Final Report

REVIEW OF SOURCE APPORTIONMENT TECHNIQUES FOR AIRBORNE PARTICULATE MATTER

Prepared for

California Air Resources Board
and the
California Environmental Protection Agency

Contract #00-332

Prepared by

Michael J. Kleeman
University of California, Davis
Davis, CA 95616

February 2003

Disclaimer

Acknowledgements

This research described in this final report was conducted as a part of the PM assessment by the North American Research Strategy for Tropospheric Ozone (NARSTO).

This report was submitted in fulfillment of 00-332 “Review of Source Apportionment Techniques for Airborne Particulate Matter” by Michael J. Kleeman under the sponsorship of the California Air Resources Board. Work was completed as of date December 2002.

Table of Contents

Disclaimer ii
Acknowledgements iii
Table of Contents iv
Abstract v
Executive Summary vi
  Background vi
  Methods vi
  Results vi
  Conclusions vii
Introduction 1
  Scope and Purpose 1
  Background 2
Materials and Methods 3
Results 4
  Unique Chemical Tracer Analysis 4
  Statistical Analysis Using Chemical Profiles 8
    Linear Statistical Models 10
    Bi-Linear Statistical Models 13
    Hybrid Statistical Models 22
  Mechanistic Analysis Techniques 23
    Emissions Inventory Analysis 24
    Source-Oriented Chemical Transport Models 25
Discussion 32
Summary and Conclusions 34
Recommendations 35
References 36
List of Inventions Reported and Publications Produced 51
Glossary of Terms, Abbreviations, and Symbols 52

Abstract

A review of source apportionment studies for airborne particulate matter reveals that unique chemical tracers are only available for a few general source categories including wood combustion, motor vehicle exhaust, coal combustion, meat cooking, crustaceous seafood cooking, candle combustion, tobacco smoke, and leaf abrasion. Unique chemical profiles can be constructed for many more sources, but their interpretation requires the uses of complex source apportionment models. Ten models that calculate the contribution that different sources make to air borne particulate matter are reviewed: Tracer Analysis, Chemical Mass Balance, Principal Component Analysis, Factor Analysis, Positive Matrix Factorization, Empirical Orthogonal Functions, Chemical Mass Balance / Factor Analysis, Target Transformation Factor Analysis, Emissions Inventory Analysis, and Source-oriented Chemical Transport Models. The mathematical basis, assumptions, and limitations for each source apportionment model are summarized. The source apportionment methods are shown to provide a continuous spectrum of capabilities with increasingly detailed information provided as the amount of input data expands.

A protocol to characterize source contributions to primary and secondary airborne particulate matter concentrations in a region of interest is recommended. Ambient data sets can first be analyzed using bilinear statistical models to identify approximate source identities and contributions. After important source profiles have been measured, linear statistical models can be applied to calculate source contributions to primary particulate matter at receptor locations. If regional source contributions to airborne particulate matter and / or source contributions to secondary particulate matter are important, emissions inventories can be developed for the region of interest and source-oriented chemical transport models can be used to identify important source contributions to these features of airborne particulate matter.

Executive Summary

Background

Airborne particles in typical urban areas are a complex mixture of inorganic and organic compounds that exist in either the solid or liquid state. Traditional strategies to reduce PM2.5 concentrations are based on a reduction of the controllable emissions of the compounds that form airborne particles. The most effective set of emissions controls is not obvious, however, since a variety of sources release fine primary particulate matter to the atmosphere including combustion sources (stationary and mobile), food preparation, activities that create dust (road travel, agriculture), and natural sources such as sea spray. Secondary particulate matter can also form in the atmosphere through chemical reactions that convert gaseous pollutants (NOx, SOx, VOC) to semi-volatile products that partition to the particle phase.

The California Air Resources Board must develop emissions control strategies that reduce the potentially harmful effects of particulate air pollution with a minimum impact on California's economy. A wide variety of source apportionment tools have been developed to aid in the design of emissions control strategies. The purpose of this review is to compare the mathematical basis of the state-of-the-art source apportionment techniques, discuss their application to different air quality problems, and determine how they can best be used in the design of emissions control strategies. A protocol for the source apportionment of airborne particulate matter using various methods then will be recommended.

Methods

A literature review of the chemical compounds typically employed by statistical source apportionment techniques was compiled to show how some sources contributions could be identified by unique chemical tracers while other sources required the application of statistical or mechanistic models. The mathematical basis for ten (10) different source apportionment techniques was summarized to illustrate similarities and differences. A literature review of the information revealed by different modeling approaches then was used to determine the type of information that each source apportionment method provides.

Results

Unique chemical tracers were identified in the literature for wood combustion, motor vehicle exhaust, coal combustion, meat cooking, crustaceous seafood cooking, candle combustion, tobacco smoke, and leaf abrasion. It was also found that unique chemical profiles could be constructed for many other sources of interest even if unique chemical tracers were not available. The various source apportionment techniques that were reviewed provide a continuous spectrum of capabilities with increasingly detailed information provided as the amount of input data expands. The statistical models that require the least amount of information require ambient measurements of the chemical composition of airborne particulate matter to provide an estimation of the number of important sources present in a region and the approximate chemical composition of particles released from those sources. The next level of statistical models require detailed information about the chemical composition of particles released from all the important sources in a region to provide a more accurate estimate of the contribution that each source makes to airborne particulate matter concentrations at specific receptor locations. Finally, mechanistic emissions inventories and meteorological data can support the application of source-oriented chemical transport models that can predict source contributions to the regional concentrations of primary and secondary particulate matter.

Conclusions

The design of emissions control strategies for airborne particulate matter can be best accomplished through the use of multiple source apportionment models applied in several phases. During the initial assessment of particulate matter in a region, bilinear statistical models can identify important sources that should be characterized by detailed testing. During the second stage of the assessment, linear statistical models can accurately calculate source contributions to primary particulate matter at specific receptor sites. During the final stage of the assessment, source-oriented chemical transport models can calculate source contributions to the regional distribution of primary and secondary particulate matter.

Introduction

There are now over 150 published articles that describe epidemiological studies considering the health effects of inhaled airborne particles (Pope, 2000). The general consensus among the majority of the epidemiological evidence indicates that particulate air pollution is an important risk factor for increased cardio-pulmonary disease and mortality. The United States Environmental Protection Agency (US EPA) has recently implemented a National Ambient Air Quality Standard (NAAQS) for PM2.5 in an attempt to reduce potential adverse health effects. Preliminary data indicate that large portions of the Western and the Eastern United States have ambient PM2.5 concentrations that exceed the new PM2.5 NAAQS resulting in unhealthful conditions for a significant fraction of the population in North America. Air quality managers at the national, state, and local levels will soon need to characterize their particulate air quality problems and devise a strategy to reduce PM2.5 concentrations to comply with new standards.

Airborne particles in typical urban areas are a complex mixture of inorganic and organic compounds that exist in either the solid or liquid state. Traditional strategies to reduce PM2.5 concentrations are based on a reduction of the controllable emissions of the compounds that form airborne particles. The most effective set of emissions controls is not obvious, however, since a variety of sources release fine primary particulate matter to the atmosphere including combustion sources (stationary and mobile), food preparation, activities that create dust (road travel, agriculture), and natural sources such as sea spray. Secondary particulate matter can also form in the atmosphere through chemical reactions that convert gaseous pollutants (NOx, SOx, VOC) to semi-volatile products that partition to the particle phase. Furthermore, airborne particles with aerodynamic diameter less than 2.5 μm can stay suspended in the atmosphere for long periods of time. As a result, some areas downwind of major urban centers may not be able to directly control the emissions that make the greatest contribution their particulate air quality problems. Faced with this complexity, decision makers need a set of tools that clearly show the relationship between emissions sources and airborne particle concentrations.

Scope and Purpose

Source apportionment techniques for airborne particulate matter are generally defined as any method that quantifies the contribution that different sources make to airborne particulate matter concentrations at receptor locations in the atmosphere. Source apportionment techniques are valuable tools that aid in the design of effective emissions control programs to reduce particulate air pollution. A comprehensive review and comparison of different source apportionment techniques for airborne particulate matter has been undertaken previously (Stevens et al., 1984; Henry et al., 1984; Currie et al., 1984; Dzubay et al., 1984; Gordon et al., 1984). Since the time of this initial review several new source apportionment techniques for airborne particulate matter have been developed and additional experience has been gained applying traditional source apportionment techniques to new locations.

Source apportionment techniques for airborne particulate matter define the total contribution that different sources make to airborne particle concentrations at a receptor site. These techniques are distinct from a set of closely related source perturbation methods that predict the change in airborne particle concentrations that would result from a change in emissions associated with different source categories. In mathematical terms, source perturbation techniques attempt to identify the local derivative of source contributions to airborne particulate matter while source apportionment techniques integrate this derivative to find the absolute source contribution. Each technique provides useful information during the design of emissions control programs. In the present study, only techniques that provide an absolute source apportionment of airborne particulate matter will be considered.

The purpose of the current study is to review the mathematical basis of the state-of-the-art source apportionment techniques, discuss their application to different air quality problems, and compare their strengths and weaknesses. A protocol for the source apportionment of airborne particulate matter using various methods then will be recommended.

Background

Source apportionment techniques for airborne particulate matter can be carried out in response to a variety of problems (health impacts, acid deposition, visibility impairment, climate change, etc.) The spatial and temporal patterns of particulate matter concentrations that define the problem of interest must be measured using an appropriate sampling network before any study for airborne particulate matter can be carried out. Spatial variability can occur on scales ranging from a few meters (e.g. roadside dispersion) to thousands of miles (e.g. climate change). Temporal patterns are linked to spatial patterns through transport phenomena, and so temporal scales can range from a few seconds to many years. The size distribution of airborne particulate matter may also be an important consideration for the problem of interest. Currently regulated size fractions of airborne particulate matter in the United States include PM10 (particle aerodynamic diameter less than 10 μm) and PM2.5 (particle aerodynamic diameter less than 2.5 μm). The particle size distribution is a continuum in which different particle sizes have different environmental impacts, and so the size fraction relevant to the problem of interest must be characterized before source apportionment studies can be undertaken. Particles at any given size are usually made up of many different chemical compounds (ionic species, organic species, metals, etc) that can be detected and quantified using a variety of different techniques (IC, FID, AA, colorimeter, GC/MS, LC/MS, ICP/MS, ATOFMS, etc.). The exact nature of the problem of interest determines the nature of particulate matter characterization that must be carried out in support of a source apportionment project. In the current study, it is assumed that an appropriate characterization of airborne particulate matter has been completed so that the relative merits of different source apportionment techniques can be examined.

Materials and Methods

A literature survey was undertaken to identify previous source apportionment studies for airborne particulate matter. The results of this survey were used to compile a list of chemical tracers and source apportionment models used to identify source contributions. The mathematical basis for each source apportionment model was summarized to highlight similarities and differences. The strengths and weaknesses of each model then were determined based on the results of the studies where they were employed.

Results

Unique Chemical Tracer Analysis

Unique chemical tracers provide the simplest method for the quantification of source contributions to PM2.5 concentrations. Assuming that the tracer species (gas or particle phase) have the same atmospheric lifetime as the particle size fraction of interest, then the contribution that each source makes to airborne particulate matter concentrations can be calculated using the equation:

$$G_k = \frac{X_j}{F_{kj}} \qquad (1)$$

where G_k is the concentration (μg m^{-3}) of airborne particulate matter at the receptor site from source k, X_j is the atmospheric concentration (μg m^{-3}) of unique chemical tracer j, and F_{kj} is the ratio of unique chemical tracer j to total airborne particulate matter released from source k. Note tracers should only be used in equation (1) to calculate source contributions to an airborne particle size fraction if the atmospheric removal rate of the tracer species is identical to the atmospheric removal rate of particles in that size fraction. Transformation rates for some compounds in the atmosphere are significant and so they cannot be used as tracer species for airborne particulate matter.

Examples of unique chemical tracers that have been identified for different sources of airborne particulate matter are summarized in Table 1. The majority of the chemical compounds listed are trace organic species identified through the use of solvent extract and subsequent analysis by gas chromatograph / mass spectrometry (GC/MS). Excellent reviews of the identification of organic tracers for particulate air pollution sources are provided by Cass (1998) and Simoneit (2000). Of particular note are the unique organic tracers identified for the combustion of different subspecies of wood (Fine et al., 2001; 2002). Betulin, Juvabione, Dehydrojuvabion, and Dehydroabietic acid have been suggested as unique tracers for paper birch, balsam fir, and conifers respectively. This level of detail for subspecies within an individual emissions category will help identify exact sources of particulate air pollution and should assist in the design of emissions control programs.

Lead has traditionally been used as a marker for particulate matter released from the combustion of leaded gasoline. Since the conversion to non-leaded gasoline in much (but not all) of the world, ambient concentrations of particulate lead have decreased sharply and a number of smaller sources including waste incineration, copper smelting, and coal combustion now account for the residual concentrations (Chiaradia and Cupelin, 2000; Hamester et al., 1994; Mukai et al., 1993). Isotopic analysis of lead concentrations is generally a useful tool for differentiating between different lead sources in this case. Artificial isotopic ratios have also been used to identify other source contributions to particulate air pollution. In one study, artificially enriched isotopes of neodymium (Nd) were released from a coal-fired power plant and concentrations of airborne particulate matter were collected at downwind locations (Ondov et al., 1992).

Notably absent from the list of unique particulate matter tracers in Table 1 is potassium. Potassium is often present at high concentrations in particulate matter released from biomass combustion and is sometimes used as a unique tracer for that source. Studies show that sources such as meat charbroiling also release particulate matter that contains potassium (Schauer et al., 1999b) and so care must be exercised when assigning source contributions based on this component. Organic markers such as levoglucosan and methoxyphenols are only released from biomass combustion sources and so they are more suitable unique tracers for wood combustion.

The general classifications of Elemental Carbon (EC) and Organic Carbon (OC) have been used in the past as unique tracers for certain source types. Great care must be exercised when interpreting ambient measurements of these components for several reasons. The operational definitions of EC and OC are not standardized across the community. Generally speaking, EC refers to carbon that is not associated with a significant amount of oxygen. EC is typically non-volatile and has a dark black color as commonly found in airborne particulate matter. OC consists of carbon compounds that have greater amounts of oxygen and hydrogen. Compounds within the OC category are generally more volatile and lighter in color than EC compounds. The standard analysis technique for EC and OC measurements involves the quantification of carbon that is evolved from a sample as the temperature is increased in an inert or slightly oxidizing atmosphere. The analysis is complicated by the pyrolysis of some OC during the initial stages of thermal evolution to form less volatile compounds that can mistakenly be classified as EC. An optical technique has been employed to detect this phenomenon and correct for pyrolyzed carbon. The transmission or reflectance intensity of a laser beam incident on the sample is monitored during the analysis. The laser intensity decreases as the sample is heated due to the pyrolysis effect, and then increases during the final stage of the analysis as the less volatile carbon evolves at hotter temperatures in an oxidizing atmosphere. The transition point between OC and EC is taken to be the point at which the laser intensity is equal to the value of the original laser intensity of the sample. This definition for EC and OC is operational, since the temperature ramp used in the early parts of the analysis determines the amount of pyrolyzed carbon and the size of the associated pyrolysis correction (Chow et al., 2001). Protocols that use relatively low temperatures during the initial stages of analysis (see for example Chow et al., 1993) typically measure EC concentrations that are approximately double those measured using protocols that specify an temperature of 850C during the initial analysis stages (see for example NIOSH 1996, 1999). This trend is most evident during the analysis of particulate matter released from biomass combustion, and is most likely caused by the pyrolysis of organic compounds that do not thermally evolve at temperatures lower than 550C. Large amounts of pyrolyzed OC remain on the sample substrate until the final stage of analysis when temperatures are increased and a small amount of oxygen is introduced into the analysis chamber. Techniques that use a relatively high initial temperature (NIOSH, 1996; 1999) evolve the pyrolyzed carbon during the initial stages of analysis, reducing the need for pyrolysis corrections during the latter stages of analysis. Potentially significant sources of EC in the ambient atmosphere include diesel engines, wood smoke, and gasoline engines. Unless independent methods can establish the dominance of one of these sources in the region of interest, EC cannot be used as a unique tracer.

Traditional markers for fuel oil combustion such as nickel and vanadium are no longer considered unique tracers for these sources because combustion of heavy fuel oil has been discontinued in many communities. As a result, no single source dominates the residual emissions of these elements to the atmosphere.

A few studies have attempted to use the measured concentration and emissions ratio of gas-phase species to quantify the emissions of particulate matter from known sources. Naeher et al. (2001) used gas-phase CO concentrations to estimate PM2.5 concentrations released from open fires and improved wood cookstoves in Guatemalan homes. CO concentrations were found to correlate well with PM2.5 concentrations in the home under polluted conditions, but not under relatively clean conditions. Poor correlations were observed between CO and PM2.5 in homes using gas stoves. In general, gas-phase tracers can only be used to estimate source contributions to PM2.5 concentrations under well-characterized conditions where the removal rates of gas-phase tracers and PM2.5 can be verified as identical.

In general, the exact quantification of source contributions to airborne particulate matter concentrations using a single chemical tracer is difficult because any measurement error that affects the concentration of the tracer species in either the atmosphere (X_j) or at the source (F_{kj}) will skew the results of the calculation. The use of multiple tracers is recommended where feasible to reduce uncertainty.

Table 1: Unique tracers that can be used for source attribution of airborne particulate matter under appropriate conditions where minor sources are not significant.

Unique Tracer (Analysis Method)	**Source**	**Comment**
Lead (XRF, ICP/MS) Lead isotopes (XRF, ICP/MS)	Leaded Gasoline Combustion Waste incineration, copper smelting, coal combustion	Only viable in locations where leaded gasoline is still used. In areas where unleaded gasoline is used.
Levoglucosan Methoxyphenols Betulin Juvabione, Dehydrojuvabion Dehydroabietic acid (solvent extraction; GC/MS)	General Wood Combustion General Wood Combustion Paper Birch Balsam Fir Conifers	Emissions rate depends on combustion conditions.
High molecular weight, odd carbon number n-alkanes (solvent extraction, GC/MS)	Leaf Abrasion	
C31 Hopanes Divanillyl 1,2-divanillylethane (solvent extraction, GC/MS)	Coal	
Choloesterol Acyl Monoglycerides (solvent extraction; GC/MS)	Grilling/Charring Meat	Emission rate depends on the cooking method: charbroiling or frying.
1,6-anhydro-2-acetamido-2-deoxyglucose (solvent extraction, GC/MS)	Grilling/Charring Crustacean Seafood	
Wax esters (solvent extraction, GC/MS)	Beeswax Candle Combustion	
Hopanes and Steranes (solvent extraction; GC/MS)	Lubricating oil in Gasoline- and Diesel powered engines	Does not distinguish between gasoline and diesel vehicles.
Isoprenoids and Trycyclic Terpans (solvent extraction; GC/MS)	Gasoline-powered Motor Vehicle Exhaust	Trace amounts present in diesel vehicle exhaust.
Iso- and anteiso-alkanes (solvent extraction, GC/MS)	Tobacco Smoke	

Statistical Analysis Using Chemical Profiles

The power of source apportionment studies can be greatly increased through the simultaneous use of multiple tracers. Even when unique tracers for a given source do not exist, the ratio of tracers may form a unique chemical "profile" for that source. The list of chemical species and other aerosol properties that have been used to produce chemical profiles for airborne particulate matter is summarized in Table 2. The most common approach to constructing chemical profiles is to create a normalized profile by calculating the chemical composition of a unit mass of particulate matter. Another common approach is to divide the mass of trace organic species by the total concentration of organic compounds found in the particulate matter sample. Isotopic ratios of elements such as sulfur and lead also have been used to create chemical profiles for different sources.

Many of the unique chemical tracers identified in Table 1 also appear in Table 2 as a part of a chemical profile. Inclusion of these species increases the resolution of the statistical techniques used to perform source apportionment of chemical profiles. Many additional properties of the airborne particulate matter also contribute to profiles commonly used for source apportionment studies, including soluble ions, metals, EC, OC, individual organics, isotope ratios, particle shape, and particle magnetic properties. These properties are not unique for each source, but the relative ratios of the properties for each source are unique. The interpretation of source profile information is more complicated than the analysis techniques employed for unique tracers. Statistical models must be used to identify source contributions to airborne particulate matter based on source profiles.

Table 2: Chemical species and other aerosol properties used to produce chemical profiles for airborne particulate matter.

Chemical Species / Aerosol Property (analysis method)	**Emissions Source**	**Comment**
Soluble Ions (ion chromatography)	All anthropogenic sources	Some common ions (sulfate, nitrate, ammonium ion) are produced by gas-phase reactions.
Metals (ICP/MS, PIXE, or XRF)	All anthropogenic sources	
Elemental Carbon (EC) (thermal optical analysis)	Combustion Sources	Operational definition must be consistent within analysis.
Organic Compounds (OC) (thermal optical analysis)	Combustion Sources Crustal Sources	Operational definition must be consistent within analysis. General category that includes a wide range of organic compounds emitted by variety of sources.
Normal Alkanes (solvent extraction; GC/MS)	Meat Cooking Vegetable Frying Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines Wood Combustion Cigarette Smoke	Subset of compounds within OC.
Branched Alkanes (solvent extraction; GC/MS)	Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines Cigarette Smoke	Subset of compounds within OC.
Saturated Cycloalkanes (solvent extraction; GC/MS)	Diesel Engines Non-catalyst Gasoline Engines	Subset of compounds within OC.
Aromatic Hydrocarbons (solvent extraction; GC/MS)	Meat Cooking Vegetable Frying Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines Wood Combustion Aromatic Hydrocarbons	Subset of compounds within OC.
Phenols (solvent extraction; GC/MS)	Wood Combustion Cigarette Smoke	Subset of compounds within OC.
Aliphatic Aldehydes (solvent extraction; GC/MS)	Meat Cooking	Subset of compounds within OC.
Ketones	Meat Cooking	Subset of compounds within OC.

(solvent extraction; GC/MS)	Vegetable Frying Diesel Engines	
Unsaturated Carbonyls (solvent extraction; GC/MS)	Meat Cooking Vegetable Frying	Subset of compounds within OC.
n-Alkanoic Acids (solvent extraction; derivatization; GC/MS)	Meat Cooking Vegetable Frying Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines Wood Combustion Cigarette Smoke	Subset of compounds within OC.
n-Alkenoic Acids (solvent extraction; derivatization; GC/MS)	Vegetable Frying Wood Combustion Cigarette Smoke	Subset of compounds within OC.
Alkanedioic Acids (solvent extraction; derivatization; GC/MS)	Meat Cooking Vegetable Frying Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines	Subset of compounds within OC.
Aromatic Acids (solvent extraction; derivatization; GC/MS)	Diesel Engines Catalyst-equipped Gasoline Engines Non-catalyst Gasoline Engines	Subset of compounds within OC.
Furanones (solvent extraction; GC/MS)	Meat Cooking Vegetable Frying	Subset of compounds within OC.
^{207}Pb/^{206}Pb; ^{208}Pb/^{206}Pb; ^{207}Pb/^{204}Pb; ^{208}Pb/^{204}Pb (ICP/MS)	Combustion sources; crustal components	Isotope ratios emitted by different sources are distinct.
^{34}S/^{32}S (ICP/MS)	DMS oxidation; combustion sources	Isotope ratios emitted by different sources are distinct.
^{14}C/^{12}C (ICP/MS)	Combustion sources	Isotope ratios emitted by biogenic vs. fossil sources are distinct
Particle shape (optical imaging; TEM)		

Linear Statistical Models

Linear statistical models assume that the concentration of a chemical species at a receptor site is proportional to some linear combination of variables:

$$X_j = \sum_k G_k * F_{kj} + E_j \qquad (2)$$

where X_j is the concentration of chemical component j measured at a receptor site (μg m^{-3}), G_k is the total concentration of particulate matter released from source k at the receptor site (μg m^{-3}), F_{kj} is the fraction of chemical component j in the particulate matter emitted from source k, and E_j is the error in the model (μg m^{-3}). In order for this model to be considered linear, it must not involve products composed of unknown variables; thus, either G_k or F_{kj} must be known in Equation (2). Using matrix notation to write Equation (2) multiple times for different chemical components produces:

$$< X > = < G > [F] + < E > \qquad (3)$$

where $<X>$ denotes the row vector with entries for individual chemical components X_1 to X_p, $<G>$ is the row vector of individual source contributions G_1 to G_n, $[F]$ is the matrix whose rows $<f>$ contain the emissions profiles for each source, and $<E>$ is the row vector of residuals or errors E_1 to E_p for each chemical component in the linear model. If we have a dataset with multiple samples (perhaps a time series of measured particle composition at a receptor site), then a matrix of observed values and emissions strengths is produced:

$$[X] = [G][F] + [E] \qquad (4)$$

where each row of $[X]$, $[G]$, and $[E]$ corresponds to a single sample. The solution to this source apportionment problem statement is the value of [G] and/or [F] that minimizes [E].

The great advantage of the linear statistical models is simplicity: no information is required concerning meteorological patterns (wind speed, wind direction, temperature, relative humidity, atmospheric stability, etc.) or emissions patterns (source location, source diurnal profiles, etc.). There is often a great deal of effort associated with collecting or estimating each of these parameters and so a technique that can predict source contributions to airborne particulate matter concentrations without using this information has obvious benefits.

Chemical Mass Balance Model

Perhaps the most widely used source apportionment technique based on equation (3) is the Chemical Mass Balance (CMB) Model. The CMB model uses measured values of $<X>$ and [F] in Equation (3) to calculate the strength from different sources [G] such that the error $<E>$ is minimized. The original least squares weighted solution to this problem is:

$$< G > = < X > [W][F]\left([F]^T [W][F]\right)^{-1} \qquad (5)$$

where the superscript T indicates a matrix transpose, the superscript -1 indicates a matrix inverse, [W] is a diagonal matrix with elements defined by

$$w_{jj} = \frac{1}{\sigma_{Xj}^2} \qquad (6)$$

and σ_{Xj} is the uncertainty associated with ambient measurements for chemical species j (Henry et al., 1984). More recently, the original least squares technique has been modified so that the diagonal matrix [W] contains weighting factors chosen to be the inverse of the combined uncertainties of the source profile matrix [F] and the measured values [X] (Watson et al., 1984):

$$w_{jj} = \frac{1}{\sigma_{Xj}^2 + \sum_{k=1}^{q} \sigma_{Fjk}^2 g_k^2} \qquad (7)$$

where σ_{Fjk} is the uncertainty associated with source profile measurements for chemical species j emitted from source k, and g_k is the estimated source contribution to ambient particulate matter from source k. An iterative solution is used to solve implicit equations (5) and (7) for the source loadings (<G>).

There are several assumptions in the formulation of the CMB model that are worthy of note. The model is applicable only for compounds that are not preferentially produced or degraded in the atmosphere. In the case of airborne particulate matter, this means that the CMB model can only resolve contributions to primary particle mass. The source strengths calculated using Equation (5) only represent the true source contributions to primary airborne particulate matter if all relevant sources have been included and the source profile measurements are accurate for the region studied. This makes it difficult to apply the CMB model in regions where the nature of characteristic sources is unknown. Finally, the formulation of the CMB model assumes that measurement uncertainties are independent, random, and normal in distribution. This may not be true for measurements of atmospheric particulate matter since biases in parameters such as sampling flow rates, sample extraction volumes, etc might lead to correlated errors.

A recent review article determined that the CMB model had been applied in 22 different studies in locations across 7 countries between 1982 – present (Chow and Watson, 2002). The vast majority of these studies employed some form of metals analysis (PIXE / XRF / ICP-MS), ions analysis (IC, AC, AAS) and carbon (EC/OC via TOR or TOT). One study used the different carbon fractions evolved at increasing temperatures as additional chemical species, and a few studies expanded the list of tracer compounds to consider detailed organic species analyzed via GC/MS methods. In general, those studies that used only metals, ions, and simple carbon (EC/OC) analysis could resolve 3-6 source contributions to ambient particulate matter concentrations with reasonable certainty. Those studies that expanded the list of tracer species to include detailed organic species could resolve 8-9 source contributions to ambient particulate matter. The increased resolution of the CMB studies conducted using detailed organic species results from the fact that unique organic tracers have been identified for certain source categories and so the statistical uncertainty associated with the quantification of those sources is greatly reduced.

Linear methods work well for airborne particulate matter that is released directly from sources, but they cannot accurately predict how the concentration of secondary particulate matter will change when source emissions of gaseous precursors are changed. Studies have shown that the formation of secondary particulate matter components such as sulfate and nitrate is a non-linear process (Ansari and Pandis, 1998; Nguyen and Dabdub, 2002) that cannot be accurately represented using approximate linear models (Foresman et al., 2002). Although this result seems obvious, it has been largely overlooked in certain analyses seeking to evaluate the feasibility of proposed control strategies for secondary airborne particulate matter concentrations.

Bi-Linear Statistical Models

Bi-Linear statistical models have the same general form shown in equation (4), but in this case both [G] and [F] are unknown. The advantage of bi-linear statistical models is that chemical profiles for important sources in a region do not have to be measured. Only ambient measurements of particulate matter composition at a receptor site are needed in the calculation. The disadvantage of bi-linear statistical models is ambiguity. The only definite information revealed from bi-linear statistical models is the number of sources with unique chemical profiles that contribute to airborne particulate matter concentrations. This makes the true determination of the source contributions to ambient particulate matter more difficult.

Several approaches to the solution of bi-linear statistical models have been applied to source apportionment of airborne particulate matter in recent years. Many of these solution techniques share common features. Each of the bi-linear models attempts to minimize the sum of the square of the residuals defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{p} \left(\frac{e_{ij}}{w_{ij}} \right)^2 \qquad (8)$$

where e_{ij} represents the residual for chemical component j in sample i (element ij of matrix $[E]$ in equation 4) and w_{ij} is a weighting factor. If the measured airborne particulate matter concentrations contain p chemical components, bi-linear statistical models can resolve p different source contributions at most. In practice, there are usually only m ($< p$) statistically significant source contributions that bi-linear statistical models seek to identify.

Accurate results can only be obtained from bi-linear statistical models if the dataset of ambient concentrations contains sufficient information about the variation of different chemical components. As a practical matter, this requirement states that bi-linear statistical models can only be applied to datasets that contain many observations of a set of particulate matter concentrations. Henry et al. (1984) recommend that the number of samples n should satisfy the following criteria

$$n > 30 + \frac{p+3}{2} \qquad (9)$$

where p is the number of chemical components in each sample.

If equation (8) is minimized using observed airborne particle measurements $[X]$ in units of μg m^{-3}, then species present in large concentrations will dominate the error term in equation (4) and the problem solution will be heavily weighted towards explaining the variation of these abundant species. This approach neglects information contained in accurate tracers present at low concentrations, and it may lead to erroneous results when measurement errors are present for abundant species. An alternative approach that gives equal weighting to all species regardless of their absolute concentration is data normalization. In this approach, the mean ($\overline{x_j}$) and standard deviation (σ_j) of each chemical species *j* is calculated and a normalized dataset of measured values is produced with elements:

$$z_{ij} = \frac{x_{ij} - \overline{x_j}}{\sqrt{\sigma_j}} \quad (10)$$

When statistical source apportionment techniques are applied to the normalized set of observations [*Z*] equal weighting is given to all species regardless of their absolute concentration. The predicted solution can be transformed back to the original variables (with units) by inverting equation (10) for each chemical species. Both the Principal Component Analysis and Factor Analysis Techniques described in the following sections can be applied to either the original data [*X*] or normalized data [*Z*].

The primary differences between the alternative bi-linear statistical models studied in the current review result from the choice of weighting factors in equation (8), the additional restrictions that are imposed on the solution to satisfy physically realistic criteria, and to a lesser extent the mathematical technique used to solve the problem. The features of several bi-linear statistical models are described in greater detail in the sections below.

Principal Component Analysis

Principal component analysis (PCA) solves equation (4) for source contributions [*G*] and source profiles [*F*] while minimizing the residual *Q* defined by equation (8). The unique features of PCA that are not common to all other bi-linear statistical techniques are:

1) The weighting function w_{ij} is set equal to unity for all chemical species in all samples.
2) All source profiles are normalized such that their sum of squares equals unity ($<f_i>\{f_i\} = 1.0$).
3) Each "source profile" is orthogonal to other source profiles ($<f_i>\{f_i\} = 0$ when $i \neq j$).
4) The first "source profile" is chosen such that it explains the maximum amount of total variation in each chemical species about the mean value for that chemical species. Subsequent "source profiles" are chosen to explain the maximum residual total variation in each chemical species about the mean value.

In PCA applied to the raw data matrix [X], the "variation" of each chemical species about the mean value is described by the $p*p$ covariance matrix [*S*] whose elements are defined by:

$$s_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}(X_{ik} - \overline{X_i})(X_{jk} - \overline{X_j}) \quad (11a)$$

In PCA applied to the normalized data matrix [Z], the "variation" of each chemical species about the mean value is described by the $p*p$ correlation matrix [R] whose elements are defined by:

$$r_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}\left(\frac{X_{ik}-\overline{X_i}}{\sqrt{\sigma_{Xj}}}\right)\left(\frac{X_{jk}-\overline{X_j}}{\sqrt{\sigma_{Xj}}}\right) \quad (11b)$$

Under the restrictions described above, an elegant analytical solution exists for the principal component analysis problem such that each "source profile" <*f*> is an eigen vector for the covariance matrix [*S*] or the correlation matrix [R]. The matrix of *p* "source profiles" containing *p* chemical species defined in this manner has the orthogonality property that $[F]^T = [F]^{-1}$ and so the "source contribution" matrix for the calculation based on [S], for example, is given by:

$$[G] = [X][F]^T \quad (12)$$

The residual terms e_{ij} are all equal to zero because the eigen vectors form a *p* dimension basis set that can explain any combination of observed chemical components with some combination of "source contributions".

In the classical solution to the PCA problem, the eigen value corresponding to each eigen vector determines the relative significance of each identified "source profile". According to the problem formulation, the most significant eigen values / vectors are found first, with progressively less significant eigen values / vectors found after that. Unfortunately, there is no obvious rule for the determination of the optimum number of significant source profiles to retain in the PCA solution. Typically only m ($< p$) eigen values / vectors (sources) are retained that explain 80-90% of the sample variation. A number of other methods also have been suggested to determine the number of significant sources (Henry et al., 1999).

It is often the case that the "source contribution" matrix [*G*] and the "source profile" matrix [*F*] identified by PCA contain negative elements. Negative values occur in the PCA solution because the method specifies that source profiles must be orthogonal to one another. In reality the true source profiles are not orthogonal (if they were orthogonal then they would be composed of only unique tracers for each source and the trivial methods described in previous sections could be used to perform an exact source apportionment). In order to find a physically meaningful solution to the source apportionment problem, the significant "source profiles" [*F*] derived by the traditional PCA method must undergo a linear transformation to produce non-orthogonal "source profiles" [*B*] using the equation

$$[B] = [T][F] \quad (13)$$

[*T*] is a non-singular matrix sometimes called a rotation although this is not strictly accurate since a true rotation could only produce [*B*] that is also orthogonal. Since the rotation matrix [*T*] is non-singular, we can write a new solution to equation (4) that also has a square of residuals equal to zero:

$$[X] = [G][T]^{-1}[T][F] \quad (14)$$

If $[T]$ is chosen such that the rotated source contribution matrix $[G][T]^{-1}$ and the rotated source profile matrix $[T][F]$ contain only positive elements, then the transformed solution satisfies the physical constraints inherent to the source apportionment problem. Unfortunately, for many problems there are an infinite number of rotations $[T]$ that satisfy this criterion and so the solution identified by equation (13) is not unique.

A number of algorithms have been suggested in the statistical literature to determine rotated source profile matrices defined by equation (12). The most widely used algorithm is the 'VARIMAX' orthogonal rotation that seeks to maximize the summation:

$$Q = \sum_{k=1}^{m} \left[\sum_{j=1}^{p} c_{jk}^{4} - \frac{1}{p} \left(\sum_{j=1}^{p} c_{jk}^{2} \right) \right] \quad (15)$$

where

$$c_{jk} = \frac{b_{jk}}{\sqrt{\sum_{k=1}^{m} b_{jk}^{2}}} \quad (16)$$

and b_{jk} are elements of the transformed source profile matrix defined by equation (12). In general terms, this approach tends to drive elements of each source profile towards either zero or extreme positive / negative limits within the range of possible transformations defined by the m significant principal components. Transformed source profiles containing primarily positive values may then bear some similarity to known profiles for common sources of airborne particulate matter. The total amount of variation in the original sample that is explained by the m rotated components is the same as that explained by the m original components, but the distribution of the variation between the components is shifted by the transformation.

A second approach that may be used to find appropriate source profile rotations is linear programming. Nonnegativity constraints and other physical constraints on predicted source compositions can be used to define a feasible solution range in eigen vector space determined from PCA (Henry and Kim, 1990). The resulting solution is usually not unique since it can undergo rotations within the feasible region. In the SAFER model described by Henry et al. (1994), the maximum and minimum composition value of each species in the source profiles are determined within the feasible region define by the problem constraints. Typically the average of the maximum and minimum value is reported as the best estimate of the species concentration in the source profile.

The choice of the number of significant components m to use in the transformed solution can have a large impact on the final result. Several alternative solutions usually must be generated using different choices for m, and then the solution that appears to make the most physical sense is selected as the final solution.

Approximately 41 studies have been conducted to identify and quantify source contributions to atmospheric particulate matter concentrations using PCA. Studies were conducted in locations throughout the world including **Europe** (12 studies: Xhoffer et al., 1991; Harrison et al., 1996; Swietlicki et al., 1996; Pio et al., 1996; Armanino et al., 1996; Rocha and Duarte, 1997; Cardoni et al, 1998; Beceiro-Gonzalez et al., 1998; Rocha et al., 1999; Prati et al., 2000; Kendall et al., 2001; Ruuskanen et al, 2001), the **United States** (9 studies: Saucy et al., 1991; Ehrman et al., 1992; Zieman et al., 1995; Veltkamp et al., 1996; Mukerjee et al., 2001; Smith et al., 2001; Jeon et al., 2001; Claiborn et al., 2002), **Canada** (3 studies: Pryor et al., 1994; Pryor et al., 1997; Breed et al., 2001), **South America** (3 studies: Andrade et al., 1994; Colombo et al., 1999; Matsumoto et al., 2002; Maenhaut et al., 2002), **Greece** (3 studies: Samara et al., 1994; Samara et al., 1994; Manoli et al., 2002), **Africa** (2 studies: Maenhaut et al., 1996; Roberts at al., 2001), **India** (2 studies: Kulshrestha et al., 1995; Balachandran et al., 2000), **Bermuda** (1 study: Anderson et al., 1996), **Hong Kong** (1 study: Lee and Sequeira, 2002), **Indonesia** (1 study: Maenhaut et al., 2002), **Korea** (1 study: Baek et al., 1997), **Singapore** (1 study: Orlic et al., 1995), and **Norway** (1 study: Anderson et al., 1992). On average, these PCA studies were able to identify and quantify approximately 3 "source" contributions to ambient particulate matter concentrations. Examples of sources that could be resolved included airborne dust (22 studies), sea salt (9 studies), motor vehicles (9 studies), biomass combustion (6 studies), fuel oil combustion (6 studies), and coal combustion (2 studies).

PCA techniques often detect the signature of secondary particulate matter sources such as particulate sulfates and/or nitrates. This information does not immediately identify the source of the secondary particulate matter since it is impossible to create a linear "source profile" that captures non-linear secondary formation processes. During periods when the apparent contribution of secondary sources is large, air mass trajectories can be integrated backwards through observed wind fields so that the general location of these sources can be identified (Poirot et al., 2001). The identification of a source location can sometimes provide enough information to infer the identity of the unknown source in question.

In many studies that employed PCA there were a number of sources identified as statistically significant that could not be resolved into a know source profile. This problem illustrates the main disadvantage of bilinear statistical techniques. Rotational ambiguity can mask the true identity of sources that contribute to airborne particulate matter concentrations, making it impossible to specify an emissions control program to reduce pollutant concentrations. On the other hand, PCA can sometimes identify the importance of unknown sources and give some information about their chemical signatures. Backwards integration of air mass trajectories can be used to identify the location of unknown sources and possibly their identity. This makes PCA a powerful tool that can be used to establish the scope of an air quality problem before more quantitative studies are performed.

Factor Analysis

The classical form of Factor Analysis bears many similarities to the PCA method described above. In the Factor Analysis model, equation (4) is solved under constraints that:

1) The weighting function w_{ij} is set equal to unity for all chemical species in all samples.
2) All source profiles are normalized such that their sum of squares equals unity ($<f_i>\{f_i\} = 1.0$).
3) Each "source profile" is orthogonal to other source profiles ($<f_i>\{f_i\} = 0$ when $i \neq j$).
4) The first "source profile" is chosen such that it explains the maximum amount of **common** variation in each chemical species about the mean value for that chemical species. Subsequent "source profiles" are chosen to explain the maximum residual **common** variation in each chemical species about the mean value.

The Factor analysis model explicitly recognizes that only m ($< p$) common sources will be significant in the final solution, and so the residual error term in equation (4) will be non-zero. Additional constraints are placed on the solution stating that

5) Each residual concentration is uncorrelated with other residuals ($<e_i>\{e_i\} = 0$ when $i \neq j$).
6) Each residual concentration is uncorrelated with "source profiles" ($<e_i>\{f_i\} = 0$).

In Factor Analysis the "variation" of each chemical species about the mean value is described by the $p*p$ covariance matrix [R] whose elements are defined by equation (11b). The initial solution to the Factor Analysis problem uses the eigen vectors of the covariance matrix [R] as estimates of the source profiles <f>. This initial solution finds source profiles <f> that explain the maximum amount of total (=common + residual) variation of each chemical species about the mean value, but it does not necessarily explain the maximum amount of common variation. Under constraints (5) and (6) listed above, residual variation terms only contribute to the diagonal elements m+1, m+2, …p of the covariance matrix [R]. The amount of this residual variation is estimated based on the trial solution and subtracted from the appropriate elements of R. This procedure is iterated until it converges to stable values of [F] and [R].

Because Factor Analysis directly solves for [F] with only m (< p) significant sources, there is no direct way to calculate the source contribution matrix [G] ([F] can not be inverted directly since it is not square). Each row of [G] should be selected to use the significant source profiles to explain the observed concentrations with the minimum possible error. Mathematically, choose, <G> such that the error <E> is minimized in the equation:

$$< X > = < G > [F] + < E > \qquad (17)$$

This is exactly the same problem defined by the Chemical Mass Balance model, with solution for each row of [G] defined by equation (5).

Approximately 59 studies have used some form of Factor Analysis to determine source contributions to airborne particulate matter concentrations. Study locations were once again located throughout the world including **Europe** (16 studies: VanBorm et al., 1990; Rojas et al., 1992; VanMalderen et al., 1992; Rojas et al, 1993; Mendex et al., 1993; Einax et al., 1994; DeBock et al., 1994; Otten et al., 1994; DeBock et al., 1996; Harrison et al., 1997; Jambers et al., 1999; Del Carmine et al., 1999; Lucarelli et al., 2000; Cancio et al., 2002; Garcia-Alonso et al., 2002), **South America** (10 studies: Paiva et al., 1993; Artaxo et al., 1993; Artaxo et al., 1994;

Artaxo et al., 1995; Artaxo et al., 1998; Artaxo et al., 1999; Artaxo et al., 2000; Cantanho and Artaxo, 2001; Kavouras et al., 2001; Kavouras et al., 2001), **United States** (7 studies: Morandi et al., 1991; Sweet et al., 1993; Gao et al., 1994; Sheffield et al., 1994; Lawrence and Koutrakis, 1996; Ames et al, 1998; Chen et al, 2002), **India** (4 studies: Sarma and Patil, 1992; Sharma and Patil, 1992; Sharma and Singh, 1992; Sinha and Banerjee, 1997), **China** (3 studies: Zhu and Wang, 1998; Wang et al., 2002; Cao et al., 2002), **Greece** (3 studies: Samara et al., 1994; Kouimtzis et al., 2002; Voutsa et al, 2002), **Africa** (2 studies: Oluyemi et al., 1994; Oluyemi et al., 2001), **Siberia** (2 studies: VanMalderen et al., 1996; VanMalderen et al., 1996), **Vietnam** (2 studies: Hien et al., 1999; Hien et al., 2001), **Australia** (1 study: Chan et al., 1999), **Arctic** (2 studies: Artaxo et al.), **Antarctic** (1 study: Artaxo et al., 1993), **Canada** (1 study: Ko and Jervis, 1992), **Hong Kong** (1 study: Fung and Wong, 1995), **Japan** (1 study: Fujimura et al., 1993), **Korea** (1 study: Kim et al., 2002), **Mexico** (1 study: Diaz et al., 2002), and **Thailand** (1 study: Nouchpramool et al., 1999). On average, these Factor Analysis studies were able to identify and quantify approximately 4 "source" contributions to ambient particulate matter concentrations. Examples of sources that could be resolved included airborne dust (24 studies), sea salt (22 studies), motor vehicles (14 studies), biomass combustion (12 studies), fuel oil combustion (7 studies), coal combustion (6 studies), incinerators (5 studies), and metal smelters (4 studies).

Positive Matrix Factorization

Positive matrix factorization (PMF) is a numerical method that solves equation (4) for source contributions [G] and source profiles [F] while minimizing the residual Q defined by equation (8). The unique features of PMF that are not common to all other bi-linear statistical techniques are:

1. The technique specifies a weighting function w_{ij} that is not necessarily equal to unity. Usually the inverse of the uncertainty for each chemical species is used to weight the error residuals.
2. All source profiles are normalized such that their sum of squares equals unity ($<f_i>\{f_i\} = 1.0$).
3. Elements of [G] and [F] are constrained to be positive.
4. The sum of all source contributions is minimized.

There is no analytical solution to the problem posed in this fashion, and so the PMF algorithm uses a numerical solution approach that minimizes the loss function

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{p}\left(\frac{e_{ij}}{w_{ij}}\right)^2 - \alpha\sum_{i=1}^{n}\sum_{k=1}^{q}\log(g_{ik}) - \beta\sum_{i=1}^{n}\sum_{k=1}^{q}\log(f_{kj}) + \gamma\sum_{k=1}^{q}\left(1-\sum_{j=1}^{p}f_{kj}^2\right)^2 + \delta\left(\sum_{k=1}^{q}\sum_{i=1}^{n}g_{ik}\right)^2 \qquad (18)$$

where e_{ij} is the residual term for species j in sample i, w_{ij} is the weighting factor for species j in sample i, α is the weighting factor for the penalty function that forces elements of [G] to be positive, β is the weighting factor for the penalty function that forces elements of [F] to be positive, γ is the weighting factor for the penalty function that forces the sum of squares for rows of $[F]$ to be unity, and δ is a weighting factor for the penalty function that minimizes the sum of all source contributions.

The major advantage of the PMF technique relative to other bi-linear statistical analysis methods is that the elements of [G] and [F] are constrained to be positive. Several numerical algorithms have been proposed to solve the non-linear problem defined by equation (16) including Gauss-Newton / Newton-Raphson iteration (Paatero, 1997) and the conjugate gradient algorithm (Paatero 1999). These solution methods affect the speed of the calculation but they should not affect the answer to the problem defined by equation (16). The mathematical condition leading to a unique solution within the space of possible positive solutions is the penalty term for the sum of the square of the elements of [G] in equation (16). This approach is somewhat arbitrary and does not necessarily correspond to a physical constraint on the system. Even when this condition is used, the solution to the minimization problem that is found may be a local minimum, not a global minimum. The author of the PMF solution algorithm recommends that each problem be solved multiple times using different initial points so that the overall global solution to the problem can be found (Paatero et al., 2002).

Approximately 11 studies have been conducted to identify and quantify source contributions to atmospheric particulate matter concentrations using PMF. Studies were conducted in locations throughout the world including the **United States** (5 studies: Huang et al., 1999; Ramadan et al, 2000; Pollisar et al., 2001; Song et al., 2001; Lee et al., 2002), **Hong Kong** (2 studies: Lee et al., 1999; Qin et al., 2002), the **Artic** (2 studies: Polissar et al., 1998; Hopke et al., 1999), **Europe** (1 study: Huang et al., 2001), and **Thailand** (1 study: Chueinta et al., 2000). On average, these PMF studies were able to identify and quantify approximately 6 “source” contributions to ambient particulate matter concentrations. Examples of sources that could be resolved included airborne dust (9 studies), sea salt (8 studies), motor vehicles (4 studies), biomass combustion (5 studies), fuel oil combustion (3 studies), coal combustion (3 studies), smelters (4 studies) and incinerators (3 studies).

Several studies have been carried out to compare the results produce by PMF to the results produced by other bi-linear statistical models (see for example Huang et al., 1999; Poirot et al., 2001). These studies have found that the alternative statistical techniques produce similar results when applied to the same data sets. A more important issue to maximize the resolution of factor analysis studies is the selection of the appropriate elements to include in the dataset that will be analyzed.

Empirical Orthogonal Functions

The bilinear statistical models discussed in the previous sections have each been used to explain variation in particle-phase chemical components collected at a single location at a number of different times. It is also possible to use bilinear statistical techniques to explain variation in the spatial distribution of a single chemical component. Equation (4) is written such that the columns of [X] represent concentrations of a single component species measured at different locations, and the rows of [X] contain concentrations measured at different times. The spatial factors [F] identified using this process are referred to as Empirical Orthogonal Functions. The product <G>[F] represents different orthogonal spatial distributions of the chemical component concentration of interest. Plotting these concentration patterns can reveal areas of net sources and sinks for that component.

The majority of the EOF studies that have been carried out have focused on source contributions to visibility reduction in the Western United States (see for example Ashbaugh et al., 1984; Green et al., 1992; Gebhart et al., 1997). These studies focused on the spatial distribution of sulfate concentrations to determine the location of sources that make the greatest contribution to visibility reduction.

Henry et al. (1991) attempted to enhance the use of Empirical Orthogonal Function Analysis by combining it with a simplified partial differential equation that governs the concentration of conserved particle-phase species in the atmosphere:

$$\frac{\partial X}{\partial t} = -u\frac{\partial X}{\partial x1} - v\frac{\partial X}{\partial x2} + Q - S \quad (19)$$

where *X* is the species concentration, *t* is time, *u* is the velocity of wind in the direction *x1*, *v* is the velocity of wind in the direction *x2*, *Q* is the source strength, and *S* is the species removal rate. If Empirical Orthogonal Analysis is used to express the concentrations at a given time and location as the product <X> = <G> {F}, then the concentration at a specific time and location can be used to modify equation (19)

$$\frac{\partial X}{\partial t} = -u < G > \left\{\frac{\partial F}{\partial x1}\right\} - v < G > \left\{\frac{\partial F}{\partial x2}\right\} + Q - S \quad (20)$$

When this equation is integrated with respect to time it yields an expression for the net source strength at a particular location:

$$\overline{Q - S} = \frac{X(T) - X(0)}{T} + \frac{1}{T}\int_0^T u < G > dt\left\{\frac{\partial F}{\partial x1}\right\} + \frac{1}{T}\int_0^T v < G > dt\left\{\frac{\partial F}{\partial x2}\right\} \quad (21)$$

where the overbar represents an average with respect to time. This formulation recognizes that the spatial Empirical Orthogonal Functions {F} are not a function of time. The integral expression for weighted velocity at each location is approximated using a simple averaging over the *n* times when observations were made:

$$\overline{Q - S} = \frac{X(T) - X(0)}{T} + \frac{1}{n}\sum_{i=1}^{n} u_i < G_i > \left\{\frac{\partial F}{\partial x1}\right\} + \frac{1}{n}\sum_{i=1}^{n} v_i < G_i > \left\{\frac{\partial F}{\partial x2}\right\} \quad (22)$$

Equation (22) is spatially interpolated from the set of non-uniform grid points where observations are available to a uniform grid using $1/r^2$ interpolation. Spatial gradients are then calculated using discrete approximations, leading to a set of equations that can be solved for the average source strength at any grid location.

Equation (22) provides a method to identify source locations in the presences of strong convective conditions. The assumptions implicit in the formulation of equation (22) include all those needed to derive the PCA or FA method used to calculate the EOF's in the system. Additional restrictions to this model require that species are homogeneously mixed in the surface layer, turbulent diffusion processes are negligible, and the density of sampling sites is sufficient to resolve important horizontal gradients. In practice these conditions are rarely met, and only one study has used the EOF formulation described by equation (22) to date (Henry et al., 1991).

Hybrid Statistical Models

Halfway between Linear and Bi-linear statistical models are hybrid statistical models. These techniques use some known information about [G] or [F] to transform the solution derived from a bi-linear statistical model into more known and quantitative terms.

Chemical Mass Balance with Factor Analysis of Residuals

The Chemical Mass Balance model described in Section 3.1 minimizes the residual error <E> in equation (3) through the calculation of source contributions <G>. If this technique is applied to a dataset that includes multiple observations, then a set of residual errors [E] is generated. Presumably these residual errors reflect the influence of random errors and unknown sources on ambient concentrations. The identity and strength of unknown sources can be identified if the residual error matrix [E] is analyzed with the bilinear statistical techniques described in Section 3.2. It should be noted, however, that the composition and strength of unknown sources suffers from the same rotational ambiguity that resides with all bilinear statistical techniques. The number of significant sources retained to explain the residual component of the CMB solution, and the rotations applied to those sources contribute significant uncertainty to the calculation.

Target Transformation Factor Analysis

The traditional bilinear statistical models (PCA and FA) identify a set of orthogonal basis source profiles that can be used to explain the variation in measured ambient concentrations. The source profiles [F] derived from these methods often do not satisfy physical constraints (all elements greater than zero) and so they must undergo a linear transformation into a more realistic form. In Target Transformation Factor Analysis (Hopke, 1988), it is recognized that if the composition of a possible source <b> is known, a transformation <r> can be found that attempts to express the known source profile as a linear combination of the source profiles contained in [F].

$$< b > = < r > [F] + < E > \qquad (23)$$

The choice of <r> that does the best job of reproducing <b> from the orthogonal basis source profiles minimizes the residual error <E>. The solution to this problem is mathematically identical to the solution to the CMB model:

$$< r > = < b > [W][F]\left([F]^T [W][F]\right)^{-1} \quad (24)$$

where [W] is a diagonal weighting matrix with elements equal to the inverse of the variance in the element concentrations. Successive target source profiles <b> can be found using associated transformation vectors <r>. If the residual error <E> in equation (21) is too large, the target source is considered to be inappropriate and it is not included in the matrix of transformed source profiles.

In addition to known target source profiles, a number of source profiles can be derived by starting with a test profile that has a 1.0 in one element location, and zero for all others. After calculation of the initial transformation matrix <r>, a revised target profile <b'>is formed using the equation (21) with <E> set equal to zero. The revised target profile is modified to set any negative values to small positive values, and then substituted into equation (22) to find a revised transformation <r'>. Iteration is continued until the solution converges to a stable combination of <b'> and <r'>.

Approximately 6 studies have been conducted to identify and quantify source contributions to atmospheric particulate matter concentrations using Target Transformation Factor Analysis. Studies were conducted in locations throughout the world including the **United States** (2 studies: Alpert and Hopke, 1980; Alpert and Hopke, 1981), **Europe** (2 studies: Borbelykiss et al., 1993; Moro et al., 1997), **China** (1 study: Zelenka et al., 1994), and **Australia** (1 study: Chan et al., 1999).

Target Transformation Factor Analysis and CMB analysis with Factor Analysis of residuals have many common features. Both methods explain as much sample variation as possible using known source composition information, and then explain residual variation using a Factor Analysis approach. A detailed study of the TTFA method determined that it does not provide unique solutions to the Factor Analysis problem due to rotational ambiguity. There is no way to remove operator judgment in choosing the number of factors and the selection of profiles to include in fitting the data.

Mechanistic Analysis Techniques

Mechanistic analysis techniques for the source apportionment of PM2.5 attempt to reconstruct the concentration of airborne particulate matter without prior knowledge of ambient concentrations. A complete mechanistic analysis includes an emissions inventory, atmospheric transport (advection and turbulent diffusion), particle removal (dry and wet deposition), chemical reaction (gas-phase and particle phase), gas-to-particle conversion (condensation and nucleation), and particle coagulation. In some situations the input data needed to support detailed mechanistic air quality calculations is not available or some atmospheric processes are judged to be negligible. In these circumstances more simplistic mechanistic calculations can be useful if the limitations of the analysis are clearly stated.

The mechanistic analysis techniques that have been used to perform a source apportionment of airborne particulate matter are described in the sections below.

Emissions Inventory Analysis

Emissions inventories describe the release rate, the spatial allocation, and the diurnal variation of pollutant emissions within a region of interest. An emissions inventory analysis is the simplest possible form of a mechanistic source apportionment for airborne particulate matter. The "mechanistic" component of this analysis is the creation of a detailed emissions inventory that describes pollutant emissions to the atmosphere separated according to source categories. No attempt is made to mechanistically represent atmospheric transformation processes, but rather it is assumed that processes such as mixing, deposition, and chemical transformation affect emissions from all sources in a similar fashion. Therefore, the emissions inventory analysis approach assumes that sources contribute to airborne particulate matter concentrations in proportion to their emissions strength. The contribution that each source makes to airborne particle concentrations is calculated as:

$$X_j^k = \left(X_j^{total} - X_j^{\sec ondary} - X_j^{background}\right) * \frac{G_k^j}{G_{total}^j} \qquad (25)$$

where X^k_j is the portion of the chemical species *j* contained in airborne particles that is attributed to the target source *k*, X_j^{total} is the total concentration of chemical species *j* measured in airborne particles, $X_j^{background}$ is the concentration of chemical species *j* that does not originate in the study region, $X_j^{secondary}$ is the concentration of chemical species *j* that forms in the atmosphere by chemical reaction, G^j_k is the emissions strength of the target source *k* for chemical species *j,* and G^j_{total} is the total emissions for chemical species *j*.

The source apportionment for airborne particulate matter determined using equation (25) assumes that $X_j^{background}$ and $X_j^{secondary}$ are well characterized (or negligible) and that the precursor emissions of particulate chemical species *j* are known. The background concentration of particulate chemical species *j* can be determined by measuring the incoming concentration at the boundary of the study area. The determination of secondary concentrations of particulate species is more problematic. In some cases there are no known primary sources of species *j* and so any atmospheric concentrations are attributed to secondary production (nitrate, sulfate, ammonium ion). In this case, the ratio of precursor species emissions are sometimes used as a crude estimate of source attribution for those species, although this method ignores non-linear aspects of atmospheric chemistry involved in secondary production. In general, the fraction of the precursor gas that reacts to form the particulate species depends on the concentration of other species in the system and the condition of the atmosphere. Estimation of secondary species source apportionment using inventory analysis is only possible over long time scales and large geographical domains where the nonlinear aspects of the atmospheric chemistry average out.

Emissions inventory analysis is one of the most basic mechanistic techniques that can be applied to a particulate air quality problem and so many studies have been applied but only a few have been published. A few recent examples at the urban scale involve the determination of contributions from different mobile sources to urban particulate air quality (Newmark, 2001) and the contribution of secondary organic aerosol production to particulate organic carbon concentrations (Cabada et al., 2002). At the global scale, emissions inventory analysis has been used to identify sources of black carbonaceous aerosol (Cook and Wilson, 1996) and overall particulate matter mass (Wolfe and Hidy, 1997). In some cases, emissions inventory estimates for source contributions to airborne particulate matter concentrations can be verified by trend analysis when emissions rates change. For example, trends in the concentration of sulfate in precipitation has been qualitatively linked to changes in anthropogenic emissions of SO2 (see for example Baier and Cohn, 1993; Shannon 1999).

Source-Oriented Chemical Transport Models

Source-oriented chemical transport models (CTMs) track pollutant emissions released from different sources separately through the mechanistic model framework. The reactive advection diffusion equation forms the basis of all chemical transport models (CTMs):

$$\frac{\partial X_i}{\partial t} + \nabla \cdot u X_i = \nabla K \nabla X_i + E_i - S_i + R_i^{gas}(X) + R_i^{part}(X) + R_i^{phase}(X) \qquad (26)$$

where X_i is the concentration of gas- or particle-phase species at a particular location as a function of time t, u is the wind vector, K is the turbulent eddy diffusivity tensor (assumed to be diagonal), E_i is the emissions rate, S_i is the loss rate, R_i^{gas} is the change in concentration due to gas-phase reactions, R_i^{part} is the change in concentration due to particle-phase reactions, and R_i^{phase} is the change in concentration due to phase change. Mechanistic air quality models specify initial concentrations $X_i^{initial}$ and boundary concentrations $X_i^{boundary}$ so that equation (26) can be integrated with respect to time to predict particulate matter concentrations. Source-oriented chemical transport models write equation (26) multiple times for chemical species contained on different particles released from different sources to determine source contributions to airborne particulate matter.

Many of the terms shown on the right side of equation (26) are evaluated in separate steps within the air quality model (operator splitting) that require multiple layers of implicit iteration to achieve convergence. Separate review articles have been written comparing various approaches for particulate matter modeling (Seigneur, 2001), photochemical models (Russell and Denis, 2000), pollutant advection (Chock 1990; Dabdub and Seinfeld 1994), chemical oxidant mechanisms (Dodge 2000), inorganic aerosol thermodynamic equilibrium modules (Zhang et al., 2000), and stiff ordinary differential equation solvers (Chock et al., 1994). In the present review, discussion is restricted to the mathematical foundations of the Lagrangian (Kleeman et al., 1997) and Eulerian (Kleeman and Cass, 2001) versions of the source-oriented CTMs that have been demonstrated to date.

Advection

The Lagrangian version of the source-oriented CTM employs a moving reference frame that is fixed to the mean horizontal wind vector. Vertical advection and horizontal turbulent diffusion are assumed to be negligible. Under these conditions equation (26) simplifies to

$$\frac{\partial X_i}{\partial t} = \frac{\partial}{\partial z}\left(K_{zz} \frac{\partial X_i}{\partial z} \right) + E_i - S_i + R_i^{gas}(X) + R_i^{part}(X) + R_i^{phase}(X) \qquad (27)$$

where *z* refers to the vertical direction. This formulation of the mechanistic model is still 3 dimensional, but it cannot represent certain atmospheric processes that may be important for certain episodes (vertical wind shear, horizontal turbulent diffusion).

The Eulerian version of the source-oriented CTM uses the Accurate Space Derivative (ASD) method (Gazdag, 1973) to advect particles and gases in the horizontal plane between fixed grid cells. The ASD method expands the time rate of change in species concentration due to advection in a Taylors Series and replaces the temporal derivatives with spatial derivatives that satisfy the advection equation to produce:

$$X_i^{t+\Delta t} = X_i^t - \frac{\partial (u_z X_i)}{\partial z}\Delta t + \frac{\partial}{\partial z}\left(u\frac{\partial (u_z X_i)}{\partial z}\right)\Delta t^2 - \frac{\partial}{\partial z}\left(u\frac{\partial}{\partial z}\left(u\frac{\partial (u_z X_i)}{\partial z}\right)\right)\Delta t^3 \quad (28)$$

Spatial derivatives are evaluated by fitting a Fourier series to the distribution of Xi, differentiating, and then transforming back to the original coordinate system.

Turbulent Diffusion

Turbulent diffusion in both the Lagrangian and Eulerian form of the source-oriented CTM is based on simple first-order K theory. Turbulent fluctuations in the produce of velocity and concentration are set equal to the mean gradient of those properties:

$$< u_z' X' >= -K_{zz}\frac{\partial X}{\partial z} \quad (29)$$

where u_z' is the turbulent component of velocity in some direction *z*, *X'* is the turbulent component of concentration, the angle brackets are the ensemble-averaging operator, and K_{zz} is the turbulent eddy diffusivity in the direction *z*. K_{zz} values in the vertical direction are parameterized as a function of atmospheric stability conditions (McRae et al., 1982). K_{zz} values in the horizontal direction are set to be equal to some constant fraction of the vertical component.

Gas-phase Chemical Reaction Mechanism

Both the Lagrangian and the Eulerian versions of the source-oriented CTM use the Carter 90 (Carter 1990) gas-phase chemical oxidant mechanism with modifications to describe the production of semi-volatile organic species (Pandis et al., 1992). The mechanism has 10 detailed inorganic species, 23 detailed organic compounds, 24 lumped organic compound classes, and 29 semi-volatile lumped organic products. Lumped organic classes are formed from compounds with similar chemical structure and reactivity. Reaction rates for lumped classes are calculated as averages weighted by the emissions rates of individual compounds within each class. The overall mechanism consists of 193 reactions that contribute to 86 ordinary differential equations.

Aqueous-phase Chemical Reaction System

During periods of high humidity a relatively large amount of water condenses onto airborne particles leading to the formation of aqueous droplets. Soluble gases dissolve in these droplets and participate in chemical reactions leading to products that remain in the particle-phase after relative humidity falls back to unsaturated conditions. Aqueous chemical reactions on each droplet are considered under these conditions based on the fog model described by Jacob (1986) and Jacob et al. (1989). The model calculates the vapor pressure of water above the particle surface using experimentally measured osmotic coefficients for electrolytes in solution (Tang and Munkelwitz, 1994). When the ionic strength of the droplets falls below 0.5, the dissolution of 22 detailed inorganic and organic gas-phase species into the aqueous phase is calculated. Species in aqueous solution are acted on by a kinetic reaction mechanism focusing on the oxidation of sulfur via pathways including iron, manganese, and copper catalysis, in addition to reaction with dissolved ozone and hydrogen peroxide. The aqueous mechanism consists of 58 active chemical species undergoing 177 kinetic reactions while constrained by 29 equilibrium relationships. Ordinary differential equations are formed for each of the 58 active chemical species on each particle.

Gas-to-particle conversion

Gas-to-particle conversion rates account for resistance to mass transfer associated with gas-phase diffusion and interfacial barriers. The rate of change in particulate matter concentrations that result from gas-to-particle conversion processes is described by the differential equation:

$$\frac{\partial X_i^{particle}}{\partial t} = 4\pi D_i \frac{RpN}{\beta + 1}\left(X_i^{gas,\infty} - X_i^{gas,surface}\right) \quad (30a)$$

$$\beta = \frac{4D_i}{\alpha_i \hat{c} Rp} \quad (30b)$$

$$\hat{c} = \sqrt{\frac{8RT}{\pi Mwt_i}} \quad (30c)$$

where D_i is the gas-phase diffusion coefficient, Rp is the particle radius, N is the particle number concentration, α_i is the accommodation coefficient, $\hat{c}$ is the molecular speed of the gas molecules, R is the gas constant, T is the ambient temperature, and Mwt_i is the molecular weight of chemical species i.

Integration of Ordinary Differential Equations

The coupled sets of stiff ordinary differential equations describing the gas-phase reaction mechanism, the aqueous-phase reaction system, the gas-to-particle conversion equations is solved using a modified form of the hybrid numerical integration technique described by Yong and Boris (1977). The technique integrates normal differential equations using the implicit Euler equation:

$$X_i^{t+\Delta t} = X_i^t + \frac{\partial X_i}{\partial t} * \Delta t \qquad (31)$$

and integrates stiff differential equations using an asymptotic technique:

$$X_i^{t+\Delta t} = \beta + (X_i^t - \beta) * \exp\left(\frac{-2\Delta t}{\tau(t) + \tau(t+\Delta t)} \right) \qquad (32)$$

where

$$\beta = \left(\frac{a(t) + a(t+\Delta t)}{2} \right) * \left(\frac{\tau(t) + \tau(t+\Delta t)}{2} \right) \qquad (33)$$

and *a, b* are coefficients in the equation $\partial X_i / \partial t = a - bX_i$.

Aerosol Thermodynamics for Inorganic and Organic Species

The inorganic aerosol thermodynamics module used in both the Lagrangian and the Eulerian versions of the source-oriented CTM is a modified version of the Aerosol Inorganic Module (AIM) (Wexler and Seinfeld, 1990; Kleeman et al., 1997). Sulfate, nitrate, chloride, sodium, ammonium ion, and hydrogen ion concentrations in the aqueous phase associated with each particle are calculated while considering the possible formation of 9 different solid species. Equilibrium inorganic concentrations on each particle are calculated by minimizing the Gibbs Free Energy of individual equilibrium reactions until the entire system reaches the point of minimum Gibbs Free Energy. Equilibrium vapor pressures for HNO3, HCl, H2SO4, and NH3 directly above the particle surface are calculated based on the particle composition. This information is used to calculate the kinetic exchange of species between the gas and particle phases.

The organic aerosol thermodynamics module used in both the Lagrangian and the Eulerian versions of the source-oriented CTM is based on organic partitioning coefficients measured during smog chamber experiments (Odum et al., 1996). Equilibrium vapor pressures (X_i^{gas}) for classes of semi-volatile organic compounds are calculated using the equation:

$$X_i^{gas} = \frac{X_i^{particle}}{K_i M_o} \qquad (34)$$

where $X_i^{particle}$ is the concentration of species *i* in the particle phase, K_i is the partitioning coefficient measured during experiments, and M_o is the concentration of all organics in the particle phase.

Emissions

The standard emissions inventories supplied by regulatory agencies for mechanistic air quality modeling contain the spatial and temporal distribution of regulatory pollutants (CO, NOx, SOx, PM10) and volatile organic compounds (VOC's). Typical resolution for emissions inventories is 4-10km in the spatial coordinate and 1-24 hrs in the time coordinate. Emissions of lumped species (VOC and PM) must be broken out into detailed species concentrations using emissions profiles [F] for individual source types. These profiles have the same format as the emissions profiles used in statistical models: they describe the amount of individual chemical compound present per unit of emitted pollutant. In previous applications of the source-oriented CTM VOC profiles are based on source measurements made by Schauer et al. (1999a, 1999b, 2001, 2002a, 2002b). Size and composition resolved particulate matter source profiles used in the model are based primarily on measurements made by Kleeman et al. (1999; 2000) and Hildemann et al. (1991a; 1991b).

Ammonia is the most important basic gas that acts to neutralize acid condensation products such as sulfuric acid, nitric acid, and hydrochloric acid in the atmosphere. Ammonia emissions are a necessary component in any mechanistic air quality model that includes a description of secondary sulfate and nitrate. In previous applications of the source-oriented CTM, ammonia emissions were calculated for the region of interest based on landuse, population, and agricultural patterns (Gharib and Cass, 1984).

The source-oriented CTMs that have been used to date have distinguished between the following classes of airborne particles: paved road dust, crustal material other than paved road dust, food cooking, catalyst-equipped gasoline powered vehicles, non-catalyst equipped gasoline powered vehicles, diesel engines, high-sulfur content fuel combustion, sea salt particles, background marine particles that contain sulfate, and other anthropogenic sources not described by one of the previous 9 categories.

Deposition

Dry deposition of particles follows the method described by Slinn and Slinn (1980) with appropriate modifications to account for the effect of atmospheric stability conditions. The resultant equation for the deposition velocity of particles is:

$$v_p = \frac{1}{r_C + r_D + r_C r_D v_S(Rp)} + v_S(Rp) \quad (35)$$

with

$$(r_D)^{-1} = u_*[(Sc)^{-2/3} + 10^{-3/St}] \quad (36)$$

where r_C is the resistance to deposition in the constant flux layer, r_D is the resistance in the sublayer immediately adjacent to the ground, u_* is the friction velocity, Sc is the Schmidt number, St is the Stokes number, and $v_S(Rp)$ is the settling velocity for particles with aerodynamic diameter Rp. The resistance in the constant flux layer is a function of atmospheric stability conditions (McRae et al., 1982) and surface roughness estimated from landuse maps (Russell et al., 1993).

Source Apportionment of Secondary Particulate Matter

The most recent version of the Lagrangian form of the source-oriented CTM tracks the formation of secondary nitrate, sulfate, and ammonium ion from different sources by following the evolution of the NO_x, SO_x, and NH_3 released from different sources in the atmosphere. The linear nature of kinetic reactions is exploited to expand the resolution of the gas-phase reaction system involving precursor species from each source. If there are two major sources of precursor species X_{j1} in the atmosphere, and X_{j1} reacts with hydroxyl radical to form X_{j2}, then the total reaction equation:

$$X_{j1} + OH \rightarrow X_{j2} \qquad (37)$$

can be divided into two reaction equations that track the evolution of X released from different sources:

$$X_{j1}^{source1} + OH \rightarrow X_{j2}^{source1} \qquad (38a)$$
$$X_{j1}^{source2} + OH \rightarrow X_{j2}^{source2} \qquad (38b)$$

The total amount of X_{j2} formed from equation (37) equals the amount of X_{j2} formed from equations (38a) and (38b). Each reaction involving n reactants that are tracked separately from m different sources produce an expanded set of $m*n$ equations. In the current version of the source-oriented CTM model the gas-phase oxidation system includes over 1100 reactions. This does not increase the computational burden of problems significantly since relative little computation effort is associated with solution of the gas-phase system. The model differentiates between 6 different sources of NOx and SOx (diesel engines, non-catalyst equipped gasoline engines, catalyst-equipped gasoline engines, high-sulfur fuel combustion, other primary sources, background) and 7 different sources of NH3 (refrigerant, residential, animals, catalyst-equipped gasoline engines, fertilizer, other primary sources, background).

Applications of Source-oriented CTM Source Apportionment

Five studies have been completed to date using source-oriented CTMs to determine contributions to airborne particulate matter concentrations (Kleeman et al., 1997; Kleeman and Cass, 1998; Kleeman and Cass, 1999; Kleeman and Cass, 2001; Mysliwiec and Kleeman, 2002). All of these studies were applied to air quality episodes that occurred in Los Angeles, California. The limited number of applications reflects the recent development of the source-oriented CTM technique and the large amount of input data needed to perform the calculations. Source-oriented CTM calculations were used to identify contributions to PM2.5 and PM10 from 9 different source types: paved road dust, crustal material other than paved road dust, food cooking, diesel engines, catalyst-equipped gasoline-powered engines, non-catalyst-equipped gasoline-powered engines, high-sulfur fuel combustion, marine particles, and background sulfate particles.

Source-oriented CTMs do not rely on statistical differences between ambient concentrations to identify source contributions. As a result, the absolute concentration of particulate matter and the relative composition of particles released from different sources do not affect the sensitivity of the source-oriented CTM calculations. Any number of emissions sources with similar or identical profiles can be tracked through the mechanistic calculation as long as they are emitted to the atmosphere at separated locations and/or times. Tracking a larger number of sources through the source-oriented CTM calculation does require more processing time and greater amounts of computer memory. The studies that have been conducted to date have resolved 10 source contributions to airborne particulate matter concentrations as a matter of computational convenience. As computers continue to increase in speed and decrease in cost, this limitation will relax. One study has already used the source-oriented CTM approach to calculate more than 50 source contributions to airborne particulate matter concentrations at a receptor site (Kleeman and Cass, 1999).

Regional source apportionment information is useful for population exposure calculations and the identification of localized areas with high concentrations. The Eulerian version of the source-oriented CTM provides a regional source-apportionment of airborne particulate matter. The spatial resolution that has been demonstrated to date is 5km in the horizontal direction over an urban area that measured approximately 100*230km. A larger number of ambient samples at many locations would be needed to produce a statistical source apportionment of particulate matter with the same resolution over the same area.

Many geographical areas with high PM2.5 concentrations are characterized by the presence of high concentrations of secondary particulate $SO_4^=$, NO_3^-, NH_4^+, and SOA. Identification of source contributions to these secondary particulate matter concentrations is necessary for the design of effective emissions control strategies. Source-oriented CTMs can directly calculate source contributions to secondary particulate matter concentrations with no ambiguity about the original source origin or location. The extra computational burden of this calculation is minimal, since most of the extra effort is associated with the gas-phase chemical reaction system.

The chief disadvantage associated with source-apportionment calculations performed using source-oriented CTMs is that these models require large amounts of detailed input data describing initial conditions, the location and diurnal variation of relevant pollutant emissions, and meteorological parameters such as wind speed, wind direction, temperature, relative humidity, total solar radiation, and ultra violet solar radiation in the region of interest. The quality of this input data determines a large part of the error and uncertainty associated with the mechanistic air quality calculation. Missing or inaccurate emissions data directly affects the results of the source-oriented CTM calculations. Model validation data that describes the size and composition distribution of airborne particulate matter in the study domain also is needed to verify the accuracy of source-oriented CTM calculations.

Discussion

Figure 1 shows the information required and the information provided by each source apportionment tool considered in the present review. The information required by each of the methods shown in Figure 1 is cumulative, while the information provided is not. For example, methods that require source profile information to predict contributions to primary particulate matter concentrations also require ambient concentrations, but they do not provide information about the number of significant sources contributing to airborne particulate matter concentrations.

Bilinear statistical techniques such as Principal Component Analysis, Factor Analysis, Positive Matrix Factorization, and Empirical Orthogonal Functions can each provide useful information to help identify important source contributions to airborne particulate matter concentrations when only ambient measurements are available. The exact identification of sources and source contributions using bilinear models is not exact due to rotational ambiguity in the solution. After important sources have been identified and characterized, the linear Chemical Mass Balance method or hybrid approaches such as Chemical Mass Balance / Factor Analysis or Target Transformation Factor Analysis can provide improved estimates of source contributions to airborne primary particulate matter concentrations. When emissions inventories and meteorological patterns have been characterized for the problem of interest, Mechanistic Source-oriented Chemical Transport Models can be used to estimate regional source contributions to primary and secondary airborne particulate matter concentrations with even greater source resolution. Mechanistic models require large amounts of input data and their results must be compared to ambient measurements before confidence can be established in their predictions.

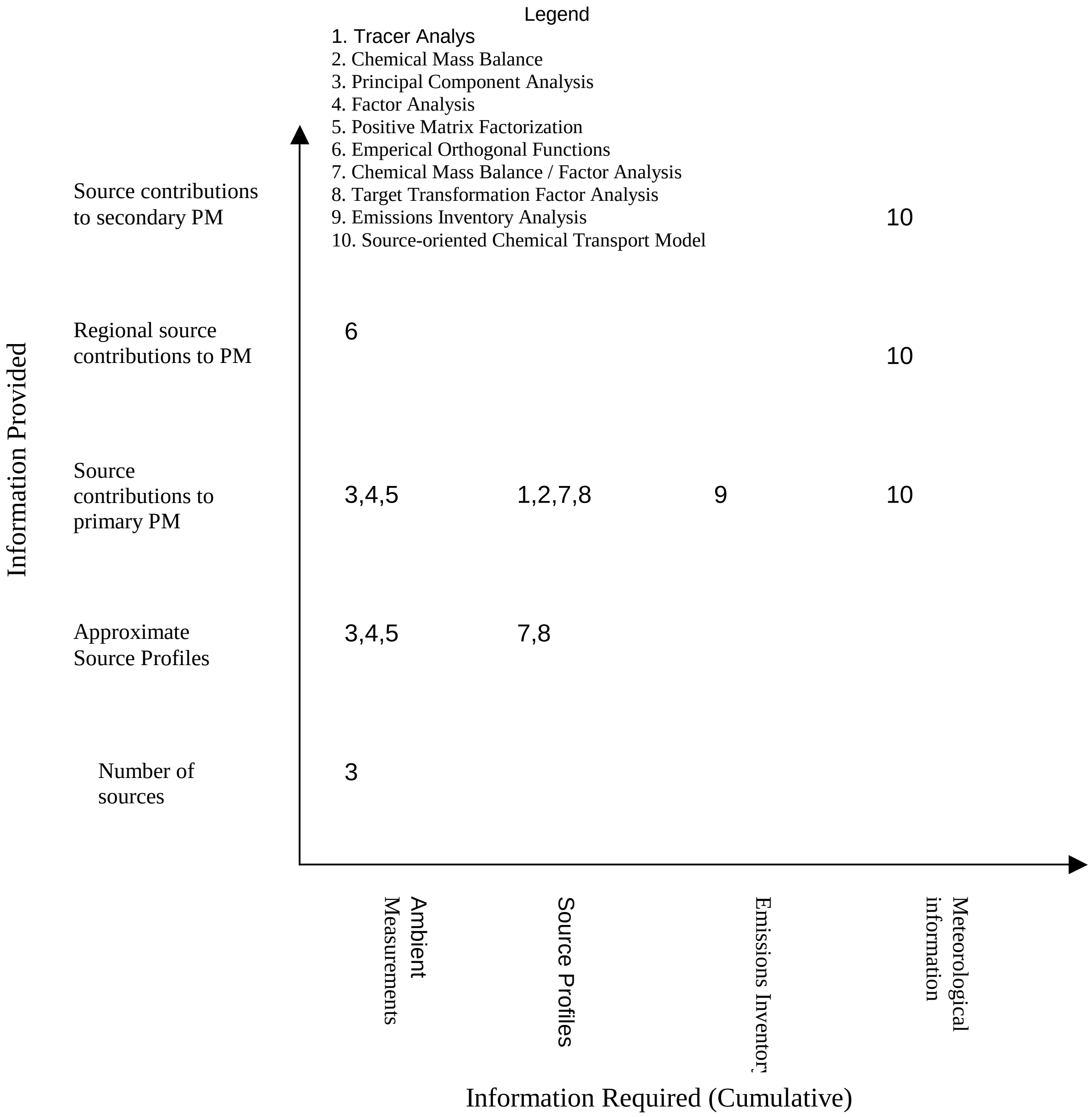

Figure 1: Summary of information required and information provided by source apportionment methods for airborne particulate matter.

Summary and Conclusions

Even when trace organic species are considered, unique chemical tracers exist for only a few important sources of airborne particulate matter, and so statistical and mechanistic source apportionment models are needed to identify source contributions to airborne particulate matter. A review of source apportionment models was conducted to compare the features of alternative methods that have been employed in previous studies. The various source apportionment techniques form an array of tools that require increasing amounts of information to provide increasing detail about source contributions to airborne particulate matter. A logical approach to source apportionment studies for airborne particulate matter would involve the following steps:

1. Characterize the airborne particulate matter concentration and composition over the spatial and temporal scales that are relevant to the problem of interest.
2. Analyze the data collected in step 1 using bilinear statistical techniques to identify the number of important sources and to infer approximate information about important source types.
3. Characterize exact source profiles for airborne particulate matter emissions identified in step 2 by collecting samples of airborne particulate matter at the source and analyzing the samples using the techniques employed in step 1.
4. Analyze the data collected in steps 1 and 3 using the chemical mass balance model. If the source profiles used in the calculation do not explain the variation of measured particulate matter concentrations, use a hybrid approach (CMB/FA or TTFA) to identify the approximate composition of missing source profiles (return to step 3 to characterize these profiles).
5. Construct an emissions inventory if the regional distribution of airborne particulate matter is of interest or the concentration of secondary particulate matter is significant. Perform a simple emissions inventory analysis and compare with CMB results to verify accuracy.
6. Apply a source-oriented chemical transport model to the region to study regional contributions to airborne particulate matter concentrations and source contributions to secondary particulate matter.

The application of these 6 steps should lead to the identification of sources that contribute to airborne particulate matter concentrations under virtually any circumstances. Emissions control programs can then be designed for primary particulate matter concentrations using simple linear-rollback calculations, or for secondary particulate matter using mechanistic chemical transport models.

Recommendations

The power of statistical source apportionment techniques is increased by the use of unique tracers associated with specific sources of interest. Future research should be conducted to identify additional unique tracers for gasoline-powered motor vehicles and diesel vehicles to help differentiate these sources from one another.

Mechanistic source apportionment models offer the best method for regional source apportionment studies of primary and secondary particulate matter. Detailed emissions inventories should be developed and validated for all parts of California to support the use of mechanistic models.

References

Alpert D.J., Hopke P.K. (1980) A quantitative dermination of sources in the Boston urban aerosol. Atmos. Environ. 14, 1137-1146.

Alpert D.J., Hopke P.K. (1981) A determination of the sources of airborne particles collected during the regional air pollution study. Atmos. Environ. 15, 675-687.

Ames M., Gullu G., Olmez I. (1998) Atmospheric mercury in the vapor phase and in fine and coarse particulate matter at Perch River, New York. Atmos. Environ. 32, 865-872.

Andrade F., Orsini C., Maenhaut W. (1994) Relation between aerosol source and meteorological parameters for inhalable atmospheric particles in Sao-Paulo City, Brazil. Atmos. Environ. 28, 2307-2315.

Anderson J.R., Buseck P.R., Patterson T.L., Arimoto R. (1996) Characterization of the Bermuda tropospheric aerosol by combined individual-particle and bulk-aerosol analysis. Atmos. Environ. 30, 319-338.

Anderson J.R., Buseck P.R., Saucy D.A., Pacyna J.M. (1992) Characterization of individual fine-fraction particles from the artic aerosol at Spitsbergen, May-June 1987. (1992) Atmos. Environ. 26, 1747-1762.

Ansari A.S., Pandis S.N. (1998) Response of inorganic PM to precursor concentrations. Environ. Sci. and Technol. 32, 2706-2714.

Armanino C., Roda A., Ius A., Casolino M.C., Bacigalupo M.A. (1996) Pattern recognition of particulate-bound pollutants sampled during a long term urban air monitoring scheme. Environmetrics 7, 537-550.

Artaxo P., Rabello M.L.C., Watt F., Grime G., Swietlicki E. (1993) Nuclear microprobe analysis and source apportionment of individual atmospheric aerosol particles. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms. 75, 521-525.

Artaxo P., Gerab F., Yamasoe M.A., Martin J.V. (1994) Fine mode aerosol composition at 3 long-term atmospheric monitoring sites in the Amazon Basin. Journal of Geophysical Research. 99, 22857-22868.

Artaxo P., Hannson H.C. (1995) Size distribution of biogenic aerosol particles from the Amazon Basin. Atmos. Environ. 29, 393-402.

Artaxo P., Fernandes E.T., Marins J.V., Yamasoe M.A., Hobbs P.V., Maenhaut W., Longo K.M., Castanho A. (1998) Large-scale aerosol source apportionment in Amazonia. Journal of Geophysical Research. 103, 31837-31847.

Artaxo P., Oyola P., Martinez R. (1999) Aerosol composition and source apportionment in Santiago de Chile. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 150, 409-416.

Artaxo P., DeCampos R.C., Fernandes E.T., Martins J.V., Xiao Z.F., Lindqvist O., Fernandez-Jimenez M.T., Maenhaut W. (2000) Large scale mercury and trace element measurements in the Amazon Basin. Atmos. Environ. 34, 4085-4096.

Artaxo P., Rabello M.L.C., Maenhaut W., VanGrieken R. (1992) Trace-elements and individual particle analysis of atmospheric aerosols from the Antarctic Peninsula. Tellus Series B-Chemical and Physical Meteorology. 44, 318-334.

Ashbaugh L.L., Myrup L.O., Flocchini R.G. (1984) A principal component analysis of sulfur concentrations in the western United States. Atmos. Environ. 18, 783-791.

Baek S.O., Choi J.S., Hwang S.M. (1997) A quantitative estimation of source contributions to the concentration of atmospheric suspended particulate matter in urban, suburban, and industrial areas of Korea. Environment International. 23, 205-213.

Bair W.G., Cohn T.A. (1993) Trend analysis of sulfate, nitrate, and pH data collected at National Atmospheric Deposition Program/National Trends Network stations between 1980 and 1991. U.S. Geological Survey Open File Report 93-56. U.S. Geological Survey, Reston, VA.

Balachandran S., Meena B.R., Khillare P.S. (2000) Particle size distribution and its elemental composition in the ambient air of Delhi. Environment International. 26, 49-54.

Beceiro-Gonzalez E., Andrade-Garda J.M., Serrano-Velasco E., Lopez-Mahia P., Prada-Rodrigues D. (1998) Characterization of anthropogenic and natural emissions of particulate heavy metals in a surrounding area of La Coruna. Afinidad. 55, 207-212.

Borbelykiss I., Koltay E., Szabo G. (1993) Apportionment of atmospheric aerosols collected over Hungary to sources by target transformation factor analysis. Nuclear Instruments and Methods in Physics Research Section B-Beam Interactions with materials and atoms. 75, 287-291.

Breed C.A., Arocena J.M., Sutherland D. (2002) Possible sources of PM10 in Prince George (Canada) as revealed by morphology and in situ chemical composition of particulate. Atmos. Environ. 36, 1721-1731.

Cabada J.C., Pandis S.N., Robinson A.L. (2002) Sources of atmospheric carbonaceous particulate matter in Pittsburgh, Pennsylvania. Journal of the Air and Waste Management Association. 52, 732-741.

Cancio J.L., Castellano A.V., Trujillo R.N., Rodriguez J.S. (2002) Present situation of air quality in the Jinamar Valley (Gran Canaria Island). Fresenius Environmental Bulletin. 11, 243-248.

Cantanho A.D.A., Artaxo P. (2001) Wintertime and summertime Sao Paulo aerosol source apportionment study. Atmos. Environ. 35, 4889-4902.

Cao L., Tian W.Z., Ni B.F., Zhang Y.M., Wang P.S. (2002) Preliminary study of airborne particulate matter in a Beijing sampling station by instrumental neutron activation analysis. Atmos. Environ. 36, 1951-1956.

Carter, W.P.L. (1990) A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. Atmos. Environ. Part A – General Topics. 24, 481-518.

Cardoni F., Lucarelli f., Mando P.A., Prati P., Zucchiatti A. (1998) Aerosol sampling and wind studies for the identification of pollutant sources. Journal of Wind Engineering and Industrial Aerodynamics 74, 327-334.

Cass G.R. (1998) Organic molecular tracers for particulate air pollution sources. Trends in Analytical Chemistry. 17, 356-366.

Chan Y.C., Simpson R.W., Mctainsh G.H., Vowles P.D., Cohen D.D., Bailey G.M. (1999) Source apportionment of PM2.5 and PM10 aerosols in Brisbane (Australia) by receptor modeling. Atmos. Environ. 33, 3251-3268.

Chen L.W.A., Doddridge B.G., Dickerson R.R., Chow J.C., Henry R.C. (2002) Origins of fine aerosol mass in the Baltimore-Washington corridor: Implications from observation, factor analysis, and ensemble air parcel back trajectories. Atmos. Environ. 36, 4541-4554.

Chiaradia M., Cupelin F. (2000) Gas-to-particle conversion of mercury, aresenic, and selenium through reactions with traffice related compounds (Geneva)? Indications from lead isotopes. Atmos. Environ. 34, 327-332.

Chock D.P. (1990) A comparison of numerical methods for solving the advection equation – III. Atmos. Environ. 25A, 853-871.

Chock D.P., Winkler S.L., Sun P. (1994) Comparison of stiff chemistry solvers for air quality modeling. Environ. Sci. Technol. 28, 1882-1892.

Chow J.C., Watson J.G., Pritchett L.C., Pierson W.R., Frazier C.A., Purcell R.G. (1993) The DRI Thermal/Optical reflectance carbon analysis system: description, evaluation, and applications in U.S. air quality studies. Atmos. Environ. 27A, 1185-1201.

Chow J.C., Watson J.G., Crow D., Lowenthal D.H., Merrifield T. Comparison of IMPROVE and NIOSH carbon measurements. Aerosol Science and Technology 34, 23-34.

Chow J.C., Watson J.G. (2002) Review of PM2.5 ane PM10 apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. Energy and Fuels 16, 222-260.

Chueinta W., Hopke P.K., Paatero P. (2000) Investigation of sources of atmospheric aerosol at urban and suburban residential areas in Thailand by positive matrix factorization. Atmos. Environ. 34, 3319-3329.

Claiborn C.S., Larson T., Sheppard L. (2002) Testing the metals hypothesis in Spokane, Washington. Environmental Health Perspectives 110, 547-552.

Colombo J.C., Landoni P., Bilos C. (1999) sources, distribution, and variability of airborne particles and hydrocarbons in La Plata area. Environmental Pollution. 104, 305-314.

Cooke W.F., Wilson J.J.N. (1996) A global black carbon aerosol model. Journal of Geophysical Research. 101, 19395-19409.

Currie L.A., Gerlach R.W., Lewis C.W., Balfour W.D., Cooper J.A., Dattner S.L., De Cesar R.T., Gordon G.E., Heisler S.L., Hopke P.K., Shah J.J., Thurston G.D., Williamson H.J. (1984) Interlaboratory comparison of source apportionment procedures: results for simulated data sets. Atmos. Environ. 18, 1517-1537.

Dabdub D., Seinfeld J.H. (1994) Numerical advective schemes used in air quality models – sequential and parallel implementation. Atmos. Environ. 28, 3369-3385.

DeBock L.A., VanMalderen H., VanGrieken R.E. (1994) Environ. Sci. Technol. 28, 1513-1520.

DeBock L.A., VanGrieken R.E., Camuffo D., Grime G.W. (1996) Microanalysis of museum aerosols to elucidate the soiling of paintings: Case of the Correr Museum, Venice, Italy. Environ. Sci. Technol. 30, 3341-3350.

DelCarmine P., Lucarelli F., Mando P.A., Valerio M., Prati P., Zucchiatti A. (1999) Elemental composition of size-fractionated urban aerosol collected in Florence, Italy; preliminary results. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 150, 450-456.

Diaz R.V., Aldape F., Flores M.J. (2002) Identification of airborne particulate sources, of samples collected in ticoman, Mexico, using PIXE and multivariate analysis. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 189, 249-253.

Dodge M.C. (2000) Chemical oxidant mechanisms for air quality modeling: critical review. Atmos. Environ., 34, 2103-2130.

Dzubay T.G., Stevens R.K., Balfour W.D., Williamson H.J., Cooper J.A., Core J.E., De Cesar R.T., Crutcher E.R., Dattner S.L., Davis B.L., Heisler S.L., Shah J.J., Hopke P.K., Johnson D.L. (1984) Interlaboratory comparison of receptor model results for Houston Aerosol. Atmos. Environ. 18, 1555-1566.

Ehrman S.H., Pratsinis S.E., Young J.R. (1992) Receptor modeling of the fine aerosol at a residential Los Angeles site. Atmos. Environ. 26, 473-481.

Einax J., Geiss S., Michaelis W. (1994) Fresenius Journal of Analytical Chemistry. 350, 614-619.

Fine P.M., Cass G.R., Simoneit B.R.T. (2001) Chemical characterization of fine particle emissions from fireplace combustion of woods grown in the northeastern United States. Environ. Sci. Technol. 35, 2665-2675.

Fine P.M., Cass G.R., Simoneit B.R.T. (2002) Chemical characterization of fine particle emissions from fireplace combustion of woods grown in the southern United States. Environ. Sci. Technol. 36, 1442-1451.

Foresman E.L., Kleeman M.J., Kear T., Niemeier D. (2002) PM10 conformity determinations: the Equivalent Emissions method. Transportation Research Part D. in press.

Fujimura M., Suzuki H., Sekine Y., Winchester J.W. (1993) Resolution of air-pollution from regional aerosol components in western Japan by factor-analysis of elemental concentrations measured by PIXE. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms. 75, 292-295.

Fung Y.S., Wong L.W.Y. (1995) Apportionment of air-pollution sources by receptor models in Hong-Kong. Atmos. Environ. 29, 2041-2048.

Gao N., Cheng M.D., Hopke P.K. (1994) Receptor modeling of airborne ionic species collected in SCAQS. Atmos. Environ. 28, 1447-1470.

Garcia-Alonso S., Perez-Pasto R.M., Quejido-Cabezas A.J. (2002) Chemometric study of selected polychlorinated biphenyls in ambient air of Madrid (Spain). Talanta. 57, 773-783.

Gasdag J. (1973) Numerical convection schemes based on accurate computation of space derivatives. J. of Computational Physics 13, 100-113.

Gebhart K.A., Malm W.C. (1997) Spatial and temporal patterns in particle data measured during the MOHAVE study. Journal of the Air and Waste Management Association. 47, 119-135.

Gharib S., Cass G.R. (1984) Ammonia emissions in the South Coast Air Basin, Open file Report 84-2, Environmental Quality Laboratory, California Institute of Technology, Pasadena CA.

Gordon G.E., William R.P., Daisey J.M., Lioy P.J., Cooper J.A., Watson J.G., Cass G.R. (1984) Considerations for design of source apportionment studies. Atmos Environ. 18, 1567-1582.

Green M.C., Flocchini R.G., Myrup L.O. (1992) The relationship of the extinction coefficient distribution to wind-field patterns in southern California. Atmos. Environ. 26, 827-840.

Hamester M. Stechmann H. Steiger M. Dannecker W. (1994) The origin of lead in urban aerosols – a lead isotopic ratio study. Science of the Total Environment. 147, 321-323.

Harrison R.M., Smith D.J.T., Luhana L. (1996) Source apportionment of atmospheric polycyclic aromatic hydrocarbons collected from an urban location in Birmingham, UK. Environ. Sci. Technol. 30, 825-832.

Harrison R.M., Smith D.J.T., Pio C.A., Castro L.M. (1997) Comparative receptor modeling study of airborne particulate pollutants in Birmingham (United Kingdom), Coimbra (Portugal) and Lahore (Pakistan). Atmos. Environ. 31, 3309-3321.

Henry R.C., Lewis C.W., Hopke P.K., Williamson H.J. (1984) Review of receptor model fundamentals. Atmos. Environ. 18, 1507-1515.

Henry R.C., Park E.S., Spiegelman C.H. (1999) Comparing a new algorithm with classic methods for estimating the number of factors. Chemometrics and Intelligent Laboratory Systems 48, 91-97.

Henry R.C., Kim B.M. (1990) Extension of self-modeling curve resolution to mixtures of more than 3 components 1. Finding the basic feasible region. Chemom. Intell. Lab. Syst. 8, 205-216.

Hien P.D., binh N.T., Truong Y., Ngo N.T. (1999) Temporal variations of source impacts at the receptor, as derived from air particulate monitoring data in Ho Chi Minh City, Vietnam. 33, 3133-3142.

Hien P.D., binh N.T., Trung Y., Ngo N.T., Sieu L.N. (2001) comparative receptor modeling study of TSP, PM2.5, and PM2-10 in Ho Chi Minh City. Atmos. Environ. 35, 2669-2678.

Hopke, P.K. (1988) Target transformation factor analysis as an aerosol mass apportionment method: review and sensitivity study. Atmos. Environ. 22, 1777-1792.

Hopke P.K., xie Y.L., Paatero P. (1999) Mixed multiway analysis of airborne particle composition data. Journal of Chemometrics. 13, 343-352.

Huang S.L., Rahn K.A., Armito R. (1999) Testing and optimizing two factor-analysis techniques on aerosol at Narragansett, Rhode Island. Atmos. Environ. 33, 2169-2185.

Huang S.L., Arimoto R., Rahn K.A. (2001) Sources and source variations for aerosol at Mace Head, Ireland. Atmos. Environ. 35, 1421-1437.

Jacob D.J. (1986) Chemistry of OH in remote clouds and its role in the production of formic acid and peroxymonosulfonate. J. Geophys. Res. 91, 9807-9826.

Jacob D.J. Gottlieb E.W., Prather M.J. (1989) Chemistry of a polluted boundary layer. J. Geophys. Res. 94, 12975-13002.

Jambers W., Dekov V., VanGrieken R. (1999) Single particle characterization of inorganic and organic North-sea suspension. Marine Chemistry. 67, 17-32.

Jeon S.J., Meuzelaar H.L.C., Sheya S.A.N., Lighty J.S., Jarman W.M., Kasteler C., Sarofim A.F., Simoneit B.R.T. (2001) Exploratory studies of PM10 receptor and source profiling by GC/MS and principal component analysis of temporally and spatially resolved ambient samples. Journal of the Air and Waste Management Association 51, 766-784.

Kavouras I.G., Koutrakis P., Cereceda-Balic F., Oyola P. (2001) Source apportionment of PM10 and PM2.5 in five Chilean cities using factor analysis. Journal of the Air and Waste Management Association. 51, 451-464.

Kavouras I.G., Koutrakis P., Tsapakis M., Lagoudaki E., Stephanou E.G., Von Baer D., Oyola P. (2001) Source apportionment of urban particulate aliphatic and polynuclear aromatic hydrocarbons (PAHs) using multivariate methods. Environ. Sci. Technol. 35, 2288-2294.

Kendall M., Hamilton R.S., Watt J. Williams I.D. (2001) Characterisation of selected speciated organic compounds associated with particulate matter in London. Atmos. Environ. 35, 2483-2495.

Kim K.H., Lee H.S., Youn Y.H., Yun S.T., Ro C.U., Oh J.M. (2002) Studies of spatial variabilities of airborne metals across four different land-use types. Water Air and soil Pollution. 138, 7-24.

Kleeman M.J., Cass G.R., Eldering A. (1997) Modeling the airborne particle complex as a source-oriented mixture. J. Geophys. Res. 102, 21355-21372.

Kleeman M.J., Cass G.R. (1998) Source contributions to the size and composition distribution of urban particulate air pollution. Atmos. Environ. 32, 2803-2816.

Kleeman M.J., Cass G.R. (1999) Identifying the effect of individual emissions sources on particulate air quality within a photochemical aerosol processes trajectory model. Atmos. Environ. 33, 4597-4613.

Kleeman M.J., Hughes L.S., Allen J.O., Cass G.R. (1999) Source contributions to the size and composition distribution of atmospheric particles: southern California in September 1996. Environ. Sci. Technol. 33, 4331-4341.

Kleeman M.J., Schauer J.J., Cass G.R. (2000) Size and composition distribution of fine particulate matter emitted from wood burning, meat charbroiling, and cigarettes. Environ. Sci. Technol. 33, 3516-3523.

Kleeman M.J., Schauer J.J., Cass G.R. (2000) Size and composition distribution of fine particulate matter emitted from motor vehicles. Environ. Sci. Technol. 34, 1132-1142.

Kleeman M.J., Cass G.R. (2001) A 3D Eulerian source-oriented model for an externally mixed aerosol. Environ. Sci. Technol. 35, 4834-4848.

Ko M.M., Jervis R.E. (1992) Atmospheric toxic metal concentrations from hospital incinerators. Journal of radioanalytical and nuclear chemistry-articles. 161, 159-170.

Kouimtzis T., Samara C., Voutsa D., Balafoutis C., Muller L. (2002) PCDD/Fs and PCBs in airborne particulate matter of the greater Thessaloniki area, N. Greece. Chemosphere. 47, 193-205.

Kulshrestha U.C., Kumar N., Saxena A., Kumari K.M., Srivastava S.S. (1995) Identification of the nature and source of atmospheric aerosols near the Taj-Mahal (India). Environmental Monitoring and Assessment. 34, 1-11.

Lawrence J., Koutrakis P. (1996) Measurement and speciation of gas and particulate phase organic acidity in an urban environment. 2. Speciation. Journal of Geophysical Research. 101, 9171-9184.

Lee E., Chan C.K., Paater P. (1999) Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmos. Environ. 33, 3201-3212.

Lee Y.L., Sequeira R. (2002) Water-soluble aerosol and visibility degradation in Hong Kong during autumn and early winter, 1998. Environmental Pollution. 116, 225-233.

Lee J.H., Yoshida Y., Turpin B.J., Hopke P.K., Poirot R.L., Lioy P.J., Oxley J.C. (2002) Identification of sources contributing to mid-Atlantic aerosol. Journal of the Air and Waste Management Association. 52, 1186-1205.

Lucarelli F., Mando P.A., Nava S., Prati P., Zucchiatti A. (1999) Elemental composition of size-fractionated urban aerosol collected in Florence, Italy. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 161, 819-824.

Maenhaut W., Salma I., Cafmeyer J., Annegarn J.J., Adreae M.O. (1996) Regional atmospheric aerosol composition and sources in the eastern Transvaal, south Africa and impact of biomass burning. Journal of Geophysical Research-Atmospheres. 101,23631-23650.

Maenhaut W., Fernandez-Jimenez M.T., Rajta I., Artaxo P. (2002) Two-year study of atmospheric aerosols in Alta Floresta, Brazil: Multielemental composition and source apportionment. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 189, 243-248.

Maenhaut W., De Ridder D.J.A., FernandezJimenez M.T., Hooper B., Nurhayati M. (2002) Long-term observations of regional aerosol composition at two sites in Indonesia. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 189, 259-265.

Manoli E., Voutsa D., Samara C. (2002) Chemical characterization and source identification/apportionment of fine and coarse air particles in Thessaloniki, Greece. Atmos. Environ. 36, 949-961.

Matsumoto E., Simabuco s.M., Perez C.A., Nascimento V.F. (2002) Atmosperic particulate analysis by synchrotron radiation total reflection (SR-TXRF). X-Ray Spectrometry 31, 136-140.

McRae G.J., Goodin W.R., Seinfeld J.H. (1982) Development of a 2nd generation mathematical model for urban air pollution. 1. Model formulation. Atmos. Environ. 16, 679-696.

Mendez J., Quejido A.J., Perezpastor R., Perezgarcia M. (1993) Chemometric study of organi pollution in the aerosol of Madrid. Analytica Chimica Acta. 283, 528-537.

Morandi M.T., Lioy P.J., Daisey J.M. (1991) Comparison of 2 multivariate modeling approaches for the source apportionment of inhalable particulate matter in Newark, NJ. Atmos. Environ. 25, 927-937.

Moro G., Lasagni M., Rigamonti N., Cosentino U., Pitea D. (1997) Critical review of the receptor model based on target transformation factor analysis. Chemosphere. 35, 1847-1865.

Mukai H., Furuta N., Fujii T., Ambe Y., Sakamoto K., Hashimoto Y. (1993) Characterization of sources of lead in the urban air of Asia using ratios of stable lead isotopes. Environ. Sci. Technol. 27, 1347-1356.

Mukerjee S., Shadwick D.S., Smith L.A., Somerville M.C., Dean K.E., Bowser J.J. (2001) Techniques to assess cross-border air pollution and application to a US-Mexico border region. Science of the Total Environment 276, 205-224.

Mysliwiec M.J., Kleeman M.J. (2002) Source Apportionment of Secondary Airborne Particulate Matter in a Polluted Atmosphere. Environ. Sci. Technol. In press.

Naeher L.P., Smith K.R., Leaderer B.P., Neufeld L., Mage D.T. (2001) Carbon monoxide as a tracer for exposures to particulate matter in wood and gas cookstove households of highland Guatemala. Environ. Sci. Technol. 35, 575-581.

Newmark G.L. (2001) Emissions inventory analysis of mobile source air pollution inTel Aviv, Israel. Energy, Air Qualityu, and Fuels 2001 Transportation Research Record. 1750, 40-48.

Nguyen K., Dabdub D. (2002) NOx and VOC controls and its effects on the formation of aerosols. Aerosol Science and Technology, 36, 560-572.

NIOSH (1996) Elemental carbon (diesel exhaust). In *NIOSH Manual of Analytical Methods*. National Institute of Occupational Safety and Health, Cincinnati, OH.

NIOSH (1999) Method 5040 Issue 3 (Interim): elemental carbon (diesel exhaust). In *NIOSH Manual of Analytical Methods*. National Institute of Occupational Safety and Health, Cincinnati, OH.

Nouchpramool S., Sumitra T., Leenanuphunt V. (1999) Characterization of airborne particulates in Bangkok urban area by neutron activation analysis. Biological Trace Element Research. 71, 181-187.

Odum J.R., Hoffmann T., Bowman R., Collins D., Flagan R.C., Seinfeld J.H. (1996) Gas-particle partitioning and secondary organic aerosol yields. Environ. Sci. Technol. 30, 2580-2585.

Oluyemi E.A., Asubiojo O.I., Oluwole A.F., Toussaint C.J.N. (1994) Elemental concentrations and source identification of air particulate matter at a Nigerian site – a preliminary study. Journal of radioanalytical and nuclear chemistry-articles. 179, 187-194.

Oluyemi E.A., Asubiojo O.I. (2001) Ambient air particulate matter in Lagos, Nigeria: a study using receptor modeling with X-ray fluorescence analysis. Bulletin of the Chemical Society of Etheopia. 15, 97-108.

Orlic I., Osipowicz T., Watt F., Tang S.M. (1995) The microanalysis of individual aerosol-particles using the nuclear microscope. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 104, 630-637.

Otten P., Injuk J., VanGrieken R. (1994) Elemental concentrations in atmospheric particulate matter sampled on the North-sea and the English Channel. Science of the Total Environment. 155, 131-149.

Paatero P. (1997) Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems. 37, 23-35.

Paatero P. (1999) The multilinear engine – a table driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. Journal of Computational and Graphical Statistics. 8, 854-888.

Paatero P., Hopke P.K., Song X., Ramadan Z. (2002) Understanding and controlling rotations in factor analytic models. Chemometrics and Intelligent Laboratory Systems. 60, 253-264.

Pandis S.N., Harley R.A., Cass G.R., Seinfeld J.H. (1992) Secondary organic aerosol formation and transport. Atmos. Environ. 26A, 2269-2282.

Paiva R.P., Munita C.S., Cunha I.L., Alonso C.D., Romano J., Martins M.H.R. (1993) A contribution to the characterization of the aerosol sources in Sao-Paulo. Journal of Radioanalytical and Nuclear Chemistry – Articles. 167, 295-307.

Pio C.A., Castro L.M., Cerqueira M.A., Santos I.M., belchoir F., Salgueiro M.L. (1996) Source assessment of particulate air pollutants measured at the southwest European coast. Atmos. Environ. 30, 3309-3320.

Polissar A.V., Hopke P.K., Paatero P. (1998) Atmospheric aerosol over Alaska – 2. Elemental composition and sources. Journal of Geophysical Research. 103, 19045-19057.

Polissar A.V., Hopke P.K., Poirot R.L. (2001) Atmospheric aerosol over Vermont: chemical composition and sources. Environ. Sci. Technol. 35, 4604-4621.

Pope, C.A. (2000) Review: Epidemiological Basis for Particulate Air Pollution Health Standards. Aerosol Science and Technology 32, 4-14.

Prati P., Zucchiatti A., Lucarelli F., Mando P.A. (2000) Source apportionment near a steel plant in Genoa (Italy) by continuous aerosol sampling and PIXE analysis. Atmos. Environ. 34, 3149-3157.

Pryor S.C., Steyn G.G., Rogak S.N. (1994) Source apportionment of visibility-degrading aerosols in the lower Fraser Valley, BC. Atmosphere-Ocean 32, 663-683.

Pryor S.C., Barthelmie r.J., Hoff r.M., Sakiyama S., Simpson R., Steyn D. (1997) REVEAL: Characterizing fine aerosols in the Fraser Valley, BC. Atmosphere-Ocean, 35, 209-227.

Qin Y., Oduyemi K., Chan L.Y. (2002) Comparative testing of PMF and CFA models. Chemometrics and Intelligent Laboratory Systems. 61, 75-87.

Ramadan Z., Song X.H., Hopke P.K. (2000) Identification of sources of Phoenix aerosol by positive matrix factorization. Journal of the Air and Waste Management Association. 50, 1308-1320.

Roberts G.C., Andreae M.O., Maenhaut W., Fernandez-Jimenez M.T. (2001) Composition and sources of aerosol in a central African rain forest during the dry season. Journal of Geophysical Research-Atmospheres. 106, 14423-14434.

Rocha T.A.P., Horvath H., Oliveria J.A.B.P., Duarte A.C. (1999) Trends in alkanes and PAHs in airborne particulate matter from Oporto and Vienna: identification and comparison. Science of the Total Environment 236, 231-236.

Rocha T.A.P., Duarte A.C. (1997) Application of chemometrics to the identification of trends in polynuclear aromatic hydrocarbons and alkanes in air samples from Oporto. Analyst. 122, 1509-1515.

Rojas C.M., VanGrieken R.E. (1992) Electron-microprobe characterization of individual aerosol particles collected by aircraft above the Southern Bight of the North-sea. (1992) Atmos. Environ. 26, 1231-1237.

Rojas C.M., VanGrieken R.E., Maenhaut W. (1993) Elemental composition of aircraft-sampled aerosols above the Southern Bight of the North-sea. Water Air and Soil Pollution. 71, 391-404.

Russell A.G., Winner D.A., Harley R.A., McCue K.F., Cass G.R. (1993) Mathematical modeling and control of the dry deposition flux of nitrogen-containing air pollutants. Environ. Sci. Technol. 27, 2772-2782.

Russell A., Dennis R. (2000) NARSTO critical review of photochemical models and modeling. Atmos. Environ. 34, 2283-2324.

Ruuskanen J., tuch T. Tenn Brink H., Peters A.,Khlystov A., Mirme A. Kos G.P.A., Brunekreef B., Wichmann H.E., Buzorius G., Vallius M., Kreyling W.G., Pekkanen J. (2001) Concentrations of ultrafine, fine, and PM2.5 particles in three European cities. Atmos. Environ. 35, 3729-3738.

Samara C., Kouimtzis T., Katsoulos G.A. (1994) Characterization of airborne particulate matter in Thessaloniki Greece. 2. A multivariate modeling approach for the source apportionment of heavy-metal concentrations within total suspended particles. Toxicological and Environmental Chemistry 41, 221-232.

Samara C., Kouimtzis T., Katsoulos G.A. (1994) Characterization of airborne particulate matter in Thessaloniki Greece. 3. Comparison of 2 multivariate modeling approaches for the source apportionment of heavy-metal concentrations within total suspended particles. Toxicological and Environmental Chemistry 44, 147-160.

Saucy D.A., Anderson J.R., Buseck P.R. (1991) Aerosol-particle characteristics determined by combined cluster and principal component analysis. Journal of Geophysical Research-Atmospheres. 96, 7407-7414.

Schauer J.J., Kleeman M.J., Cass G.R., Simoneit B.R.T. (1999a) Measurement of emissions from air pollution sources. 2. C-1 through C-30 organic compounds from medium duty diesel trucks. Environ. Sci. Technol. 33, 1578-1587.

Schauer J.J., Kleeman M.J., Cass G.R., Simoneit B.R.T. (1999b) Measurement of emissions from air pollution sources. 1. C-1 through C-29 organic compounds from meat charbroiling. Environ. Sci. Technol. 33, 1566-1577.

Schauer J.J., Kleeman M.J., Cass G.R., Simoneit B.R.T. (2001) Measurement of emissions from air pollution sources. 3. C-1-C-29 organic compounds from fireplace combustion of wood. Environ. Sci. Technol. 35, 1716-1728.

Schauer J.J., Kleeman M.J., Cass G.R., Simoneit B.R.T. (2002a) Measurement of emissions from air pollution sources. 4. C-1-C-27 organic compounds from cooking with seed oils. Environ. Sci. Technol. 36, 567-575.

Schauer J.J., Kleeman M.J., Cass G.R., Simoneit B.R.T. (2002b) Measurement of emissions from air pollution sources. 5. C-1-C-32 organic compounds from gasoline-powered motor vehicles. Environ. Sci. Technol. 36, 1169-1180.

Seigneur C. (2001) Current status of air quality models for particulate matter. J. of the Air & Waste Management Association. 51, 1508-1521.

Shannon J.D. (1999) regional trends in wet deposition of sulfate in the United States and SO2 emissions from 1980 through 1995. Atmos. Environ. 35, 807-816.

Sharma V.K., Patil R.S. (1992) Chemical-composition and source identification of Bombay aerosol. Environmental Technology. 13, 1043-1052.

Sharma V.K., Patil R.S. (1992) Size distribution of atmospheric aerosols and their source identification using factor analysis in Bombay, India. Atmos. Environ. 26, 135-140.

Sharma P.K., Singh G. (1992) Distribution of suspended particulate matter with trace-element composition and apportionment of possible sources in the Raniganj coalfield, India. Environmental Monitoring and Assessment. 22, 237-244.

Sheffield A.E., Gordon G.E., Currie L.A., Riederer G.E. (1994) Organic, elemental, and isotopic tracers of air pollution sources in Albuquerque, NM. Atmos. Environ. 28, 1371-1384.

Simoneit B.R.T. (2000) Biomass burning – a review of organic tracers for smoke from incomplete combustion. Applied Geochemistry 17, 129-162.

Sinha S., Banerjee S.P. (1997) Characterization of haul road dust in an Indian opencast iron ore mine. Atmos. Environ. 31, 2809-2814.

Slinn S.A., Slinn W.G.N. (1980) Predictions for particle deposition on natural waters. Atmos. Environ. 14, 1013-1016.

Smith L.A., Mukerjee S., Monroy G.J., Keene F.E. (2001) Preliminary assessments of spatial influences in the Ambos Nogales region of the US-Mexican border. Science of the Total Environment 276, 83-92.

Song X.H., Polissar A.V., Hopke P.K. (2001) Sources of fine particle composition in the northeastern US. Atmos. Environ. 35, 5277-5286.

Stevens R.K., Thompson G.P. (1984) Overview of the mathematical and empirical receptor models workshop (Quail Roost II). Atmos. Enviorn. 18, 1499-1506.

Sweet C.W., Vermette S.J., Landsberger S. (1993) Sources of toxic trace elements in urban air in Illinois. Environ. Sci. Technol. 27, 2502-2510.

Swietlicki E., Puri S., Hansson H.C., Edner H. (1996) Urban air pollution source apportionment using a combination of aerosol and gas monitoring techniques. Atmos. Environ. 30, 2795-2809.

Tang I.N., Munkelwitz H.R. (1994) Water activities, densities, and refractive indices of aqueous sulfates and sodium nitrate droplets of atmospheric importance. J. Geophys. Res., 99, 18801-18808.

Van Borm W.A., Adams F.C., Maenhaut W. (1990) Receptor modeling of the Antwerp aerosol. Atmos. Environ. 24, 419-435.

Van Malderen H., Rojas C., VanGieken R. (1992) Characterization of individual gian aerosol particles above the North-sea. Environ. Sci. Technol. 26, 750-756.

Van Malderen H., Van Grieken R., Khodzher T., Obolkin V., Potemkin V. (1996) Composition of individual aerosol particles above Lake Baikal, Siberia. Atmos. Environ. 30, 1453-1465.

Van Malderen H. Van Grieken r., Bufetov N.V., Koutzenogii K.P. (1996) Chemical characterization of individual aerosol particles in central Siberia. Environ. Sci. Technol. 30, 312-321.

Veltkamp P.R., Hansen K.J., Barkley R.M., Sievers R.E. (1996) Principal component analysis of summertime organic aerosols at Niwot Ridge, Colorado. Journal of Geophysical Research-Atmospheres 101, 19495-19504.

Voutsa D., Samara C., Kouimtzis T., Ochsenkuhn K. (2002) Elemental composition of airborne particulate matter in the multi-impacted urban area of Thessaloniki, Greece. Atmos. Environ. 36, 4453-4462.

Wang J., Zhu L.Z., Liu Y.J. (2002) Pattern of polycyclic aromatic hydrocarbons in communication air of Hangzhau, China. Journal of Environmental Sciences-China. 14, 145-150.

Watson J.G., Cooper J.A., Huntzicker J.J. (1984) The effective variance weighting for least squares calculations applied to the mass balance receptor model. Atmos. Environ. 18, 1347-1355.

Wolf M.E., Hidy G.M. (1997) Aerosols and climate: anthropogenic emissions and trends for 50 years. Journal of Geophysical Research. 102, 11113-11121.

Xhoffer D., Bernard P., VanGrieken R., VanDerauwera L. (1991) Chemical characterization and source apportionment of individual aerosol particles over the North-sea and the English-channel using multivariate techniques. Environ. Sci. Technol. 25, 1470-1478.

Yong T.R., Boris J.P. (1977) A numerical technique for solving stiff ordinary differential equations associated with the chemical kinetics of reactive-flow problems. J. Physical Chemistry, 81, 2424-2427.

Zelenka M.P., Wilson W.E., Chow J.C., Lioy P.J. (1994) A combined TTFA CMB receptor modeling approach and its application to air pollution sources in China. Atmos. Environ. 28, 1425-1435.

Zhang Y., Seigneur C., Seinfeld J.H., Jacobson M., Clegg S.L., Binkowski F.S. (2000) A comparative review of inorganic aerosol thermodynamic equilibrium modules: similarities, differences, and their likely causes. Atmos. Environ. 34, 117-137.

Zhu G.H., Wang G.F. (1998) Investigations of the particulate derived from indigenous zinc smelting using PIXE analytical technique. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms 137, 966-969.

Zieman J.J., Holmes J.L., Connor D., Jensen C.R., Zoller W.H., Hermann D.M., Parrington J.R., Gordon G.E. (1995) Atmospheric aerosol trace element chemistry at Mauna Loa observatory .1. 1979-1985. Journal of Geophysical Research-Atmospheres 100, 25979-25944.

List of Inventions Reported and Publications Produced

Kleeman M.J. A review of source apportionment techniques for airborne particulate matter. To be submitted, 2003.

Glossary of Terms, Abbreviations, and Symbols

CMB – Chemical Mass Balance Model. A linear statistical source apportionment tool that requires information about ambient particulate matter concentrations and source profiles to provide an accurate source apportionment of airborne particulate matter.

FA – Factor Analysis. A bilinear statistical source apportionment tool that requires information about ambient particulate matter concentrations to provide an approximate number of important sources and their approximate chemical source profile.

PCA – Principal Component Analysis. A bilinear statistical source apportionment tool that requires information about ambient particulate matter concentrations to provide an approximate number of important sources and their approximate chemical source profile.

PMF – Postive Matrix Factorization. A bilinear statistical source apportionment tool that requires information about ambient particulate matter concentrations to provide an approximate number of important sources and their approximate chemical source profile.

TTFA – Target Transformation Factor Analysis. A hybrid linear – bilinear statistical source apportionment tool that requires information about ambient particulate matter concentrations and some important source profiles to provide an approximate number of important sources and their approximate chemical source profile.