4.5

THE EMISSIONS PROCESSING SYSTEM FOR THE ETA/CMAQ AIR QUALITY FORECAST SYSTEM

George A. Pouliot*
Atmospheric Sciences Modeling Division, ARL, NOAA
(On assignment to Office of Research and Development, US EPA)

1.0 INTRODUCTION

The U. S. Environmental Protection Agency in cooperation with the National Weather Service has developed a National Air Quality Forecasting System to provide timely forecasts of ozone, particulate matter, and other pollutants to prevent or reduce adverse effects of these hazards. The planned capabilities of the modeling system can be divided into the initial phase, the medium-term plan, and the long-term plan. The initial phase is to develop and validate 1-day forecasts of ozone (O_3) for the Northeastern US, and to expand nationwide within 5 years. The medium-term (5-7 years) plan is the development and deployment of a nationwide capability to forecast particulate matter concentration for particulate size less than or equal 2.5 microns (PM 2.5). In the long-term plan (within 10 years), the forecast range of the air quality system will be extended to 48-72 hours and include a broader range of pollutants.

The Initial Operating Capability (IOC) of the Air Quality Forecasting System, was deployed in September 2004 for the Northeast US, with one-day ozone forecasts of ground-level hourly ozone concentrations and eight-hour averages of ozone in parts per billion (ppb). Forecasts will be delivered twice daily. The National Weather Service (NWS) will be responsible for the operational forecasts, while the EPA will be developing the air quality modeling system. The NWS operational requirements specify that the post-processing of the modeled meteorological data, emission data processing, and the air quality chemistry model simulation be completed in less than five and a half hours. Since the post-processing of the meteorological data and the execution of the air quality chemistry model are typically very time consuming, the processing of the emissions needs to be minimized to less than 15 minutes per run. A typical emission-processing scenario in a research environment on a national domain for one day may take up to 3 hours of time on a single CPU. Therefore, in an operational environment, the emission processing needs to be streamlined with all non-meteorological dependent operations completed well in advance of the forecast. Only the meteorology dependent component of the emission processing needs to be done operationally. Thus, the accuracy of the emission processing can be maintained and the forecast can be completed within the required time constraints.

An emission processing system requires meteorological inputs, an emission inventory, and other ancillary input files. The ancillary information is used for temporal and spatial allocation, and for chemical speciation of the emission inventory to create time-dependent gridded input for an air quality model. An emission processing system requires both meteorology-dependent inputs (such as winds and temperature) as well as static non-meteorological inputs (spatial surrogates, land-use data, emission inventories). These inputs are used independently by different components of the emission processing resulting in many opportunities for parallelism. In an operational environment, the non-meteorological component of the emission processing can be separated from the meteorological dependent component, thus reducing the amount of processing that has to be done on an operational basis. The emission processing system used for the ETA/CMAQ system is the Sparse Matrix Operator Kernel Emission (SMOKE) system

*Corresponding Author Address: George A Pouliot, Mail Drop E243-04, Atmospheric Modeling Division, US EPA, Research Triangle Park, NC 27711 e-mail:pouliot.george@epa.gov

(Coats and Houyoux, 1996; Houyoux and Vukovich, 1999; Houyoux et. al., 2000; Community and Modeling Analysis System, 2004). SMOKE uses high performance computing sparse matrix algorithms, which permit rapid and flexible processing of emission data.

The meteorological dependent components of SMOKE have been incorporated into a single program that interfaces the output from the National Weather Service's ETA model with the Community Multiscale Air Quality (CMAQ) chemistry model. This program is called the preprocessor for CMAQ (PREMAQ).

2.0 EMISSION SOURCE INVENTORIES

In this section, we describe the four source categories of emissions used in the Air Quality Forecasting System and how the existing inventory for these categories were incorporated into the system. Area Sources are emission sources that have no specific latitude and longitude and are estimated on a county level. Biogenic Sources are naturally occurring sources of emissions coming primarily from vegetation. Point Source Emissions are from sources for which a specific location and information about a stack is known such as a power plant or a commercial industrial plant. Mobile Source emissions are those originating from all types of vehicles on roadways. Off-road mobile sources are included in the area source category.

2.1 Area Sources

The EPA has developed a 2001 inventory for the Clean Air Interstate Rule (CAIR) (US EPA, 2004). This inventory includes both projected emissions from the 1999 National Emission Inventory (NEI) version 3 inventory (EPA) and recalculated emissions. Source Categories that were recalculated for 2001 include residential wood combustion, paved and unpaved roads, residential, commercial, and roadway construction, mining and quarrying, open burning, structure fires, wild land fires and prescribed burning. All other area sources were projected from the 1999 NEIv3 to 2001 using growth factors (US EPA, 2004). The 1995 Canadian inventory was used with wildfires, prescribed burning, and windblown dust removed since these categories are year-specific. The 1995 Canadian area source inventory is the only Canadian inventory available at the EPA and was used for the Air Quality Forecast System without any modifications. Since these sources are not dependent on meteorology, they were calculated *a priori* using version 2.0 of SMOKE for all the days during the ozone season in 2004. Surrogates are used to spatially map all area sources to the grid. Examples include residential wood combustion, dry cleaning facilities, residential heating, paved and unpaved road dust, off-road machinery. PREMAQ merges the pre-computed area source emissions with the other source types.

2.2 Biogenic Sources

Since biogenic sources of emissions are dependent on meteorology, the Biogenic Emission Inventory System (BEIS) version 3.12 was included directly within PREMAQ. The Biogenic Emission Inventory System (BEIS3) is a major update to its predecessor BEIS2 and includes (1) a 1-km spatial resolution vegetation database for the contiguous United States, which resolves canopy coverage by tree species; (2) normalized emission factors for 34 chemical species, including 14 monoterpene compounds; (3) a soil nitric oxide emissions algorithm that accounts for soil moisture, crop canopy coverage, and fertilizer activity; and, (4) speciation factors for the CBIV, RADM2, and SAPRC99 lumped chemical mechanisms (Pierce et. al., 2002; Vukovich and Pierce, 2002)..

2.3 Mobile Sources

On-road Mobile Source emissions are typically modeled using EPA's motor vehicle emission model, MOBILE6 (US, EPA 2003). This model has been integrated into SMOKE. To accelerate the speed of processing, the temperature dependence of the mobile source emission estimate was separated from all other non-meteorological dependencies in the SMOKE/MOBILE6 system by running the SMOKE/MOBILE6 model retrospectively for a 2 month summer period in 2003 using temperature data from the ETA model. A nonlinear least-squares approach was used to create a relationship between temperature and emission rates for each species from SMOKE/MOBILE6 for each grid cell and for each hour of the week:

$$E(x,y,t,s) = f_s(x,y,t,T(x,y,t),T_0)\,. \quad (1)$$

In Equation (1), T is the temperature at the grid cell, T_0 is a reference temperature, s is a chemical species (e.g. NO, NO2, CO, or a VOC species) from the chemical mechanism of the air quality model, and *E* is the emission rate at a grid cell. The relationship between *E* and temperature is approximated by performing a nonlinear least-square fit to the temperature and emission rates. For all the chemical species, we use a second order polynomial for the function $f_s(x,y,t,T,T_0)$ at each grid cell for each hour of the week:

$$\begin{aligned} & f_s(x,y,t,T(x,y,t),T_0) = \\ & a_2(x,y,t)(T(x,y,t)-T_0)^2 \\ & + a_1(x,y,t)(T(x,y,t)-T_0) \\ & + a_0(x,y,t) \end{aligned} \quad (2)$$

Evaporative and exhaust VOC emissions are calculated by MOBILE6 for the following processes: diurnal emissions (emissions generated by the rise in temperature during the course of a day when the vehicle is not being driven); hot soak emissions (emissions occurring after the end of a vehicle trip, due to the heating of the fuel, fuel lines, and fuel vapors); running losses (emissions occurring while a vehicle is driven, due to the heating of the fuel and fuel lines); resting losses (small but continuous seepage and minor leakage of gasoline vapor through faulty connections, permeable hoses, and other components of the fuel system); and, refueling emissions (vapors that escape into the atmosphere when incoming liquid fuel displaces vapors in the vehicle fuel tank).

The nonlinear least-squares calculation is performed using the retrospective emission factors and the temperature field for each species and for each grid cell in the domain for a sufficient time period of available meteorological information. The coefficients a_0, a_1, a_2 for each grid cell and for each hour of the week are stored. These coefficients together with the 2 meter temperature from the ETA model are then combined to generate the operational estimate of the mobile source emission according to (1) and (2). This final emission calculation is extremely fast and is the only part of the mobile source emission processing that has to be performed operationally.

We chose a non-linear least squares approach because the relationship between emission rates and temperature was clearly not linear and the idea was to create a robust and stable method for estimating future emissions rate from a modeled forecast temperature. Other statistical methods may produce more accurate results or may be more efficient.

The nonlinear least-squares calculation is undertaken with MINPACK (Moré et. al., 1984). MINPACK is a free, portable library for solving nonlinear systems of equations and nonlinear least-squares problems. The specific algorithm from MINPACK uses an analytic specification of the Jacobian matrix and minimizes the sum of the squares of the nonlinear functions by a modification of the Levenberg-Marquardt algorithm (Moré, 1978). This algorithm is well suited to this problem since it allows the user to provide the subroutine, which calculates the analytic functions and the Jacobian of the analytic function.

2.4 Point Sources

For point sources, the 2001 Clean Air Interstate Rule emission inventory (US EPA, 2004) was used with one modification. The NO_x emissions from Electric Generating Units (EGU) were adjusted for the year 2004. Regulations to the electric utility in effect after 2001 have resulted in reductions in NO_x emissions from Electric Generating Units (EGUs). Since NO_x is a key precursor to the formation of ozone, it was important to project the NO_x emissions accurately from 2001 to 2004. The Energy Information Administration of the Department of Energy's Annual Energy Outlook (US, DOE 2004) presents forecast and analysis of US energy supply, demand, and prices from 2001 through 2025. This report estimates emissions on a regional basis for the Electric Utility sector for each future year. Since the base year for this projection was also the same year as the base inventory for the CAIR rule, the NO_x emission estimates contained in the Annual Energy Outlook for 2001 were very close to those used for the CAIR rule. Using the regional emission estimates for 2003 and 2004 contained in the Annual Energy Outlook, a projected inventory for EGU NO_x for 2003 and 2004 was

created using the ratio of emissions of 2004/2001 or 2003/2001. Figure 1 shows the regional projection of factors of NO_x derived from the Annual Energy Outlook for 2004.

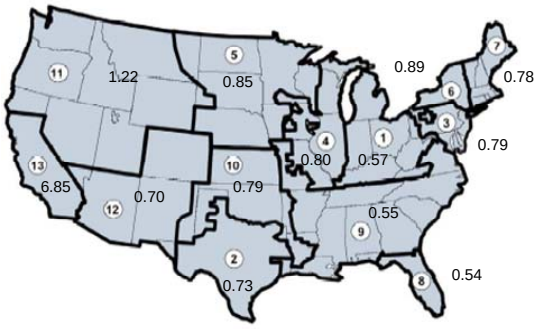

Figure 1 The Regional 2004 NOx adjustments applied to the 2001 EGU point source inventory Source: Department of Energy Annual Energy Outlook
http://www.eia.doe.gov/oiaf/aeo/index.html

3.0 Evaluation of Mobile Source Emission Estimates

The nonlinear least squares approach used in estimating the mobile source emissions was evaluated by comparing with estimates provided by the SMOKE/MOBILE6 system. These two emission estimates were then compared on a grid cell by grid cell basis for the 12Z July 20 forecast. Figure 2 shows a scatter plot of the NO_x emissions for a 48 hour time period for all non-zero grid cells in the operational NE domain. Figure 3 shows a scatter plot of the VOC emissions for the same 48 hour period. In both cases the correlation was very close to 1 and the normalized mean bias was less than 0.2%. Statistics for these two cases are summarized in the following table.

Table 1 Summary Statistics for 48 hour 7/20/04 12Z forecast of mobile source emissions

Species	Correlation (r)	Normalized Mean Bias	Count
NO_x	> 0.99	-0.0468%	840,912
VOC	> 0.99	0.1453%	840,912

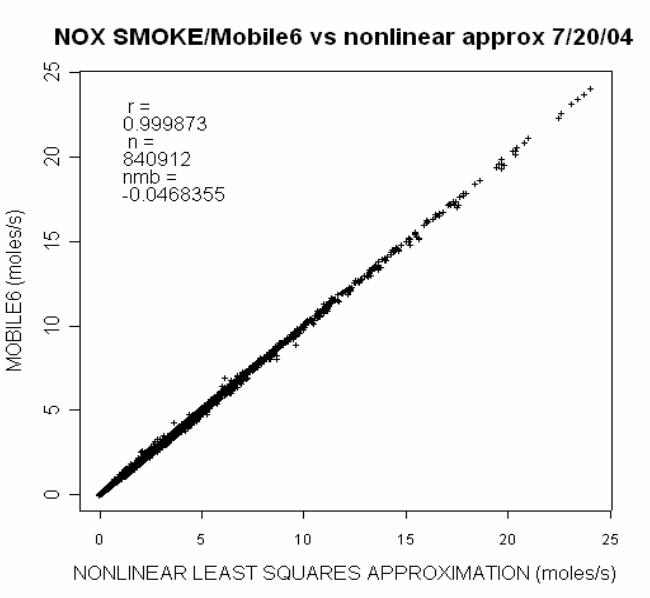

Figure 2: Scatter plot of SMOKE/MOBILE6 NOx emissions vs. operational AQF NOx emissions for the 48 hour forecast

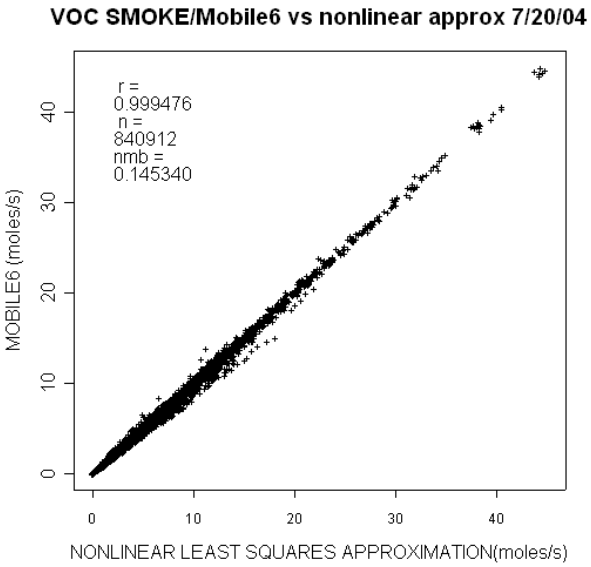

Figure 3: Scatter plot of SMOKE/MOBILE6 VOC emissions vs. operational AQF NO_x emissions for the 48 hour forecast

4.0 Evaluation of Point Source Emissions

Evaluation of the NO_x adjustments is not yet possible for 2004 since the Continuous Emission Monitoring (CEM) data for the entire year are not yet available. However, annual emission estimates for 2003 from the CEM data are

available from the Clean Air Markets Division (U.S. EPA, 2004). Using the same technique for 2004, a set of regional adjustment factors for 2003 were calculated for all the EGU sources in the continental US. These estimated annual NO_x emissions were then compared to the 2003 annual NO_x emissions measured at the facilities. A scatter plot of 758 facilities for 2003 is shown in figure 4. Only 758 out of 1448 facilities were easily matched with the inventory because of the difficulty of matching source identifiers in the NEI and CEM inventories. However these 758 facilities accounted for 96.6% of the total NO_x emissions from EGU sources for 2003. Using the DOE regional adjustments underestimated the NO_x emissions on a national basis by about 7% or about 300,000 tons of NO_x per year. The DOE NO_x projection to 2003 estimated that national EGU NO_x emissions would be reduced 22% from 4.75 million tons per year in 2001 to 3.7 million tons per year in 2003. However, based on the CEM data, national NO_x emissions from EGUs were actually reduced to 4.27 million tons per year in 2003, a reduction of only 10%. Since the DOE projection regions did not follow state boundaries and we approximated the projections at state boundaries, the DOE projection for 2003 did not match our 2003 estimate exactly. Our estimate for 2003 resulted in only a reduction of EGU NO_x emissions of 17% from 2001. If the 2001 emission inventory had been used without adjusting the EGU NO_x emissions, then EGU NO_x emissions would have been overestimated by 10% nationally. Using the DOE projection resulted in an underestimation of 7%.

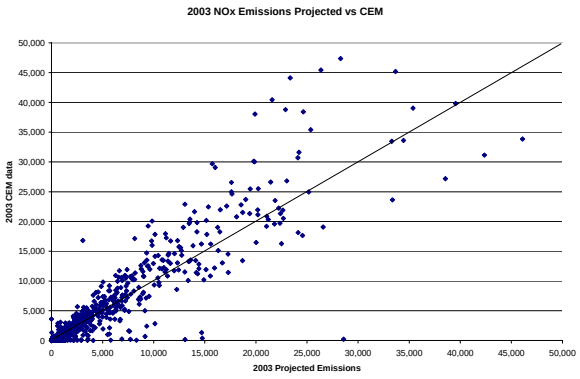

Figure 4: CEM annual NOx emissions compared to 2003 projected emissions for 758 EGU facilities in the continental US.

5.0 CONCLUSIONS

An emission processing methodology has been developed for NOAA/EPA's Air Quality Forecast System. Two major modifications have been made to traditional emission estimation schemes used in retrospective modeling. For mobile source emissions, we used a nonlinear least squares approximation to a more complex mobile emission model (MOBILE6) that accurately reproduces the emission estimates of NO_x and VOC. These two pollutants are important precursors to ozone formation. For EGU point sources, we applied projection factors for NO_x for both 2003 and 2004 that resulted in a 17% decrease in NO_x for 2003 and this compares reasonably well with the CEM measurement of a 10% NO_x reduction for 2003. We plan to evaluate of estimates for 2004 when CEM data for 2004 become available.

The research presented here was performed under the 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 number DW13921548. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

6.0 REFERENCES

Coats, C.J., Jr. and M. R. Houyoux, 1996: Fast Emission Modeling with the Sparse Matrix Operator Kernel Emission Modeling System, *Proceedings: AWMA Specialty Conference on the Emission Inventory: Key to Planning, Permits, Compliance, and Reporting*. New Orleans, LA, Air & Waste Management Assoc.

Community and Modeling Analysis System (CMAS) cited 2004: Sparse Matrix Operator Kernel Emissions (SMOKE) System. [Available online at http://www.cmascenter.org/html/models.html]

Houyoux, M.R. and J. M. Vukovich, 1999: Updates to the Sparse Matrix Operator Kernel Emission (SMOKE) Modeling System and Integration with Models-3. *Proceedings: AWMA Specialty Conference on The Emission*

Inventory: Regional Strategies for the Future, Raleigh, NC, Air & Waste Management Assoc.

Houyoux, M. R., J. M. Vukovich, C. J. Coats, Jr., N. M. Wheeler, and P. S. Kasibhatla, 2000: Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *J. Geophys. Res*., **105**, 9079-9090.

Moré, J.J., D. C. Sorensen, K. E. Hillstrom, and B. S. Garbow, 1984: The MINPACK Project, in *Sources and Development of Mathematical Software*, W. J. Cowell, ed., Prentice-Hall

Moré, Jorge J, 1978: The Levenberg-Marquardt Algorithm, Implementation and Theory. Numerical Analysis, G. A. Watson, ed. *Lecture Notes in Mathematics 630,* Springer-Verlag, 105-116

Pierce, T. C. Geron, G. Pouliot, E. Kinnee, and J. Vukovich, 2002: Integration of the Biogenic Emissions Inventory System (BEIS3) into the Community Multiscale Air Quality (CMAQ) Modeling System. Preprints *25th Conference on Agricultural and Forest Meteorology*, Norfolk VA, Amer. Meteor. Soc., CD-ROM, J5.11

U.S. DOE Energy Information Administration, 2004: Annual Energy Outlook 2004 with Projections to 2025. DOE/EIA-0383. [Available online at http://www.eia.doe.gov/oiaf/aeo/index.html.]

U.S. EPA Office of Air and Radiation, 2004: The Interstate Air Quality Rule Technical Documents. [Available online at http://www.epa.gov/air/interstateairquality/technical.html]

U.S. EPA Clean Air Markets Division, 2004: Emission data sets and published reports [Available online at http://cfpub.epa.gov/gdm/]

U.S. EPA Office of Transportation and Air Quality, 2003: User's guide to MOBILE6.1 and MOBILE6.2 (Mobile Source Emission Factor Model). EPA420-R-03-010,262 pp. [Available online at http://www.epa.gov/otaq/m6.htm]

Vukovich, J. and T. Pierce, 2002: The Implementation of BEIS3 within the SMOKE modeling framework. Proc. *11th International Emission Inventory Conference: Emission Inventories - Partnering for the Future,* Atlanta, GA, US EPA, CD-ROM, 10.7