Supporting Real-Time Air Quality Forecasting using the SMOKE modeling system

Jeffrey M. Vukovich, John N. McHenry, Carlie J. Coats, Atanas Trayanov
MCNC-Environmental Programs
PO Box 12889
Research Triangle Park, NC 27709
vukovich@mcnc.org

ABSTRACT

Emission inventories and associated data have been prepared for twice daily processing using the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system. The twice-daily emissions estimates from SMOKE were used in the Multiscale Air Quality Simulation Platform (MAQSIP) modeling system to produce daily ozone predictions for the eastern United States. The forecasts were generated for both regional and urban scale modeling domains for the summer of 2000. This paper provides an overview of the inventory data used, the issues encountered during preparation, the data flows to support the forecasting system, and the uses of the ozone forecasts.

SMOKE is used to process area, non-road, on-road mobile, biogenic and point sources data to create three dimensional gridded, hourly emissions files for input into MAQSIP. Meteorological forecast data are used as input for processing mobile, biogenic and point sources. This paper will discuss inventory improvements needed in the future, potential other uses for the forecasts, and future plans for forecasting the summer of 2001.

INTRODUCTION

The Multi-scale Air Quality Simulation Platform (MAQSIP)[1] has been used during the summers of 1998 through 2000 to generate ozone forecasts for the eastern United States. During these three summers, MAQSIP simulations have been conducted quasi-continuously for regional and urban scales. Generating these forecasts in a timely manner where results can be useful presents a tremendous challenge. Numerous tools to optimize model performance have been used to meet these demands. This paper will focus on one of these tools called the Sparse Matrix Operator Kernel Emissions (SMOKE)[2] modeling system.

The background section of this paper will summarize the SMOKE modeling system. The modeling data and issues section will discuss the input data used, modeling domains, and issues encountered when using the SMOKE modeling system. The applications section will summarize how the SMOKE/MAQSIP results were used during the summer of 2000. The future applications section will discuss plans for the use of the SMOKE modeling system for ozone forecasting for the summer of 2001 and other possible applications of the SMOKE/MAQSIP results.

BACKGROUND

In the recent past, a real limitation in air quality modeling applications has been the time required to process emissions inventories to produce data to be used in photochemical modeling. A prototype, emissions modeling system has been created to significantly reduce the time required to perform emissions preprocessing. The Sparse Matrix Operator Kernel Emissions (SMOKE) model was developed to demonstrate using matrix-vector multiplication for efficient emissions processing. It reproduces the core functions of emissions processing (i.e., spatial allocation, temporal allocation, chemical speciation, and growth and control of inventory emissions). For a five-day test case for a regional scale application, SMOKE has previously been shown to be 27 times faster than Emissions Modeling System '95 (EMS-95) for point sources and 38 times faster for area sources[3]. A test case comparison was made with the Emissions Preprocessing System version 2 (EPS2) versus SMOKE for an urban scale domain. SMOKE produced significant reductions in computing time and disk space usage[4]. The SMOKE modeling system prototype has been used in numerous regulatory and research applications. The computational performance of SMOKE has made the process of producing timely emissions forecasts for input into air quality models possible.

MODELING DATA AND ISSUES

The Real-Time Ozone Forecasting System for the year 2000 (RTOFS 2000)[5] used the SMOKE prototype modeling system for generating twice daily point, area/nonroad, biogenic and mobile source emission forecasts for use in MAQSIP simulations. This section will describe the modeling domains, input data used, data flows and the issues encountered while supporting the RTOFS 2000.

Modeling domains

The modeling domains used in the RTOFS 2000 include regional and urban scale domains. SMOKE was used to generate gridded three-dimensional emissions data for three modeling domains with horizontal resolutions of 45, 15 and 5 kilometers (Figures 1 and 2). The 45 km domain contains most of the eastern United States and extreme northwestern parts of Mexico. The 15km domain contains portions of eastern Texas, western Louisiana, southern Oklahoma and southwestern Arkansas. The 5km domain contains the eight-county Houston non-attainment area and the Beaumont-Port Arthur area. All modeling domains for the RTOFS 2000 used a Lambert conformal map projection.

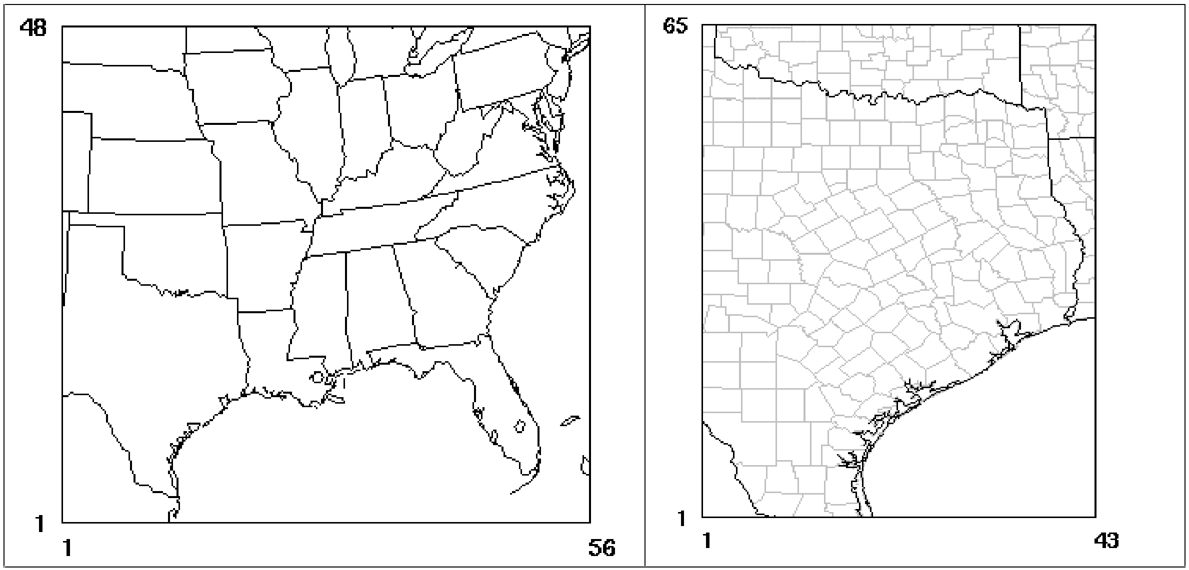

Figure 1. RTOFS2000 45 km resolution domain (left) and 15 km resolution domain (right).

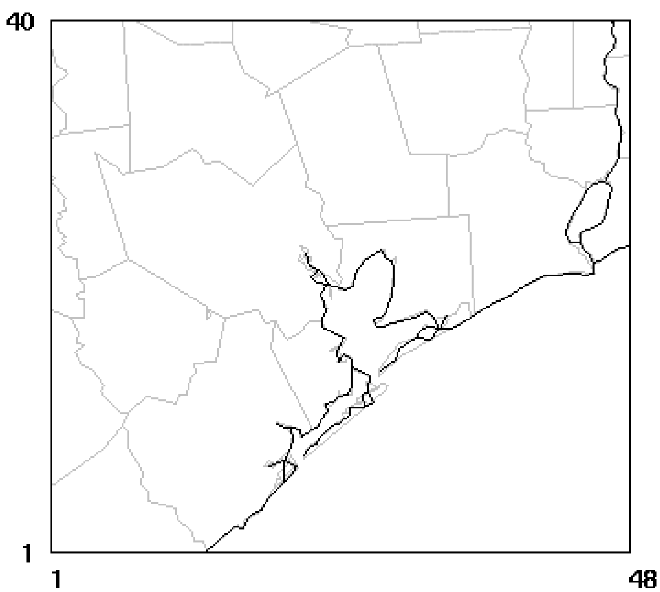

Figure 2. RTOFS2000 5km resolution domain centered over Houston, Texas.

Input Data

This section will describe the input data used in SMOKE to create gridded emissions. The meteorological data, emissions inventory and other associated data will be described.

Meteorology

The Penn State/National Center for Atmospheric Research Mesoscale Modeling System, version 2.12 (MM5)[6] was executed twice daily to generate meteorological forecasts to aide in the estimate of emissions and as a driver for the MAQSIP simulations. The MM5 was initialized at 00z and 12z each day. The MM5 modeling domains included a 45km domain that covered most of North America, the 15km domain covered the eastern two-thirds of the United States and the 5km domain covered Houston, Texas and surrounding counties. The 45km MM5 domain was executed for a 60-hour forecast, the 15km MM5 domain for a 24-hour forecast and the 5km MM5 domain for a 12-hour forecast. All MM5 modeling domains included 31 sigma-coordinate layers in the vertical. The meteorological forecast data was used to estimate and spatially allocate point source, mobile and biogenic emissions. For more information on the MM5 configuration for RTOFS 2000 refer to: http://envpro.mcnc.org/projects/SECMEP/secmep.html.

Emission Inventory

The point source inventory used in RTOFS 2000 consisted of the National Emissions Inventory (NEI) 1996 version 3.11[7], a Texas specific inventory, and a 1990 offshore source inventory. The area and non-road source inventory data were also acquired from the NEI 1996 version 3.11. The mobile source inventory data used in the Seasonal Model for Regional Air Quality (SMRAQ)[8] 1995 was also used for the RTOFS 2000. SMOKE was configured to use the MOBILE version 5b[9] when estimating mobile source emissions. SMOKE-Biogenic Emission Inventory System version 2 (BEIS2)[10] was used to estimate biogenic emissions. The land use data input into SMOKE-BEIS2 was acquired by aggregating the 1km horizontal resolution Biogenic Emissions Landcover Data version 3 (BELD3)[11]. These data were used because they were readily available and required little additional effort to process using SMOKE.

High-level data flows

The overall data flows of the RTOFS2000 are illustrated in Figure 3. The MERGE step involves merging the gridded emissions files created by SMOKE. Low-level data flows specific to SMOKE processing will be covered in the modeling issues section.

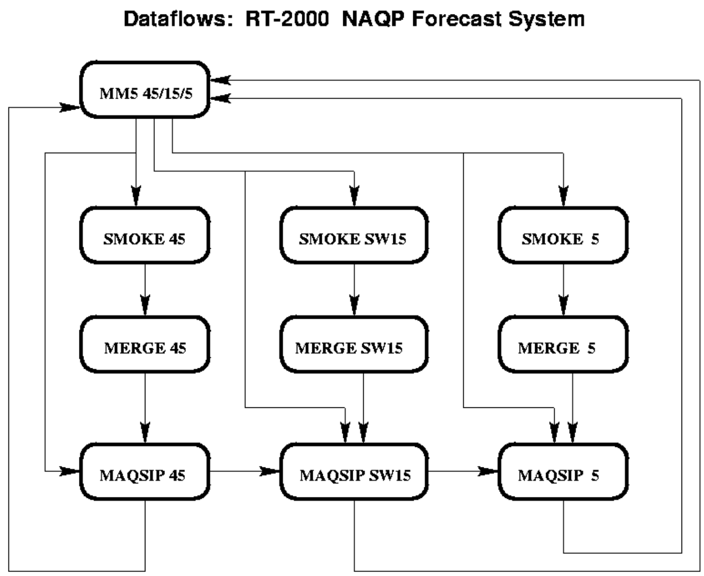

Figure 3. High-level data flows for RTOFS2000.

Modeling issues

SMOKE modeling issues for each emissions component will be described in this section. Specific data flows for each emissions component will be illustrated. One of the advantages of SMOKE is that much of data used in producing daily forecasts is static data. For example, the grid resolution and domain did not change over the forecast period so the gridding matrices only had to be calculated once. Then, the gridding matrix for each domain was used over and over again to generate emissions forecast for the associated domain. This same example can also be applied to the speciation matrix. The speciation mechanism in MAQSIP, Carbon IV[12], did not change over the forecast period. The only difference is that a speciation matrix must be created for each emissions component once. Assumptions made to further simplify the SMOKE processing are described for each emissions component. The emissions components covered in this section consist of the biogenic, area/non-road, on-road mobile and point source emission categories.

Point source inventory

The point source inventory consisted of emission inventory data in different formats. The NEI 1996 version 3.11 and offshore inventories had the EMS-95 format, but the Texas-specific inventory was in EPS2 format. The Texas point sources in the NEI 1996 inventory were removed before use with the SMOKE model. The SMOKE prototype system cannot process these inventories with different formats at the same time. Figure 4 illustrates all of the static and non-static data created and used by SMOKE while processing the point source data to produce each emissions forecast. Note that the temporal allocation factors (PTMP) have already been computed for a typical summer weekday, Saturday, Sunday and Monday. The Monday temporal allocation factors are necessary because the output from SMOKE must be in Greenwich Mean Time (GMT). These typical summer day temporal allocation factors were used throughout the summer so point source emissions were not allowed to vary week-to-week or month-to-month. The 1996 emissions data were not projected to the year 2000 due to limited resources.

The only programs needed to be executed daily to produce the emissions forecasts were the SMOKE point source programs LAYPOINT and CSGLDAYMRG. The program LAYPOINT spatially allocates point source emissions vertically using plume rise algorithms. The EMS-95 and EPS2 formatted inventories were processed separately through LAYPOINT and then CSGLDAYMRG to produce a three-dimensional, gridded point source data file for each inventory. The three-dimensional, gridded point source data were merged together with the other anthropogenic and biogenic gridded emissions data.

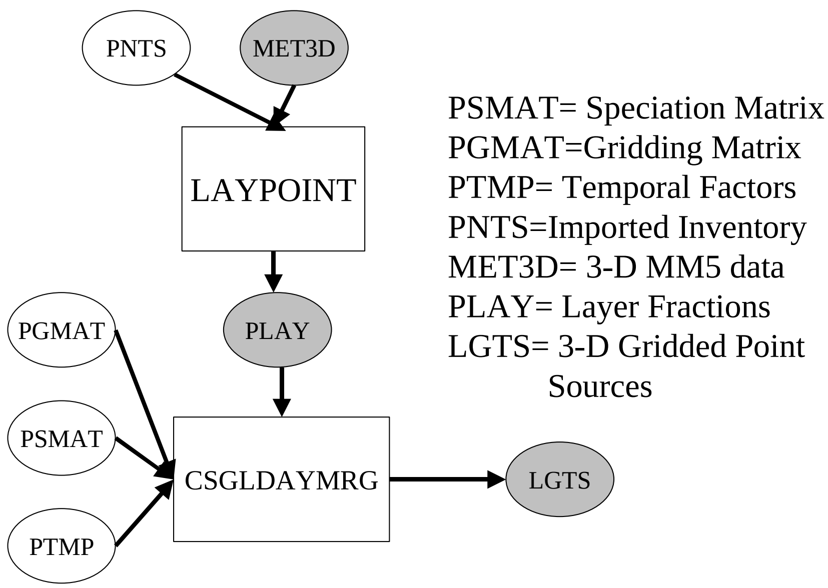

Figure 4. Data flows for point source emissions for RTOFS2000 (non-static data is shaded).

<u>Area/Non-road source inventory</u>

The area/nonroad source inventory consisted of the NEI 1996 v3.11 inventory in EMS-95 format. The area/nonroad sources were not dependent on the meteorological forecasts, so no area/nonroad-specific SMOKE program needed to be executed while producing real-time forecasts. Note that the temporal allocation factors have been already computed for a typical summer weekday, Saturday, Sunday and Monday for the same reasons mentioned in the point source inventory section. These typical summer day temporal allocation factors were used throughout the summer so area/non-road source emissions were not allowed to vary week-to-week or month-to-month. The 1996 inventory was not projected to the year 2000. Gridded, speciated, temporally allocated area/non-road source emissions were used in the merging of gridded emissions process for each emissions model forecast.

<u>Mobile source inventory</u>

The mobile source inventory consisted of the SMRAQ 1995 inventory which included MOBILE version 5b input data. 1995 vehicle miles traveled (VMT) data was used in RTOFS2000. The VMT data were not projected to the year 2000. Figure 5 displays all of the static data created by SMOKE while processing the mobile source data. There is a minor component to the mobile Volatile Organic Compounds (VOCs) emissions that is dependent on the daily minimum and maximum temperature. Since SMOKE would have had to wait until the meteorological forecast was completed before finding these temperature data, a method for estimating these VOC emissions was created for use in RTOFS2000. Average maximum and minimum temperatures were created for the modeling domains. These temperature data were allowed to vary spatially. For example, the maximum and minimum temperatures were not the same in Chicago as those used in Houston. These average maximum and minimum temperature data did not change over the summer long modeling effort.

The mobile emissions are dependent on temperature, so the MM5 forecasted temperature data was used to create mobile emissions forecasts. The program TMPMOBIL was executed using the MM5 temperature data and produced new temporal allocation factors to use for mobile emissions. The next program executed was CSGMOBIL which takes the new temporal allocation factors along with other static data (e.g. speciation and gridding matrices) to produce gridded mobile emissions.

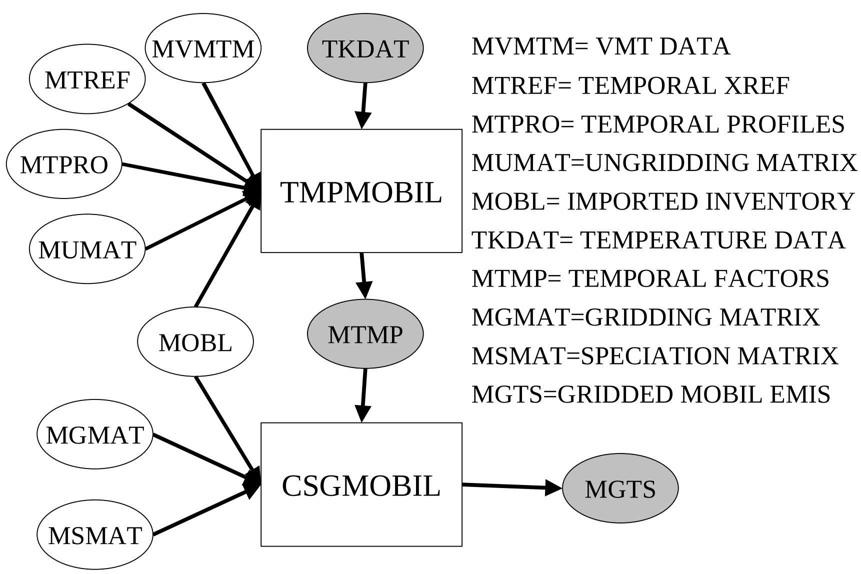

Figure 5. The data flows for mobile source emissions for RTOFS 2000 (non-static data is shaded)

Biogenic source inventory

The Biogenic Emissions Landcover Database version 3 (BELD3) consists of land use data at 1km resolution for 230 different land use types. BELD version 2 (BELD2) consists of mainly county land use data and 127 different land use types. The BELD2 land use types are those used in SMOKE-BEIS2 modeling. The modeling effort wanted to make use of the higher resolution BELD3 data while performing SMOKE-BEIS2 for RTOFS2000. So a cross-reference dataset was created to map the 230 BELD3 land use types to the 127 BELD2 land use types. This cross-reference dataset was used with a program called BELD3TO2 to generate 1km BELD2 formatted datasets. Then the CROSSBEIS program aggregated and projected the 1km datasets to produce SMOKE-BEIS2 ready land use files for each of the modeling domains. Figure 6 illustrates the programs and data flows for biogenic emission sources. The SMOKE program, RAWBIO was used to process the gridded landuse data to produce gridded, normalized emissions(BGRD). The only program required to be executed to produce the biogenic emissions forecasts was TMPBIO. The TMPBIO program uses the MM5 predicted first layer temperature (average height of 20 meters) and the solar radiation reaching the ground (Watts/m^2) to produce hourly, gridded biogenic emissions.

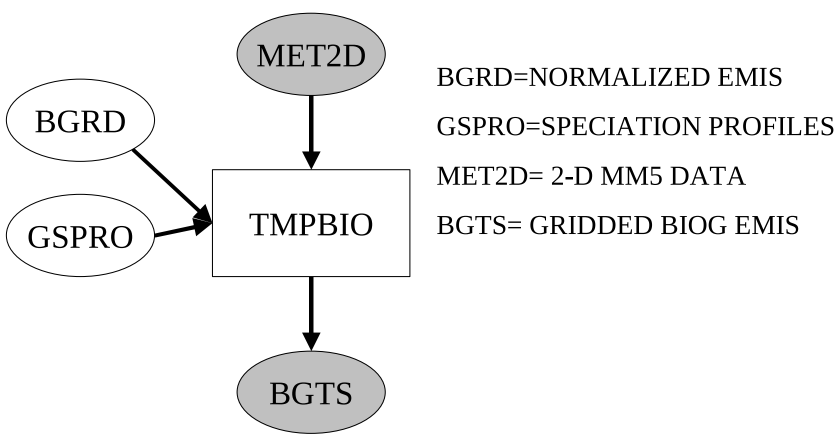

Figure 6. Data flows for biogenic emissions sources for RTOFS2000 (non-static data is shaded).

APPLICATION OF RESULTS

The RTOFS2000 modeling system described in the previous sections was used throughout the summer of 2000. The up-to-date MM5 and MAQSIP forecasts were displayed continuously for the 45km domain at the URL: http://envpro.mcnc.org/projects/SECMEP/secmep.html. The 15km and 5km domain forecasts were also available via a password protected website. These forecasts were used mainly during the Texas Air Quality study 2000 (TXAQS 2000) field program. The TXAQS 2000 program conducted an intensive monitoring program for Houston and the surrounding areas from August 15 through September 15, 2000. The field program consisted of an "all hands" meeting each morning at 7AM LST where many ozone-forecasting tools were analyzed. The 15km and 5km MAQSIP forecasts were one of these tools used by the monitoring program to determine where to send aircraft and to produce current and next day ozone forecasts. A preliminary evaluation on the MAQSIP forecasts during the TXAQS 2000 has been carried out [4]. The evaluation revealed the MAQSIP forecasts do have the potential to exhibit forecast skill at the 5km resolution.

FUTURE APPLICATIONS

The SMOKE prototype model has been used for real-time air quality forecasting purposes the past three summers (1998-2000). The SMOKE version 1 (SMOKEv1) was released in early 2000 and will be used as part of the RTOFS for the summer of 2001 (RTOFS2001). Up-to-date SMOKE emissions forecast data will be displayed graphically on the RTOFS2001 website. Updated inventory data will be acquired and evaluated to determine the level of effort needed for use with SMOKEv1. Efforts will be made to project inventory data to the current year of modeling. The SMOKE BEIS version 3 (SMOKE-BEIS3) prototype is undergoing an evaluation during the spring of 2001. If available for the summer of 2001, BEIS3 may be used for RTOFS2001. The MOBILE version 6 (MOBILE6) may be available for use sometime during the summer of 2001. Additional nested grids may also be added over the eastern United States to focus on other areas in need of ozone forecast data.

The RTOFS2000 data was used to help with flight plans for monitoring programs and for general forecast knowledge. Future applications of real-time ozone forecasting systems may also benefit other activities and the regulatory community. The continued use and evaluation of forecasting systems will improve our understanding of the strengths and weaknesses in air quality models. This increased understanding will most likely generate ideas on to improve the skill of the air quality models. Also, the speed of SMOKEv1 and the increased speed in computing hardware can allow for more than one emissions/air quality forecast to be produced. This would perhaps allow government agencies to simulate possible control strategies (e.g. Ozone Action Days) and be able to compare the actual air quality forecast versus these control strategy forecasts. This forecast information then may be able to introduce new ideas on how to produce lower ozone values for a certain short-term period. Some of these new short-term control strategies may be able to be extended to produce lower ozone values for longer periods of time (e.g. entire ozone season).

Other future applications of real-time forecasting systems will include modeling particulate matter (PM) and possibly toxics. A prototype PM forecasting application may be conducted sometime in the year 2002.

CONCLUSIONS

The SMOKE prototype model has been applied effectively to generate twice-daily emissions forecast for a real-time ozone forecasting system. The SMOKE application included producing emissions data for regional (45km and 15km) and urban scale (5km) modeling domains. SMOKE used MM5 forecast data while producing emissions forecasts for biogenic, point, and on-road mobile sources. The gridded emissions data created by SMOKE were input into MAQSIP along with MM5 data to generate the ozone forecasts.

The SMOKE/MM5/MAQSIP modeling system provided ozone forecasts over much of the summer of 2000. A special 5km horizontal resolution domain was used from August 15-September 15 for air quality field program in southeastern Texas. The MAQSIP forecasts were one of the tools used to help create flight plans for this air quality monitoring program. After preliminary evaluation, the MAQSIP forecasts do have the potential to exhibit forecast skill at urban scales.

Future applications of SMOKE in real-time air quality forecasting include adding recent emission inventory data, upgrading to SMOKEv1, and possibly including BEIS3 and MOBILE6. Other applications could include producing numerous emissions forecasts to produce multiple air quality forecasts to simulate possible control strategies (e.g. Ozone Action Days). The application of the SMOKE/MAQSIP to produce forecasts for other pollutants besides ozone might also be attempted in the near future.

REFERENCES

1) Odman, M.T., and C.L. Ingram, 1996: Multiscale Air Quality Simulation Platform (MAQSIP): Source code documentation and validation. *Technical Report ENV-96TR002-V1.0*, MCNC-North Carolina Supercomputing Center, RTP, NC

2) Houyoux, M., Vukovich, J., 1999: "Updates to the Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System and Integration with Models-3", Presented at The Emissions Inventory: Regional Strategies for the Future Conference, Air & Waste Management Association, Raleigh, NC.

3) Coats, C. J., Jr.; Houyoux, M. R., 1996: “Fast Emissions Modeling with the Sparse Matrix Operator Kernel Emissions Modeling System,” Presented at *The Emissions Inventory: Key to Planning, Permits, Compliance, and Reporting,* Air & Waste Mgt. Assoc., New Orleans, LA.

4) Houyoux, M.; Vukovich, J., 1998: “Emissions Inventory Development and Processing for the Seasonal Model for Regional Air Quality” Presented at *The Emission Inventory: Living in a Global Environment,* Air & Waste Mgt. Assoc., New Orleans, LA.

5) McHenry, J., Coats, C., Cameron, C., Vukovich, J., Trayanov, A., and T. Smith, 2001: “High-Resolution Real-Time Forecasts for the August-September Texas AQS-2000 (Houston) Field Study: Forecast Process and Preliminary Evaluation,” American Meteorological Society 81st Annual Conference, Albuquerque, NM.

6) Grell, G. A.; Dudhia, J.; Stauffer, D. R., 1994: “A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5),” National Center for Atmospheric Research, Boulder, CO, NCAR/TN-389+STR.

7) United States Environmental Protection Agency, National Emissions Trends (NET) ’96 Emissions Inventory, 2000: ftp://ftp.epa.gov/EmisInventory/net_96 .

8) Houyoux, M.R., J.M. Vukovich, C. Coats, N. Wheeler, and P. Kasibhatla, 2000: Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *JGR Atmospheres*, **105:D7**, 9079-9090.

9) United States Environmental Protection Agency, 1996: EPA/OMS MOBILE5 Vehicle Emission Modeling Software: MOBILE5b; URL: http://www.epa.gov/oms/m5.htm#m5b.

10) Vukovich, J., 1998: Development of Biogenic Emissions for the 1995 Summer Season. *Emission Inventories Conference*. December, New Orleans, LA, Air & Waste Management Association Conference.

11) Pierce, T., Kinnee, E. and C. Geron, 1998: Development of a 1-km resolved vegetation cover data base for regional air quality modeling. American Meteorological Society's 23rd Conference on Agricultural and Forest Meteorology, November 1998, Albuquerque, New Mexico.

12) Gery M.W., G.Z. Whitten, J.P. Killus, and M.C. Dodge, 1989: A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Res*., **94**, 12925-12956.

13) United States Environmental Protection Agency, 2001: EPA/OMS MOBILE6 Vehicle Emissions Modeling Software: MOBILE6, URL: http://www.epa.gov/otaq/m6.htm#overview.

KEYWORD

Emissions modeling
Ozone
SMOKE
Urban scale
Regional scale
Forecasting
Emissions Inventories