Data Analysis within EPA’s OAQPS: Thoughts on how source apportionment modeling and data analyses can be used to influence policymaking

MARAMA 2011 Data Analysis Workshop
January 20, 2011

Part 1: Source Apportionment Modeling – Kirk Baker
Part 2: Data Analyses – OAQPS Air Quality Analysis Group
Overview of Photochemical Modeling Source Apportionment

• Background
  – Photochemical models
  – Source Apportionment

• How can be used to influence policymaking
  – Contributions by State
  – Contributions by County
  – Contributions by Sector
  – Single source contributions
  – International contributions

• Future Work
Background: Photochemical Modeling

• Three-dimensional grid models simulate atmospheric processes including transport, chemistry, and deposition

• State of the science models simulate the formation and transport of many primary pollutants and secondarily formed pollutants such as **ozone** and **particulate matter**

• Many toxics are also treated in the model

• Used to replicate historical episodes of high pollution

• Used to assess atmospheric processing and transport of pollution

• Some photochemical models have been enhanced with source apportionment technology
Background: Model Source Apportionment

• Photochemical model source apportionment tracks the formation and transport of ozone and PM2.5 from specific emissions sources and allows the calculation of contributions to receptors

• Source groups may be single sources, groups of sources (sector), entire States, or entire Counties

• Receptors are each individual grid cell--which may be matched to any monitor located in the model domain

• This approach is different from other methods to estimate source contribution: observation based source apportionment (e.g., PMF, CMB), brute-force zero out simulations, decoupled direct method (DDM), and adjoint modeling

• Photochemical model source apportionment is very efficient for estimating culpability for many sources and does not perturb important atmospheric chemical processes
State Contributions

- State to State ozone and PM2.5 contributions used for NOx SIP Call, CAIR, & upcoming Transport rules
- State contribution examples below are for illustrative purposes only
  - Shown below: hourly ozone contribution to each grid cell from California (left) and Nevada (right)
County Contributions

- Ozone and PM2.5 designations support

- Source apportionment photochemical modeling for ozone designations tracked all anthropogenic NOX and VOC emissions from specific counties for contribution to ozone estimated at nonattainment monitor locations

- Contributions are the average of all modeled days at a receptor location where 8-hr max O3 is greater than 70 ppb

- Examples shown for Charlotte, Greenville, and Atlanta at right
Sector Contributions

- Source sector prioritization pilot project
- Support specific sector initiatives like cement kiln NESHAP
Sector Contributions

- 8-hr ozone, Daily PM2.5, annual PM2.5 design values
- Regional haze visibility baseline

---

SECA-C3 contribution to Annual PM2.5 Design Values

SECA-C3 contribution to Regional Haze Baseline Visibility
Short tangent: EPA efforts on ocean-going vessels

- Shipping-related emissions ~ 60,000 deaths annually
  - (Corbett et al, 2007; Winebrake and Corbett, 2008)
  - Air pollution from international shipping is difficult to regulation due to jurisdictional issues (Shrader, 2008)

- United Nations / International Maritime Organization established MARPOL convention (Annex VI) to prevent air pollution from ships

- U.S. promulgated MARPOL Annex IV into law July 08
  - Maritime Pollution Prevention Act of 2008; Allowed U.S. to apply for ECA designation

- Emissions Control Areas (ECAs)
  - Stringent fuel sulfur and engine NOx limits for ships operating in a designated ECA.
  - The U.S., and Canada applied for an ECA designation for specific portions of coastal waters in March 2009.
  - Final approval in April 2010
  - Further limits on US flagged vessels in April 2010

- Bottom line: significant reductions in Sox, NOx, and PM from this sector phased in between 2010 and 2020 (controls outpace growth)
Short tangent: Impacts of ECA on PM2.5 levels

- Large improvements in PM2.5 air quality are projected to occur as a result of an ECA designation
  - AQ benefit is largest in coastal areas, exceeding 1.0 ug/m³ annually in some locations
  - ECA designation is estimated to result in benefits ranging from $27-60 billion dollars (reduced health costs / premature mortality)
Single Source Contributions

- Tracked contribution from NOx, SO2, and primary PM2.5 emissions from 146 select point sources to support 24-hr PM2.5 designations process
- Some of these modeling results used in successful legal proceedings by US Dept. of Justice in their Wabash River and Cinergy cases
International Contributions

- Boundary (outside the modeling domain) contribution to annual average PM2.5 at bottom
- Source apportionment is useful for estimating international transport to key receptors
Concluding Remarks: Model Source Apportionment

- OAQPS plans to continue using CAMx for source apportionment of ozone, PM2.5, and mercury

- Currently working with US EPA ORD to implement source apportionment in CMAQ

- The CMAQ implementation will not be available until late 2011
Use of Data Analyses to Influence AQ Decisionmaking (Overview)

• Background
  – Air Quality Analysis Group in OAQPS
  – Membership and areas of expertise

• Cursory set of example of data analyses that have influenced policymaking
  – Source attribution via data analyses
    • Special Case Studies
    • Exceptional Events
    • Proximity Analyses
    • Greater use of GIS
  – Development of conceptual models of AQ formation
    • PM2.5 speciation
  – Data analyses to assist air monitoring
    • Analyses of co-located datasets
    • Tools to assist network assessments
  – Data analyses to assist AQ modeling
    • Diagnostic model evaluation
    • Spatial fusion of modeling/monitoring data
  – Air Data Access

• Concluding thoughts
Air Quality Data Analysis Group

• Group Responsibilities:
  – Analyzes air quality data for policy and program development
  – Analyzes complex air quality databases to identify patterns and understand cause-and-effect relationships for trend and accountability detection
  – Develops, maintains and reports findings from regulatory, ambient air quality databases (i.e., criteria pollutant design values) for attainment/nonattainment decision-making and program evaluation

• Current Group Members
  – Phil Lorang (Group Leader), Ellen Baldridge, Halil Cakir, Dennis Doll, Pat Dolwick (Acting GL), Barbara Driscoll, Mark Evangelista, Neil Frank, David Mintz, Adam Reff, Michael Rizzo (R5), Mark Schmidt, Rhonda Thompson, and Ben Wells
A “zone of influence” analysis starts with backward wind trajectories that show where an air parcel came from on its approach to the location of an odor complaint. Review of daily trajectories shows that people reported odors most frequently on days when surface oil was close to the coastline\(^1\). (based on methodology developed by R. Poirot)

1. [http://www.epa.gov/airtrends/studies.html](http://www.epa.gov/airtrends/studies.html) - Animation illustrates relationships between odor complaints, the proximity of the oil to the coastline, and the path of the winds.
As part of an on-going air toxics monitoring initiative, EPA, State and local air pollution control agencies are monitoring the outdoor air around schools for air toxics. A sample analysis from Marietta OH is shown below.
Identification of Exceptional Events

- States may request EPA to exclude data showing a violation of the NAAQS that are directly due to an exceptional event. Data analyses are required for three elements of the EE demo:
  - Demonstrating a clear, causal connection
  - Demonstration that there would have been no exceedance or violation but for the event.
  - Demonstration data is in excess of historical fluctuations.

- EPA guidance document is forthcoming:
  - How to submit complete technical analyses
  - Will be opportunity to comment prior to final (NAACA)

- For more information see:
  http://vista.cira.colostate.edu/improve/Activities/Meetings/2008Pres/F RANK_for_IMPROVE_mtg.pdf
Environmental Justice – Demographic Proximity Analysis

• Goal: to compare demographics of populations near emissions sources to the demographics of the whole US

• Census blocks - smallest areas for which demographics are available

• Problem: Which census blocks are “near” emissions sources?

• General Strategy
  – Choose a “proximity radius” within which nearby residents may be affected by local emissions sources
  – Calculate the distance between the “center” of each census block and each emissions source
  – Count “nearby” populations as those census blocks whose distance from one or more emissions sources is less than the proximity radius

Example: Census blocks which would be counted in this proximity analysis are highlighted in grey
Demographic Proximity Analysis Example:
Large Municipal Waste Combustion (LMWC) Facilities
Proximity Radius = 3 miles

<table>
<thead>
<tr>
<th>Facility</th>
<th>State</th>
<th>Population</th>
<th>White or Caucasian</th>
<th>African American</th>
<th>Native American</th>
<th>Other or Multiracial</th>
</tr>
</thead>
<tbody>
<tr>
<td>North County MWC</td>
<td>FL</td>
<td>17,637</td>
<td>11,205</td>
<td>5,322</td>
<td>28</td>
<td>1,262</td>
</tr>
<tr>
<td>Honolulu MWC</td>
<td>HI</td>
<td>2,376</td>
<td>558</td>
<td>37</td>
<td>6</td>
<td>1,775</td>
</tr>
<tr>
<td>Baltimore MWC</td>
<td>MD</td>
<td>219,153</td>
<td>78,913</td>
<td>130,787</td>
<td>732</td>
<td>8,721</td>
</tr>
<tr>
<td>Portland MWC</td>
<td>ME</td>
<td>43,390</td>
<td>40,562</td>
<td>650</td>
<td>150</td>
<td>2,028</td>
</tr>
<tr>
<td>Camden MWC</td>
<td>NJ</td>
<td>194,299</td>
<td>119,169</td>
<td>46,295</td>
<td>550</td>
<td>28,285</td>
</tr>
<tr>
<td>Essex MWC</td>
<td>NJ</td>
<td>204,501</td>
<td>103,859</td>
<td>40,550</td>
<td>823</td>
<td>59,269</td>
</tr>
<tr>
<td>Adirondack MWC</td>
<td>NY</td>
<td>26,154</td>
<td>25,590</td>
<td>145</td>
<td>54</td>
<td>365</td>
</tr>
<tr>
<td>Charleston MWC</td>
<td>SC</td>
<td>29,774</td>
<td>7,667</td>
<td>21,299</td>
<td>58</td>
<td>750</td>
</tr>
<tr>
<td><strong>Near Source Total</strong></td>
<td></td>
<td><strong>737,284</strong></td>
<td><strong>387,343</strong></td>
<td><strong>245,085</strong></td>
<td><strong>2,401</strong></td>
<td><strong>102,455</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of total</td>
<td>52.54%</td>
<td>33.24%</td>
<td>0.33%</td>
<td>13.89%</td>
</tr>
<tr>
<td><strong>National Total</strong></td>
<td></td>
<td><strong>285,339,128</strong></td>
<td><strong>214,539,706</strong></td>
<td><strong>35,043,873</strong></td>
<td><strong>2,489,515</strong></td>
<td><strong>33,083,760</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% of total</td>
<td>75.19%</td>
<td>12.28%</td>
<td>0.87%</td>
<td>11.66%</td>
</tr>
</tbody>
</table>

*** Figures based on 2000 census data. Other demographic categories include age (0-18 yrs, 65+ yrs), ethnicity (hispanic/latino), and income (below national poverty line)
Use of GIS in policy determinations

- Systematic linkage of geographic data sets with monitoring locations
  - Use GIS-based techniques to attempt to define site types and spatial scales.
  - Potential factors in making policy determinations:
    - Population/EJ
    - Landuse
    - Emissions
    - Vehicular traffic
    - Other demographic data
    - Dispersion/photochemical model data
    - Fused AQ surfaces
    - Health-effect information (e.g., ER visits)
Use of Data Analyses to Evaluate AQ Trends
(Meteorological Adjustments)

• Ozone
  – Currently, the model adjusts the seasonal (May – Sept.) mean of the daily maximum 8-hour average ozone concentrations for weather
  – Annual 4th highest daily maximum is of more interest to regulators
  – AQAG exploring a technique called quantile regression to adjust high-end concentrations for weather (4th highest daily max ~ 98th percentile)

• PM2.5
  – Currently, the model adjusts the annual mean of the daily 24-hour PM2.5 concentrations for weather by season (May-Sept. = warm, Oct.-April = cool)
  – AQAG exploring the development of individual species (sulfate, nitrate, etc.) models to better capture the seasonal variability
  – Could be coupled with quantile regression approach to adjust the 98th percentile for meteorology, for application to the 24-hour NAAQS
Use of Data Analyses to Identify Patterns (Conceptual Models)

- Quarterly analyses of fractional PM2.5 concentrations show patterns in AQ response (program accountability) as well as allow for the development of site-specific “conceptual models” to frame
Use of Data Analyses to Identify Patterns
(Conceptual Models / Accountability)

- **Trends in Elemental Carbon**
  - 18 sites, w/ at least 50 observations per year, 2002-2009
  - Did the trend change direction in 2007 because of the monitoring change?

- **Additional analyses confirmed decline in urban EC trends beginning ~ 2007.**
Use of Data Analyses to support AQ Monitoring
(Co-located monitors)

- Several comparisons of co-located continuous PM2.5 data against FRM instrumentation.
  - Multiple sets of continuous monitoring methodologies
  - Multiple years and locations

- Tendency for continuous PM2.5 data to show higher daily concentrations than filter-based FRM measurements under specific conditions.
Use of Data Analyses to support AQ Monitoring
(Network Assessments)

- A series of tools were created to aid State agencies in reviewing their ambient monitoring networks for the five year assessment
  - Concentrated on ozone and particulate matter
  - Provided information on site “redundancy” or possible “holes” in the network where more monitoring may be needed
  - Estimated the population served by individual monitors

- Tools will continue to evolve and improve so that other pollutants can be analyzed, users can input their own data, and analytical techniques can be updated
Air quality models are frequently used to estimate the impact of a change of emissions on air quality over a specified domain.
- Requires a performance evaluation.
- Four types of evaluations: operational, dynamic, diagnostic, and probabilistic.

This study conducted a diagnostic evaluation of the CMAQ model
- How well does the model capture changes in ozone concentrations related to meteorological variations?

CMAQ model tended to underestimate the sensitivity of ozone to temperature and relative humidity.
Use of Data Analyses to support AQ Modeling
(Combining modeling & measurements to estimate spatial patterns)

- Blends ambient monitoring data with CMAQ model values to create a “fused” surface which retains the information from the ambient data but applies the texture from the CMAQ data to areas particularly where monitoring data is scarce. This information can be used in conjunction with Federal/SIP modeling, permit modeling (background), and benefit assessments.
Air Data Access - Quick Reference Guide

This list of EPA’s locally managed federal air data systems is at http://www.epa.gov/air/airpolldata.html

- AQS: http://www.epa.gov/ttn/airs/airsaqs/index.htm
- AirNow: http://airnow.gov/
- AirNowTech: http://www.airnowtech.org/
- CASTNET: http://www.epa.gov/castnet/
- AirData: http://www.epa.gov/air/data/
- AirExplorer: http://www.epa.gov/airexplorer/
- AirCompare: http://www.epa.gov/aircompare/
- Air Trends: http://www.epa.gov/air/airtrends/
- AQS Data Mart: http://www.epa.gov/ttn/airs/aqsdatalmart/index.htm
- NATA: http://www.epa.gov/ttn/atw/natamain
- Air Emissions: http://www.epa.gov/air/emissions
- Ambient Toxics Data: http://www.epa.gov/ttn/amtic/toxdat.html
Data Analyses made easier via Air Data Access
(Exploring the impact of newly-revised AQ standards)

- One can plot SO2 and NO2 AQI values based on the newly revised AQI break points
  - [http://www.epa.gov/cgi-bin/htmSQL/airexplorer/trend_tile.hsql](http://www.epa.gov/cgi-bin/htmSQL/airexplorer/trend_tile.hsql)

- One can see how the number of Unhealthy days is affected by the proposed changes to the ozone standard (break points are hypothetical)
  - [http://www.epa.gov/cgi-bin/htmSQL/airexplorer/trend_aqi_temp.hsql](http://www.epa.gov/cgi-bin/htmSQL/airexplorer/trend_aqi_temp.hsql)

- One can get ozone W126 data summaries (not available anywhere else)
  - [http://www.epa.gov/ttn/analysis/w126.htm](http://www.epa.gov/ttn/analysis/w126.htm)

- One can view air quality monitor information in Google Earth and download daily and annual data
  - [http://www.epa.gov/airexplorer/monitor_kml.htm](http://www.epa.gov/airexplorer/monitor_kml.htm)

- One can view speciation data and SANDWICH estimates via AirExplorer (difficult to access elsewhere)
  - [http://www.epa.gov/cgi-bin/htmSQL/mxplorer/trend_spe.hsql](http://www.epa.gov/cgi-bin/htmSQL/mxplorer/trend_spe.hsql)
Concluding Thoughts

- There are a wide variety of instances (almost limitless) in which data analyses can be used to support air quality planning.
  - Great opportunity for collaboration: Regional organizations – divide and conquer
  - Prioritization is key (ideal world: match efforts to max health benefits)
  - Analysis efforts should be matched to problem scale

- Additionally, data analyses can be used to inform other important parts of the air quality planning process (emissions, monitoring, modeling, benefit assessments).

- The OAQPS Air Quality Analysis Group is open to working w/ State and Regional data analysts on issues of joint interest (noting everyone’s busy schedules). What can we do better to help you?
  - Data access?
  - Tool development?
  - Guidance?
  - Conferences/workshops?