EPA’s New Emissions Modeling Framework

Emission Inventory Conference
Las Vegas, Nevada  April 14th, 2005
Marc Houyoux
houyoux.marc@epa.gov
(next slide for additional credits)
Credits

This presentation was adapted from a longer version with the following developers and commenters:

- Emissions, Modeling, and Analysis Division (EMAD): Madeleine Strum, Norm Possiel, Rich Mason*, Doug Solomon, Phil Lorang, Tyler Fox
- Air Quality Strategies and Standards Division (AQSSD): Larry Sorrels, Art Rios, Darryl Weatherhead
- Office of Research and Development (ORD): Bill Benjey*, George Pouliot*, Dan Loughlin

* On assignment to EPA from NOAA, Atmospheric Sciences Modeling Division
Overview

- Motivations
- What is the Emissions Modeling Framework (EMF)?
- Hardware needs
- Expected uses in EPA
- Expected benefits to modeling community
- Status and timeline
Motivations

- Improve quality, consistency across programs, transparency and prevent mistakes in emissions modeling at EPA
  - For air quality modeling
  - For other purposes, e.g., Trends Reports, PM/Ozone brochures

- Better handle emissions modeling complexity
  - Many areas of expertise coming together into a “single” step (e.g., inventories, chemistry, meteorology, data handling)
  - Emissions modelers put together data sources from many disparate sources, but need a polished result
  - At EPA, different people do different parts (multi-user environment), sometimes across Divisions and Offices

- Provide a tool for the modeling community
What is the EMF?

- Software you access from your desktop, with:
  - Protocols for high-quality and timely emissions modeling (definable by user)
  - SMOKE GUI: Tool for setting up, managing SMOKE runs & run SMOKE in multi-computer environment
  - Tool for managing emissions modeling data in a multi-user environment

- Access point for emissions results

- Integrated criteria/PM/toxics emissions modeling platform for both grid-based and dispersion models

- QA will be tightly integrated with EmisView tool
Emissions Modeling Framework

New User Protocols
EMF interfaces
Database
EMF application
SMOKE-based computational software
SMOKE Improvements
Computational computers
Emissions Modeling Framework

NEI database

Emissions modeling database

Emissions modeling interface

Data QC & tracking

Data sharing

AQM-ready emissions

Creating grown & controlled ELs

Emissions summaries & graphics

Control programs database

Growth database
Data Management (1)

- Perform emissions modeling consistently across projects with diverse purposes
- Share data in a multi-user environment
  - Version control, problem tracking, record keeping for changes, metadata entry and review
  - Notification of data changes by “subscription”
- QA protocols: automate where possible and integrate closely with data
- Facilitate blending of multiple data sources
Data Management (2)

- Help prevent data handling and quality errors
- Store SMOKE emissions modeling data inputs in a shared database
  - Master data in central location
  - Multi-user approach allows accessing the same information across EPA
- Create spatial surrogates, speciation data, land use, & other SMOKE inputs from raw data
- Future: automated database coordination with EPA’s new NEI data system
Case Management (1)

- Case is a run to create a emissions modeling output
  - SMOKE input data (e.g., speciation or spatial inputs)
  - Grown and/or controlled inventory
  - Full set of model-ready inputs for entire episode

- Provide SMOKE interface with a much shorter learning curve than script-based approach
  - Setup based on real-world decisions, not SMOKE settings (e.g., user specifies what AQ model, what chemical mechanism, what time period)
  - Can run for all sectors and an entire episode in a single “case”
Case Management (2)

- Fosters reuse and coordination of emissions modeling cases
- Notify case creators of data updates by other users for data they are using
- QA protocols: automated and integrated with case setup and use
- SMOKE support for CMAQ, CAMx, REMSAD, ASPEN, and AERMOD
- More robust tools for future-year projection
Emissions Modeling Framework

What will the EM database include?

- Emissions modeling ancillary files (e.g., chemical speciation data, spatial surrogates)
- Emission modeling case configuration details (e.g., grid, time period, input files used)
- Some emissions summaries and reports for reuse
- Metadata for data and cases
  - Origin, history (including QA), & documentation (including links to documents), access history, purposes for use
  - Data status (including QA) & fate
  - Who can access file and for what purpose
- At least phase 1: Copies of relevant parts of inventories for emissions modeling (before new NEI data system)
EPA hardware configuration

Desktop PCs (Windows & Linux) at OAQPS, ORD

OAQPS Compute Servers

Application Server

Database Server

Emission Modeling Database

Growth Database

Controls Database

Other OAQPS Compute Servers (e.g., stand-alone Linux workstations)

Other Compute Servers in EPA (e.g., at ORD)

GUI

Data Server (node on cluster)

Shared Disk

Compute Server Head Node (Queue Manager)

Compute Nodes
Alternative hardware configuration

Desktop PCs (Windows & Linux)

Application, Database, and Compute Server

- Emission Modeling Database
- Growth Database
- Control Programs Database
- Compute Node(s)
- Large disk

Example of simpler hardware configuration supported by EMF flexible design
Expected uses in EPA

- Access by all of EPA to obtain emissions modeling data & information for both criteria and toxics
- EPA staff will be able to obtain descriptive information about emissions inputs for added transparency
- Economists and other analysts (not SMOKE experts) will be able to generate their own projected inventories
- EPA employees and contractors will be able to obtain and prepare data summaries of speciated and temporally allocated emissions
- Emissions modelers will be able to prepare, QA, and track emissions data for AQ models in a more transparent way
Expected benefits to community

- Although EMF is being developed first as an EPA resource, we are planning also to make it useable and available outside of EPA
- A new emissions data management resource
- A new emissions modeling QA resource
- Tools to create criteria/toxics speciation and spatial surrogate data independent of SMOKE
- Help run SMOKE
  - AQ model preparation (criteria/PM/toxics)
  - Future-year inventories
  - Data summaries and analysis
Project benefits timeline

- Initial Graphical QA tool
- More hardware for increased throughput
- EMF-Integrated Graphical QA tool
- EMF Public Beta Release
- EMF Public Release

- Protocols Complete
- SMOKE updates for 2002 NATA
- EMF: EPA Data management
- EMF: EPA SMOKE run management
- Speciation and Spatial allocation tools
Status

- Now accepting input, needs, and questions from outside EPA to improve end result
- High-level design is complete
- QA protocols in development by EPA staff
- Planned first applications are:
  - 2002 12-km criteria, PM, toxics CMAQ modeling
  - 2002 National Air Toxics Assessment
- Detailed design underway
- EmisView will have added functionality due to this effort
The research presented here was performed in part under a Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce’s National Oceanic and Atmospheric Administration (NOAA) and under agreement DW13921548. This work constitutes a contribution to the NOAA Air Quality Program.