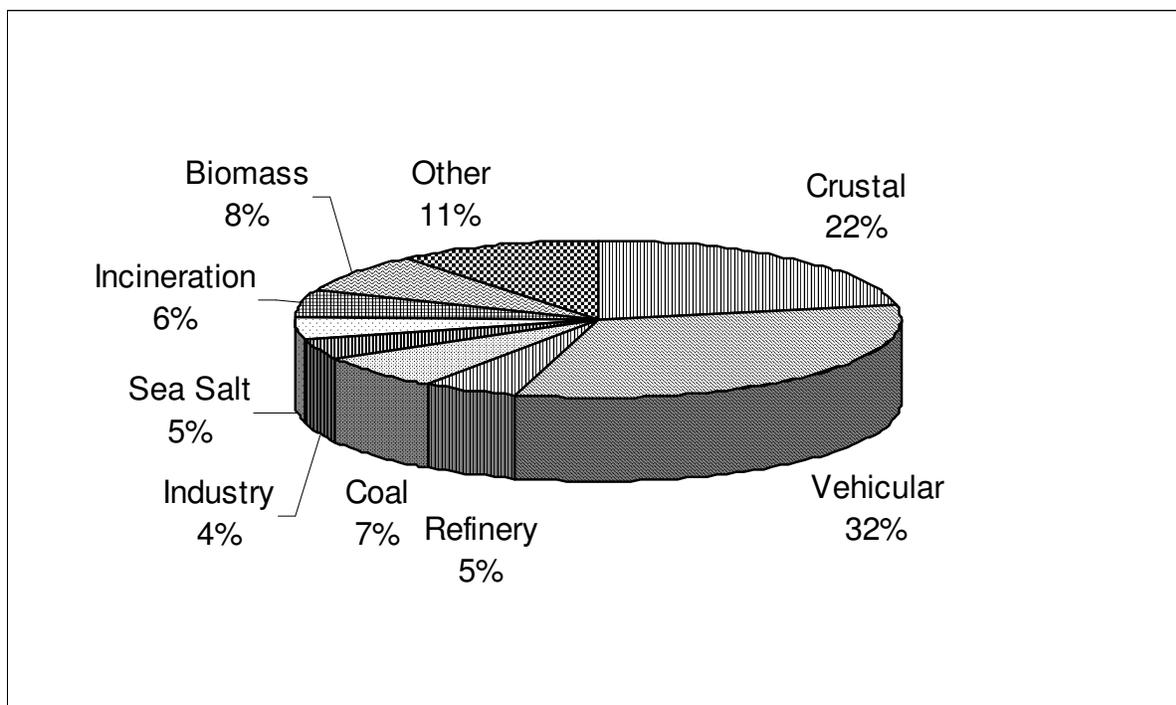
### 3.3.4 Source Apportionment of Respirable Suspended Particulate Matter:

Varimax rotated factor analysis was applied to PM<sub>10</sub> dataset, yielding the components represented in Table 3.5. For PM<sub>10</sub> nine components have eigenvalue greater than one and accounted 79 % of the variance in the data set.

In case of PM<sub>10</sub> samples, the first factor explains about 23 % of the variance with a high loading for Benzo(k)Fluranthene, Benzo(b)Fluranthene, Phenanthrene, Perylene, EC and OC. According to Hwang et al. (2003) the presence of PAHs with 3 and 4 aromatic rings, such as Benzo(k)Fluranthene, Benzo(b)Fluranthene and Phenanthrene, indicates the dominance of diesel combustion. Ogulei et al. (2005) has pointed out that a diesel source is generally characterized by relatively higher concentrations of EC than OC. Therefore, the presence of these compounds can be attributed to diesel exhausts. The second factor explains 11% of the variance with high loads in crustal components (K, Si, Pb, Mn and Ca). The crustal load may comprise both natural resuspension of soil particles and anthropogenic road dust. The third factor explains 9% of the variance with high loads in Benzo(a)Anthracene, Benzo(a)Pyrene K and Ca. Kulkarni and Venkataraman (2000) have shown that Benzo(a)Pyrene can be a good tracer for wood burning while Freeman and Cattell (1990) showed that this compound may be associated with combustion of vegetation. Watson and Chow (2001) found that biomass burning is characterized by high loading of K. Therefore, it is more appropriate to attribute this factor to biomass burning. The fourth factor has the characteristic of having high loading for anthracene, phenanthrene and benzo(a)anthracene. This factor is responsible for 8 % of the total variance. Khalili et al. (1995) identified that Anthracene, Phenanthrene and Benzo(a)Anthracene were tracer for coal combustion. Therefore coal combustion can be characterized as the source for this factor. The fifth factor explains 7 % of the variance with high loads in lower molecular PAHs (acenaphthalene, fluorine, fluoranthene and pyrene). Yang and Chen (2004) concluded that the refinery emission is characterized by acenaphthalene, fluorine, fluoranthene and pyrene. Therefore, the presence of these compounds can be attributed to refinery emissions. The sixth factor is responsible for 5 % of the total variance. This factor is weighted in Benzo(ghi)Perylene and Fluranthene. Miguel et al. (1998) have shown that Benzo(ghi)Perylene is characteristic of vehicular emissions from gasoline engines. Therefore this factor can be attributed to the

vehicular emissions for gasoline exhausts. The seventh factor explains 4 % of the variance with high loads in dibenzo(a,h)anthracene and anthracene. Masclet et. al. (1986) has shown that Dibenzo(a,h)Anthracene can be a typical tracer for incineration. Therefore this factor can be attributed to the incineration activities. The eighth factor is having high load for Cu, Fe and Mn. This factor in general named as industry which included contributions from different metallurgical activities. The last factor has the characteristics of having high Ca, Na and Mg suggesting that this factor may be attributed to sea salt.

The multiple linear regressions were performed to determine the percent contribution of different sources (Fig. 3.4). For  $PM_{10}$  the mean percent contribution from the diesel exhausts was found to be 20.6 % while that from gasoline exhausts was 11.5 %. So, the vehicular exhaust contributed almost 32 % at the sampling site. Crustal source contribute 22.3 % of the  $PM_{10}$ . Biomass burning and coal combustion contributed 8.1 % and 7.2 % respectively. 5.0 % of the  $PM_{10}$  were coming from refinery sources. The incineration activities contributed 5.7 % to the  $PM_{10}$  while the industry contributed 4.1 %. The sea salt contribution for  $PM_{10}$  is 4.8 %. The remaining 10.7 % is contributed by other sources.



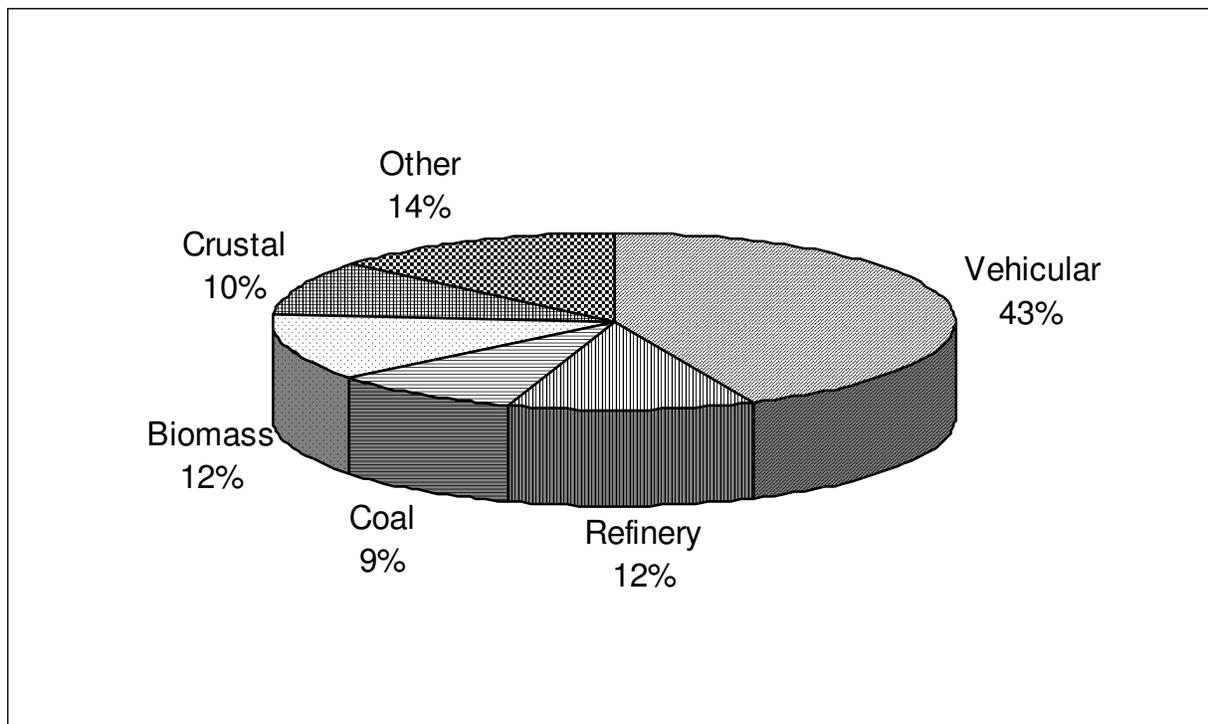
**Fig 3.4 Percentage Contribution from Different Sources to  $PM_{10}$  Mass**

### 3.3.5 Source Apportionment of Fine Suspended Particulate Matter:

Varimax rotated factor analysis was applied to PM<sub>2.5</sub> dataset, yielding the components represented in Table 3.6. For PM<sub>2.5</sub> six components have eigenvalue greater than one and accounted 76 % of the variance in the data set.

In case of PM<sub>2.5</sub> the factor analysis resolved six sources contributing to the particulate matter samples. The first factor is highly loaded with Benzo(k)Fluranthene, Benzo(b)Fluranthene, Phenanthrene, Perylene, Dibenzo(a,h)Anthracene and EC. This factor is attributed to diesel exhausts. The second factor has characteristics of having high loading in lower molecular weight PAHs, and therefore assigned as refinery source. The third factor is having high K, Ca, OC, Fluoranthene and Benzo(ghi)Perylene. This factor can be assign as biomass burning. The fourth factor is attributed to coal combustion sinceit has high loading for Anthracene, Phenanthrene and Benzo(a)Anthracene. The fifth factor is highly loaded with crustal components (K, Ca, Si, Mn, Mg, Pb and Na). The last factor is characterized as gasoline exhaust as it has high loading for Pyrene, Perylene and Benzo(ghi)Perylene.

For PM<sub>2.5</sub> (Fig. 3.5) the mean percent contribution from the diesel exhausts was found to be 28.9 % while that from gasoline exhausts was 14.3 %. So, the vehicular exhaust contributed almost 43 % at the sampling site. Crustal source contribute 9.6 % of the PM<sub>2.5</sub>. Biomass burning and coal combustion contributed 12.2 % and 9.4 % respectively. 11.7 % of the PM<sub>2.5</sub> was coming from refinery sources. The remaining 13.9 % is contributed by other sources.



**Fig 3.5 Percentage Contribution from Different Sources to PM<sub>2.5</sub> Mass**

The receptor modeling results clearly indicated that there is a decreased in contribution from crustal source in fine fraction of particulate matter (PM<sub>2.5</sub>) in comparison to coarse fraction (PM<sub>10</sub>). The contribution of sea salt for PM<sub>2.5</sub> is very negligible as compare to that for PM<sub>10</sub>. Thus, the sources contributing for PM<sub>2.5</sub> are mainly anthropogenic in nature whereas in case of PM<sub>10</sub> natural sources contribute significantly along with anthropogenic activities. From the source pattern it is clearly understood that organic pollutants play an important role in fine fraction of particulate matter (PM<sub>2.5</sub>).

Table 3.5: Varimax rotated Factor analysis of PM<sub>10</sub> data

Pollutants	Factor-1	Factor-2	Factor-3	Factor-4	Factor-5	Factor-6	Factor-7	Factor-8	Factor-9
Cu	0.12	0.22	0.36	0.21	0.12	0.27	0.37	<b>0.68</b>	0.22
K	0.32	<b>0.55</b>	<b>0.81</b>	-0.41	0.12	0.20	0.31	0.34	0.11
Fe	0.19	-0.13	-0.12	0.10	0.15	0.11	0.18	<b>0.50</b>	0.34
Mn	0.21	<b>0.60</b>	-0.23	0.32	0.45	0.27	0.18	<b>0.72</b>	0.14
Si	0.26	<b>0.86</b>	0.25	0.18	-0.41	0.13	-0.35	0.37	-0.18
Pb	0.35	<b>0.52</b>	0.18	-0.12	0.24	0.37	0.18	-0.12	0.29
Na	0.11	0.24	-0.16	0.27	0.12	0.19	0.27	0.18	<b>0.89</b>
Mg	0.14	-0.31	0.12	0.42	0.05	0.13	-0.08	0.14	<b>0.63</b>
Ca	0.29	<b>0.76</b>	<b>0.60</b>	0.37	-0.06	0.09	0.02	0.17	<b>0.57</b>
EC	<b>0.86</b>	0.16	0.30	0.19	0.42	-0.09	0.08	0.27	0.31
OC	<b>0.80</b>	0.08	0.43	-0.27	0.05	0.16	0.40	-0.37	0.21
AcNap	0.14	0.11	0.31	0.24	<b>0.84</b>	0.15	0.04	0.24	-0.38
Flu	0.26	-0.27	0.14	0.15	<b>0.73</b>	0.09	0.18	0.43	0.07
Phe	<b>0.74</b>	0.16	0.18	<b>0.52</b>	0.02	0.18	0.37	0.28	-0.05
Ant	-0.15	0.14	0.20	<b>0.87</b>	0.42	-0.05	<b>0.52</b>	0.15	0.21
Flt	0.26	0.28	0.02	0.32	<b>0.68</b>	<b>0.63</b>	-0.03	0.14	0.27
Pyr	0.13	0.24	0.27	0.18	<b>0.72</b>	-0.08	0.16	0.32	0.10
BaA	0.28	0.31	<b>0.71</b>	<b>0.63</b>	0.07	0.39	0.12	0.17	0.22
Pery	<b>0.92</b>	0.20	0.14	0.17	0.29	-0.06	0.37	0.15	-0.10
BbF	<b>0.50</b>	0.18	0.31	0.19	0.12	0.05	0.30	0.31	0.27
BkF	<b>0.78</b>	0.10	-0.40	0.24	0.08	0.34	-0.26	-0.45	0.01
BaP	0.14	0.16	<b>0.83</b>	-0.16	0.10	-0.42	0.22	0.37	0.14
DiahA	0.21	-0.22	0.27	0.37	-0.17	0.11	<b>0.76</b>	0.40	-0.05
BghiP	0.43	0.30	0.22	0.25	0.10	<b>0.76</b>	0.17	0.13	0.24
% Varance	<b>23.73</b>	<b>11.70</b>	<b>9.59</b>	<b>8.44</b>	<b>7.08</b>	<b>5.77</b>	<b>4.41</b>	<b>4.30</b>	<b>3.92</b>
Origin	Diesel exhaust	Crustal	Biomass burning	Coal	Refinery	Gasoline exhaust	Incineration	Industry	Sea Salt

Table 3.6: Varimax rotated Factor analysis of PM<sub>2.5</sub>data

Pollutants	Factor-1	Factor-2	Factor-3	Factor-4	Factor-5	Factor-6
Cu	0.02	0.08	0.07	0.27	-0.02	0.09
K	0.12	-0.18	<b>0.67</b>	-0.04	<b>0.59</b>	0.34
Fe	0.21	0.16	-0.04	0.19	-0.36	-0.21
Mn	-0.34	-0.05	0.14	0.22	<b>0.86</b>	0.07
Si	0.26	0.41	0.31	0.10	<b>0.72</b>	0.15
Pb	-0.05	0.31	-0.42	0.02	<b>0.64</b>	-0.27
Na	0.18	0.22	0.06	-0.37	<b>0.58</b>	0.11
Mg	0.27	0.04	0.17	0.40	<b>0.61</b>	0.18
Ca	0.20	0.26	<b>0.76</b>	0.07	<b>0.68</b>	0.24
EC	<b>0.68</b>	0.34	-0.09	0.15	0.02	-0.09
OC	0.06	0.15	<b>0.63</b>	0.17	0.14	0.06
AcNap	0.10	<b>0.73</b>	0.10	0.04	0.10	0.25
Flu	-0.06	<b>0.64</b>	0.18	-0.05	-0.07	0.01
Phe	<b>0.81</b>	-0.08	0.12	<b>0.59</b>	0.21	0.21
Ant	-0.05	0.30	-0.20	<b>0.76</b>	0.40	0.18
Flt	0.20	0.20	<b>0.61</b>	0.17	0.28	-0.40
Pyr	0.03	<b>0.81</b>	0.21	0.11	0.20	<b>0.65</b>
BaA	0.08	0.13	0.10	<b>0.74</b>	0.01	0.29
Pery	<b>0.74</b>	-0.09	-0.04	0.07	0.25	<b>0.56</b>
BbF	<b>0.61</b>	0.10	0.04	0.10	0.11	0.31
BkF	<b>0.69</b>	0.28	0.37	0.04	-0.31	0.06
BaP	0.14	-0.18	0.23	0.12	0.17	0.37
DiahA	<b>0.76</b>	0.27	0.20	-0.27	0.02	-0.14
BghiP	0.33	0.40	<b>0.83</b>	0.15	0.34	<b>0.84</b>
% Varance	<b>26.67</b>	<b>16.31</b>	<b>11.84</b>	<b>8.56</b>	<b>7.08</b>	<b>5.37</b>
Origin	Diesel exhaust	Refinery	Biomass burning	Coal	Crustal	Gasoline exhaust

## Chapter-4

### References

1. Alpert D.J., and Hopke P.K. (1980) a quantitative determination of the sources in the Boston urban aerosol. *Atmospheric Environment*, 14, 1137-1146.
2. Bilfford I.H., and Meeker G.O. (1967) A factor analysis model of large scale pollution. *Atmospheric Environment*, 1, 147-157.
3. Bogen J. (1973) Trace elements in the atmospheric aerosol in the Heidelberg area, measured by instrumental neutron activation analysis. *Atmospheric Environment*, 7, 1117.
4. Cass G.R. (1981) Sulfur air quality control strategy design. *Atmospheric Environment*, 15, 1227.
5. Cattell R.B. (1966) *Handbook of Multivariate Experimental Psychology*. Rand McNally, Chicago, pp. 174-243.
6. Chiapello I., Bergametti G., Chatenet B., Bousquet P., Dulac F., and Santos Soares E. (1997) Origins of African dust transported over the northeastern tropical Atlantic. *Journal of Geophysical Research*, 102, 13701-13709.
7. D'Abreton P.C., and Tyson P.D. (1996) Three-dimensional kinematic trajectory modeling of water vapour transported over South Africa. *Water SA*, 22, 279-306.
8. Deshpande V.A. and Pandit V.I. (2001); Use of PCA- A prerequisite to multiple regression analysis for ascertaining independent variables conditions. *Indian Journal of Environmental protection*, 21(8), 696-702.
9. Friedlander S.K. (1973) Chemical element balances and identification of air pollution sources. *Envir. Sci. Technol.*, 7, 235-240.
10. Freeman, D.J. and Cattell, C.R. (1990) Woodburning as a source of atmospheric polycyclic aromatic hydrocarbons in urban air of Hon King. *Atmos. Environ.*, 24, 1581-1585.

11. Gartrell G., and Friedlander S.K., (1975) Relating particulate pollution to sources: the 1972 California aerosol characterization study. *Atmospheric Environment*, 9, 279-299.
12. Gatz D.F. (1975) Relative contributions of different sources of urban aerosols: application of a new estimation method to multiple sites in Chicago. *Atmospheric Environment*, 9, 1-18.
13. Harman H.H. (1976) *Modern Factor Analysis*. (3<sup>rd</sup> edition), University of Chicago press,
14. Heindryckx R., and Dams R. (1974) Continental, marine, and anthropogenic contributions to the inorganic composition of the aerosol of an individual zone. *J. Radioanal. Chem.*, 19, 339-343.
15. Henry R.C, and Hidy G.M. (1979) Multivariate analysis of particulate sulfate and other air quality variables by principal components-Part I, Annual data from Los Angeles and New York. *Atmospheric Environment*, 13, 1581-1596.
16. Henry R.C., Lewis C.W., Hopke, P.K., and Williamson H.J. (1984) Review of receptor model fundamentals. *Atmospheric Environment*, 18, 1507-1515.
17. Henry R.C., Wang Y.J., and Gebhart K.A. (1991a) The relationship between empirical orthogonal functions and sources of air pollution. *Atmospheric Environment*, 25A, 503-509.
18. Henry R.C. (1991b) Multivariate Receptor Models, in *Receptor modeling for Air Quality Management*, P.K. Hopke (ed.), Elsevier Science Publishers, Amsterdam, pp. 117-147.
19. Hopke P.K., Gladney E.S., Gordon G.E., Zoller W.H., and Jones A.G. (1976) The use of multivariate analysis to identify sources of selected elements in the Boston urban aerosol. *Atmospheric Environment*, 10, 1015-1021.
20. Hopke P.K., Lamb R.E., and Natusch D.F.S. (1980) Multielemental characterization of urban road ways dust. *Envir. Sci. Technol.*, 14, 164-172.
21. Hwang, H.M., Wade, T.L., Sericano, J.L. (2003) Concentration and source characterization of polycyclic aromatic hydrocarbons in pine needles from Korea, Mexico and United States. *Atmos. Environ.*, 37, 2259-2267.
22. Kaiser H.F., and Hunka S. (1973) *Edu. Psych. Measurement*, 33, 99.
23. 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. *Atmos. Environ.*, 29, 533-542.

24. Kowalczyk G.S., Choquette C.E., and Gordon G.E. (1978) Chemical element balances and identification of air pollution sources in Washington. *Atmospheric Environment*, 12, 1143.
25. Kulkarni, P. and Venkataraman, C. (2000) Atmospheric polycyclic aromatic hydrocarbons in Mumbai, India. *Atmos. Environ.*, 34, 2785-2790.
26. Kulkarni P.S. (1997); Chemical mass balance for source apportionment of particulate polycyclic aromatic hydrocarbons in Indian Cities. M. Tech. Dissertation, IIT, Bombay.
27. Masclat, P., Nikolaou, K. and Mouvier, G. (1986) Relative decay index and sources of Polycyclic aromatic hydrocarbons. *Atmos. Environ.*, 20, 439-446.
28. Mayrshon H., and Crabtree J.H. (1976) Source reconciliation of atmospheric hydrocarbons. *Atmospheric Environment*, 10, 137-143.
29. Meenakhy V., Negi B.S., and Sadasivan S. (1996) PM-10 aerosols at chembur, Bombay-their elemental concentrations and sources. 5<sup>th</sup> National Symposium on Environment, Calcutta, 227-230.
30. Miguel, A.H. and Pereira, P.A.P. (1998) Benzo(k)fluoranthene, benzo(ghi)perylene and indeno(1,2,3-cd)pyrene: new tracer of automotive emissions in receptor modeling. *Aerosol Science Technology*, 10, 292-295.
31. Miller, M. S., Friedlander S.K., and Hidy G.M. (1972) A chemical element balance for the Pasadena aerosol. *J. Colloid Interface Sci.*, 39, 165-176
32. Negi B.S., Sadasivan S., and Mishra U.C. (1987); Aerosol composition and sources in urban area in India. *Atmospheric Environment*, 21,1259-1266.
33. Negi B.S., Sadasivan S., and Mishra U.C. (1988); Factor Analysis in the interpretation of aerosol composition data. *Indian Journal of Environmental Health*, 31(1), 32-42.
34. Paatero P. and Tapper U. (1994) Positive Matrix Factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5, 111-126.
35. Paatero P. (1996) User's Guide for PMF programs PMF2.EXE and PMF3.EXE. University of Helsinki, Helsinki.
36. Paatero P. (1997) Least square formulation of robust non-negative factor analysis. *Chemometrics and Intelligent Laboratory Systems*, 37, 23-35.

37. Pandit G.G. (1990); Source Reconciliation of Atmospheric Hydrocarbons. Seminar on Factor Analysis and Its Applications to Environmental Assessment, BARC, Bombay, September 24.
38. Peterson J.T. (1970) distribution of sulfur dioxide over metropolitan St. Louis, as described by empirical eigen vectors, and its relation to meteorological parameters. *Atmospheric Environment*, 4, 501-518.
39. Poirot R.L., and Wishinski P.R. (1986) Visibility, Sulfate and air mass history associated with the summertime aerosol in Northern Vermont. *Atmospheric Environment*, 20, 1457-1469.
40. Sadasivan S., Negi B.S., and Mishra U.C. (1984) Composition and sources of aerosols at Trombay, Bombay. *The Science of Total Environment*, 40, 279-286.
41. Sharma V.K., and Patil R.S. (1992a) Size distribution of atmospheric aerosols and their source identification, using factor analysis in Bombay, India. *Atmospheric Environment*, 26B, 135-140.
42. Sharma V.K., and Patil R.S. (1992b); Receptor models for source apportionment of aerosols. National Symposium on Environment, BARC, Bombay, Feb. 3-5, 138-141.
43. Sharma V.K., and Patil R.S. (1994) Chemical Mass Balance model for source apportionment of aerosols in Bombay. *Environmental Monitoring and Assessment*, 29, 75-88.
44. Steinacker R., (1984) Airmass and frontal movement around the Alps. *Rivista di Meteorologia Aeronautica*, 44, 85-93.
45. Stohl A. (1996) Trajectory statistics- a new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *Atmospheric Environment*, 30, 579-587.
46. Venkataraman C., Reddy C.K., Josson S., and Reddy M.S. (2002) Aerosol size and chemical characteristics at Mumbai, India, during the INDOEX-IFP (1999). *Atmospheric Environment*, 36, 1979-1991.
47. Vinod Kumar A., Patil R.S., and Nambi K.S.V. (2001) Source apportionment of suspended particulate matter at two traffic junctions in Mumbai, India. *Atmospheric Environment*, 35, 4245-4251.

48. Winchester J.W., and Nifong G.D. (1971) Water pollution in lake Michigan by trace elements from pollution aerosol fallout. *Water Air Soil Pollut.*, 1, 50-64.
49. Zeng Y., and Hopke P.K. (1989) A study of sources of acid precipitation in Ontario, Canada. *Atmospheric Environment*, 23, 1499-1509.