

# **Sources of PM and Precursor Gases in New York State**

Philip K. Hopke Center for Air Resources Engineering and Science Clarkson University Potsdam, NY

### Introduction



- In the next year, areas will be determined to be in non-attainment of the National Ambient Air Quality Standard (NAAQS) for Particulate Matter.
- As a result of these designations, it will be necessary to prepare state implementation plans (SIPs) that outline an air quality management strategy to bring the areas in question into attainment

### Introduction



- To ensure that an effective and efficient strategy is developed, it is important to identify the major sources of the particulate matter and the precursor gases that can be oxidized to produce additional  $PM_{2.5}$  mass.
- Receptor models applied to a variety of air quality data can help to develop that strategy by identifying the source types and by combining these results with back trajectory ensemble methods also identify the likely locations of those sources.



• Receptor models are focused on the behavior of the ambient environment at the point of impact as opposed to the source-oriented models that focus on the transport, dilution, and transformations that begin at the source and follow the pollutants to the sampling or receptor site.





#### **PRINCIPLE OF AEROSOL MASS BALANCE**

• The fundamental principle of receptor modeling is that mass conservation can be assumed and a mass balance analysis can be used to identify and apportion sources of airborne particulate matter in the atmosphere.

### Mass Balance



A mass balance equation can be written to account for all m chemical species in the n samples as contributions from p independent sources

$$\mathbf{x}_{ij} = \sum_{k=1}^{p} \mathbf{g}_{ik} \mathbf{f}_{kj}$$

Where i = 1,..., n samples, j = 1,..., m species and k = 1,..., p sources



• Need to identify the nature of the sources, how much they contribute to the measurement ambient particulate matter mass and where those sources are located.



• To identify the nature of the factors, the methods available are:

- Chemical Mass Balance
- Factor Analysis

- Factor Analysis
  - Principal Components Analysis
  - Absolute Principal Components Analysis
  - SAFER/UNMIX
  - Positive Matrix Factorization



Most factor analysis has been based on an eigenvector analysis. In an eigenvector analysis, it can be shown [Lawson and Hanson, 1974; Malinowski, 1991] that the equation estimates **X** in the least-squares sense that it gives the lowest possible value for

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \left( e_{ij} \right)^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj} \right)^2$$



The problem can be solved, but it does not produce a **unique** solution. It is possible to include a transformation into the equation.

#### X=GTT<sup>-1</sup>F

where T is one of the potential infinity of transformation matrices. This transformation is called a rotation and is generally included in order to produce factors that appear to be closer to physically real source profiles.



# Positive Matrix Factorization Clarkson

- Explicit least-squares approach to solving the factor analysis problem
- Individual data point weights
- Imposition of natural and other constraints, and
- Flexibility to build more complicated models



• The Objective Function, Q, is defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[ \frac{x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{\sigma_{ij}} \right]^{2}$$

where  $s_{ij}$  is an estimate of the uncertainty in  $x_{ij}$ 

# Application of PMF



• To illustrate what can be done with Positive Matrix Factorization, it will be applied to IMPROVE data from Brigantine, NJ





- A total of 910 samples collected between March 1992 and May 2001 and 36 species were used in this study.
- •The measured variables are:

•PM<sub>2.5</sub>, OC1, OC2, OC3, OC4, OP, EC1, EC2, EC3, S, NO<sub>2</sub><sup>-</sup>, NO<sub>3</sub><sup>-</sup>, Al, As, Br, Ca, Cl, Cl-, Cr, Cu, Fe, H, K, Mg, Mn, Na, Ni, P, Pb, Rb, Se, Si, Sr, Ti, V, Zn, Zr







Clarkson







### **Method description: PSCF**

- Potential source contribution function
  - Combination of particulate matter measurements with air parcel back trajectories to estimate regional source impact.
- Five-day back trajectories were reconstructed by the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model
  - The region covered by the trajectories was divided into 2664 grid cells of 1°×1° latitude and longitude





- If a trajectory endpoint lies in the *ij*th cell, the air parcel assumes to collect PM emitted in the cell and transports along the traj. to the monitoring site
- PSCF<sub>*ij*</sub> is the conditional probability that an air parcel that passed through the *ij*th cell had a high concentration upon arrival at the monitoring site

$$PSCF_{ij} = \frac{m_{ij}}{n_{ij}}$$

 $n_{ij}$ : total number of end points that fall in the *ij*th cell  $m_{ij}$ : number of end points that exceeded the threshold criterion (in this study, average concentration of each species was used for the threshold criterion)

#### Clarkson UNIVERSITY

### **PSCF**

- Small values of  $n_{ij}$  produce high PSCF values with high uncertainties
- To minimize the artifacts, PSCF values were downweighted with a arbitrary weight function (W) when n<sub>ij</sub> was less three times the average n<sub>ij</sub>

#### Clarkson UNIVERSITY

### **PSCF**

- PSCF describes the spatial distribution of probable geographical source locations
- Grid cells which have high PSCF values are the potential source area whose emissions can be contribute to the monitoring site
- For the secondary pollutant, the high PSCF area may also include areas where secondary formation is enhanced

### **PSCF: OC**



- Area of peak influence:
  - Southeast Hudson Bay indicates Quebec forest fire
  - Northeast of Lake Huron area could be additional fire zone
  - Pittsburgh to Michigan City

### **PSCF: EC**



- Area of peak influence:
  - Similar with OC results
  - Sources in Illinois, Missouri, and Iowa are uncertain

#### **Quebec Fire Location**



#### **PSCF: Sulfate**



- Area of peak influence:
  - Southern Indiana, Illinois & northern Kentucky
  - Midwestern coal fired power plant in Ohio River Valley

### **PSCF: PM**<sub>2.5</sub>



- Area of peak influence:
  - Ohio River Valley
  - Tailing into the Gulf of Mexico represents the increased influence of humidity on PM<sub>2.5</sub> mass
  - Southeast Hudson Bay in Quebec

# **PSCF: b**<sub>sp</sub>



- Area of peak influence:
  - Ohio River Valley
  - Gulf of Mexico indicates greater influence of humidity on b<sub>sp</sub> measurements
- No contribution from Quebec area due to the Nephelometer failure

### Data Relevant To NYS

IMPROVE Sites	NYC Supersite
Lye Brook, VT	
Underhill, VT	
Brigantine, NJ	
Speciation Network	Other Programs
South Bronx	Potsdam
Queens	Stockton
Botanical Garden	Hunter College
Rochester	Tuxedo
Buffalo	
Pinnacle State Park	
Whiteface Mountain	

• One other area of interest to use source apportionment is to relate observed adverse health effects to apportioned source contributions.



- EPA PM Centers have organized an intercomparison of receptor modeling methods and the relationship of the apportioned source contributions and the observed adverse health effects.
- Results were presented on May 28 and 29 and they are being compiled and analyzed now.

• There were statistically significant relative risks associated with the source contributions, but we need more time to examine the consistency and patterns of the relationships.

### Conclusions



- Good tools are available to help with the source identification and apportionment
- Method development continues and better tools can be expected in the near future
- Apportionment can assist in SIP development, and
- Potentially can be used to assist in health effects epidemiology