**APPENDIX 8-2: MODELING PROTOCOL FOR THE CENRAP 2002 ANNUAL EMISSIONS AND AIR QUALITY MODELING**

Draft 2.0

**Modeling Protocol for the CENRAP**
**2002 Annual Emissions and Air Quality Modeling**

Prepared for:

Annette Sharp
Technical Director
Central Regional Air Planning Association
10005 S. Pennsylvania, Ste. C
Oklahoma City, OK 73159

Prepared by:

Ralph E. Morris
Gerard Mansell
Bonyoung Koo

ENVIRON International Corporation
101 Rowland Way
Novato, California  94945

Gail Tonnesen
Mohammed Omary
Zion Wang

UC Riverside
CE-CERT
1084 Columbia Avenue
Riverside, California 92507

December 8, 2004

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

# 1.0 INTRODUCTION

This report constitutes the second draft of the Air Quality Modeling Protocol and Quality Assurance (QA) plan for the CENRAP 2002 annual Emissions and Air Quality Modeling activities to be performed by the contractor team of ENVIRON International Corp and the University of California at Riverside. Development of this second draft Modeling Protocol governing the CENRAP 2002 annual Emissions and Air Quality Modeling Study was performed as Task 2 of the study.

Note that at this writing we have not received all the relevant data and materials necessary to completely define the CENRAP annual modeling approach. Consequently, this draft Modeling Protocol will not necessarily be complete and will change as we obtain more information. The second draft Modeling Protocol incorporates comments from CENRAP on the first draft Modeling Protocol.

## 1.1 Background

CENRAP is one of five Regional Planning Organizations (RPOs) that have responsibility for coordinating development of State Implementation Plans (SIPs) and Tribal Implementation Plans (TIPs) in selected areas of the U.S. to address the requirements of the Regional Haze Rule (RHR). The RHR visibility SIPs/TIPs are due in 2007/2008. CENRAP modeling results may also form the regional component for 8-hour ozone and fine particulate ($PM_{2.5}$) SIPs/TIPs that are also expected to be due in 2007/2008. CENRAP is a regional partnership of states, tribes, federal agencies, stakeholders and citizen groups established to initiate and coordinate activities associated with the management of regional haze and other air quality issues within the CENRAP states. The CENRAP region includes states and tribal lands located within the boundaries of Arkansas, Iowa, Kansas, Louisiana, Minnesota, Missouri, Nebraska, Oklahoma and Texas. The CENRAP Emissions and Air Quality Modeling Team is comprised of staff from ENVIRON International Corporation (ENVIRON) and the University of California, Riverside (UCR). The ENVIRON/UCR Team performs the emissions and air quality modeling simulations for states and tribes within the CENRAP region, providing analytical results used in developing implementation plans under the EPA Regional Haze Rule. Figure 1-1 identifies the various Regional Planning Organizations in the U.S, including CENRAP. Table 1-1 lists the Class I areas within the CENRAP states.

The Clean Air Act establishes special goals for visibility in many national parks, wilderness areas, and international parks. Through the 1977 amendments to the Clean Air Act, Congress set a national goal for visibility as “the prevention of any future, and the remedying of any existing, impairment of visibility in mandatory Class I Federal areas which impairment results from manmade air pollution” (40 CFR 51.300). States are required to develop State Implementation Plans (SIPs) to attain visibility standards, and Tribes also may opt to assume responsibility for visibility programs under 40 CFR Part 49 by developing Tribal Implementation Plans (TIPs). The goal of the Regional Haze Rule (RHR) is to achieve natural visibility conditions at Federally mandated Class I areas by 2064. To achieve this goal, the RHR has set up several milestone years of 2018, 2028, 2038, 2048, 2058 and 2064 to monitor progress toward natural visibility conditions. Section 308 requires the first visibility SIP/TIP be submitted to

EPA by 2007 - 2008 to demonstrate progress toward natural visibility conditions in 2018 using the 2000-2004 five-year baseline. The CENRAP Long Range Plan (CENRAP, 2003) is to have an evaluation of initial strategies by July 2006 that States and Tribes can start their planning process for submitting their plans to EPA by December 2007.

Regional haze is linked to fine particulate ($PM_{2.5}$) for which EPA has a new standard. $PM_{2.5}$ SIPS are to be submitted 3 years after designation of $PM_{2.5}$ nonattainment areas. EPA intends to perform the final designations of the $PM_{2.5}$ nonattainment areas by December 2004, which would make the $PM_{2.5}$ SIP due by December 2007. As regional haze is intricately linked to $PM_{2.5,}$ the $PM_{2.5}$ and regional haze SIPs are in the process of being aligned so they would both be due by December 2007, if the $PM_{2.5}$ final designations occur as planned. EPA designated 8-hour ozone nonattainment areas in April 2004 that makes 8-hour ozone SIPs also due in 2007. States may use the integrated one-atmosphere CENRAP base year modeling for ozone modeling in addition to $PM_{2.5}$ and regional haze. Those decisions will be based on schedule constraints, model performance, appropriateness of the episodes, feasibility and need.


**Figure 1-1.** Regional Planning Organizations engaged in regional haze modeling.

**Table 1-1.** Federal mandated Class I areas in the CENRAP States.

| Class I Area | Acreage | Federal Land Manager | Public Law |
|---|---|---|---|
| **Arkansas** | | | |
| Caney Creek Wilderness Area | 4,344 | USDA-FS | 93-622 |
| Upper Buffalo Wilderness Area | 9,912 | USDA-FS | 93-622 |
| **Louisiana** | | | |
| Breton Wilderness Area | 5,000+ | USDI-FWS | 93-632 |
| **Minnesota** | | | |
| Boundary Waters Canoe Area Wilderness Area | 747,840 | USDA-FS | 99-577 |
| Voyageurs National Park | 114,964 | USDI-NP | 99-261 |
| **Missouri** | | | |
| Hercules-Glade Wilderness Area | 12,315 | USDA-FS | 94-557 |
| Mingo Wilderness Area | 8,000 | USDI-FWS | 95-557 |
| **Oklahoma** | | | |
| Wichita Mountains Wilderness | 8,900 | USDI-FWS | 91-504 |
| **Texas** | | | |
| Big Bend National Park | 708,118 | USDI-NP | 74-157 |
| Guadalupe Mountains National Park | 76,292 | USDI-NP | 89-667 |

## 1.2 CENRAP Organization Structure and Emissions and Air Quality Modeling

The governing body of CENRAP is the Policy Oversight Group (POG) that is made up of 18 voting members representing states and tribes within the CENRAP region and non-voting members representing local agencies, the EPA and other federal agencies. The work of CENRAP is accomplished through six standing workgroups:

- Monitoring;
- Emissions Inventory;
- Modeling;
- Communications;
- Implementation and Control Strategies; and
- International.

Participation in workgroups is open to all interested parties and the POG may form additional ad hoc workgroups to address specific issues (e.g., a Data Analysis workgroup was formed).

The Regional Haze Rule (RHR) requires the states, and the tribes may elect to, submit the first SIPs and TIPs in 2007/2008 that address progress toward natural conditions at federally mandated Class I areas. 40 CFR 51.308 (Section 308) discusses the following four core requirements to be included in SIPs/TIPs and Best Available Retrofit Technology (BART) requirements:

1. Reasonable progress goals;
2. Calculations of baseline and natural visibility conditions;
3. Long-term strategy for regional haze;
4. Monitoring strategy and other implementation plan requirements; and

5. BART requirements for regional haze visibility impairment.

One of CENRAP's goals is to provide support to states and tribes to meet each of these requirements of the RHR and to develop scientifically supportable, economical and effective control strategies that the states and tribes may adopt to reduce manmade effects on visibility impairment at Class I areas. One component of CENRAP's support to states and tribes as part of compliance with the RHR is performing emissions and air quality modeling to obtain a better understanding of the causes of potential mitigation measures for visibility impairment at Class I areas, to evaluate the effects of alternative control strategies for improving visibility and for projecting future-year air quality and visibility conditions. In October 2004, CENRAP selected a team of ENVIRON and UCR to perform their 2002 annual Emissions and Air Quality Modeling.

The CENRAP Emissions and Air Quality Modeling Team performs regional haze analyses by operating regional scale, three-dimensional air quality models that simulate the emissions, chemical transformations, and transport of gaseous and particulate matter (PM) species and consequently effects on visibility in Class I Areas in the central U.S. A key element of this work includes the integration of emissions inventories and models with regional transport models. The general services provided by the CENRAP Emissions and Air Quality Modeling Team include, but are not limited to:

- Emissions processing and modeling;
- Air quality and visibility modeling simulations;
- Analysis, display, and reporting of modeling results; and
- Storage/quality assurance of the modeling input and output files.

The CENRAP 2002 annual Emissions and Air Quality Modeling Team performs work for the CENRAP Modeling Workgroup through direction from the CENRAP Project Manager and CENRAP Administrative Project Manager.

## 1.3 CENRAP Long Range Plan

CENRAP adopted an initial Long Range Plan in October 2003 (CENRAP, 2003) and will review this plan each spring and fall and update it as needed. The CENRAP Long Range Plan is organized into four primary tiers of work efforts: (1) Tasks; (2) Activities; (3) Projects; and (4) Deliverables.

### 1.3.1 Tier One: Tasks

The first tier of the CENRAP Long Range Plan consists of a series of work efforts or Tasks that comprise the statuary requirements of the Regional Haze Rule (RHR):

- Establish baseline and natural conditions;
- Develop reasonable progress goals;
- Develop a long term strategy for reducing regional haze;
- Develop a monitoring strategy and other SIP/TIP requirements; and

- Establish BART requirements.

Each of these Tasks must be accomplished to achieve the goal of providing states and tribes the information upon which to develop supportable and effective RHR SIPs and TIPs.

### 1.3.2 Tier Two: Activities

The second tier (Activities) consists of broad, categorized work efforts that must be completed in order to complete a Task. For example, the 2002 annual Emissions and Air Quality Modeling effort is an activity that needs to be completed in order to perform the "develop reasonable progress goals" and "develop a long term strategy for reducing regional haze" Tasks listed above.

### 1.3.3 Tier 3: Projects

The third tier consists of specific Projects aimed at completing an Activity. For example, to perform the 2002 annual modeling, 2002 emissions are needed, thus a specific project would be to compile 2002 emissions for the CENRAP states and provide to the Emissions and Air Quality Modeling Team. Details on specific projects are identified in Attachment 1 of the CENRAP Long Range Plan (CENRAP, 2003).

### 1.3.4 Tier Four: Deliverables

The fourth and final tier is the deliverables that consist of the individual components of the work effort necessary to complete a given project. For example, Deliverables for the 2002 annual Emissions and Air Quality Modeling activity would include a Modeling Protocol, a QAPP, and Base Case Modeling and Model Performance Evaluation Report and the modeling databases.

### 1.3.5 Critical Milestone Dates for CENRAP

The CENRAP Long Range Plan has developed a series of milestone dates by which critical decisions must be made to address a regulatory or statutory deadline. Critical milestone dates were assigned to the following activities:

- SIP submittal;
- SIP and TIP drafting and approval;
- Conducting future-year modeling; and
- Conducting base case modeling.

The CENRAP Long Range Plan back calculated critical milestone dates from the date that the RHR SIPs/TIPs are required to be submitted to EPA and developed the following timeline:

SIP/TIPs Adopted and States Submit SIP to EPA – December 31, 2007: Although the actual SIP/TIP submittal date may be extended to 2008, CENRAP is adopting the December 2007 submittal date in developing their critical milestone dates to assure that CENRAP obtains the information necessary to develop the RHR SIP/TIPs in a timely fashion.

States and Tribes Begin to Draft and Adopt SIPs and TIPs – July 1, 2006: A survey of the timelines for SIP approval for states in the CENRAP region indicates that the states can draft and adopt a SIP as long as the technical analysis has been completed 18 months in advance.

Future-Year Modeling Begins – March 1, 2005: The CENRAP Modeling Workgroup estimates that it will take fifteen months to complete the future-year modeling and to conduct and evaluate alternative future-year control strategies. Thus, the base year base case and model performance evaluation needs to be completed and the future-year modeling initiated by March 1, 2005.

Base Case Modeling Begins – October 1, 2004: The CENRAP Modeling Workgroup estimates that it will take six months to perform base year base case modeling and model performance evaluation.

### 1.3.6 Role of Emissions and Air Quality Modeling in CENRAP Long Range Plan

As seen above, the 2002 annual Emissions and Air Quality Modeling is a critical part of the overall CENRAP efforts and is an essential component for the development of the RHR SIPs and TIPs. The modeling information must be of high quality and reliability in order to develop effective RHR control strategies. Thus, comprehensive and exhaustive quality assurance (QA) and quality control (QC) techniques are necessary.

## 1.4 Past Related Regional Modeling Studies

The CENRAP 2002 annual Emissions and Air Quality Modeling activities are built off of previous regional emissions, photochemical PM and visibility modeling performed in the Central States and across the United States. The procedures used in these previous studies help guide the design and form the initial basis for the plan for the CENRAP annual modeling approach. We are in the process of reviewing these other studies and data and not all information have been fully assimilated in this draft Modeling Protocol. Information considered will include, but not be limited to, the following:

Big Bend Regional Aerosol sand Visibility Observational Study (BRAVO): The BRAVO study examined the causes and source of regional haze at the Big Bend National Park, the most southwesterly Class I area in the CENRAP states. It performed data collection activities, modeling and used numerous techniques to estimate PM source apportionment (Pitchford et al., 2004).

CENRAP Scoping Study: CENRAP commissioned a scoping study to identify the causes of visibility impairment at Class I areas in the CENRAP states and identify the analytical tools that are available to investigate regional haze (Green et al., 2002).

CENRAP Ammonia Emissions Inventory Study: CENRAP sponsored a study to develop an improved ammonia emissions inventory for the CENRAP states (Coe and Reid, 2003).

CENRAP Agricultural and Prescribed Burns Study: In this study improved emissions inventories for prescribed burns and agricultural burning were developed for the CENRAP states (Reid et al., 2004a).

Evaluation of CMAQ and CAMx Models Over the CENRAP States for Three Episodes: CMAQ and CAMx model simulations of a January 2002, July 1999 and July 2001 episodes were evaluated using measurement data in the CENRAP states (Tonnesen and Morris, 2004).

Development of Enhanced Mobile Source and Agricultural Dust Emissions for CENRAP: This study developed on-road and non-road mobile source and agricultural dust emission inventories for the CENRAP states (Reid et al., 2004b).

Development of 2002 Base Case Modeling Inventory for CENRAP: CENRAP sponsored this study to prepare a 2002 Base Case emissions inventory for the CENRAP states that can be used in emissions and photochemical modeling of the 2002 annual period (Strait, Roe and Vukovich, 2004).

Preliminary PM and Visibility Modeling for CENRAP: Under this study preliminary regional PM and visibility modeling was conducted focused on the CENRAP region using the CMAQ and CAMx models (Pun, Chen and Seigneur, 2004).

VISTAS Phase I Model Sensitivity and Evaluation Study: This study, sponsored by VISTAS, performed extensive model sensitivity testing and evaluation analysis using the CMAQ and CAMx models and three episodes, January 2002; July 1999 and July 2001 (Morris et al., 2004a).

WRAP Section 309 SIP/TIP Modeling Analysis: The WRAP performed a study to generate the necessary modeling data needed to develop Section 309 SIP/TIP for states that opt-in to this program (Tonnesen et al., 2003).

VISTAS Phase II 2002 Annual Modeling: VISTAS is performing annual modeling of 2002 using a continental US 36 km domain and eastern US 12 km domain with attendant model evaluation and sensitivity analysis (Morris et al., 2004b).

MRPO Modeling and Analysis: The Midwest RPO is also performing regional haze modeling and analysis that will be integrated into the CENRAP modeling (Baker, 2004).

IDNR 2002 MM5 Modeling: The Iowa DNR is performing 2002 MM5 modeling that will be used in the CENRAP annual modeling (Johnson, 2004).

EPA VII MM5 Modeling: EPA Region VII is performing 12 km MM5 modeling for the Central States and portions of 2002 that will be used in the CMAQ/CAMx modeling analysis.

Many of the above studies above are providing data (e.g., emissions) that will be used directly in the CENRAP Emissions and Air Quality Modeling Study. Consequently, the quality assurance (QA) and quality control (QC) procedures are directly relevant to this QAPP. Others are companion modeling studies (e.g., BRAVO, VISTAS and WRAP) that provide information that is used in the development of this QAPP and the CENRAP Emissions and Air Quality Modeling Protocol (ENVIRON and UCR, 2004).

## 1.5 Overview of 2002 Annual Emissions and Air Quality Modeling Approach

The CENRAP 2002 annual modeling will include annual PM/regional haze simulations on a 36 km continental US modeling domain and additional shorter duration episodes on a 12 km domain covering the central states. After detailed performance testing, the modeling system will then be exercised with a variety of emissions control scenarios aimed at enabling CENRAP to assess the effects of future year emission control strategies on visibility and other air quality issues. The modeling system will also allow CENRAP to track reasonable progress toward regional haze goals. More specifically, the CENRAP 2002 annual modeling will focus on the use of the SMOKE emissions and CMAQ and CAMx air quality modeling systems for calendar year 2002 over a 36 km horizontal grid system. The CENRAP annual Emissions and Air Quality Modeling activities are being performed in two parts corresponding to fiscal years ending on April 30, 2005 (FY1) and April 30, 2006 (FY2). During FY1 the base year case modeling, model performance evaluation and future year base case modeling will be conducted. Whereas in FY2 future year sensitivity and control strategy modeling will be performed. A potentially large number of annual (and episodic) model simulations will be performed; the list below reflects the types of simulations that will be carried out:

- **2002 Initial Annual Run.** The initial annual model simulations and performance evaluations using a 2002 actual emissions inventory provided by CENRAP and other sources and 2002 MM5 meteorology provided by CENRAP. Multiple iterations of the 2002 annual simulation may be required to confirm the appropriateness of the model science configuration(s), to evaluate updates to the model and model inputs and to refine model performance.

- **2002 Actual Base Case Annual Run.** A subsequent annual 2002 simulation using actual 2002 emissions would be carried out using a final model configuration and inputs identified by the initial 2002 runs. The primary objective of this run is model performance demonstration using updated model inputs and best science model configurations. Additional sensitivity tests may be conducted using the actual base case year annual run.

- **2002 Annual Run with "Typical Year" EGU/Fire Inventory.** An annual 2002 simulation representing the 2000-2004 baseline period for EGU and fire emissions and using 2002 emissions inventory for all other source sectors. The primary objective of this inventory is to provide the baseline modeled air quality condition against which future year modeling runs will be compared to develop relative reduction factors for each pollutant species.

- **Future Year Annual Runs.** Future year simulations involving a base case inventory of typical EGU and fire emissions for CENRAP-selected future-year period, for which 2018 is the current thinking. Additional future year inventories may also be modeled. The objective of these future year model runs is to establish the modeled air quality basis against which the effectiveness of emissions control strategies will be evaluated.

- **Future Year Emission Control Strategies and Sensitivities.** Prescription of the future year emissions sensitivity and control strategies to be performed would occur during the second part (May 2005 to April 2006) of the CENRAP 2002 annual Emissions and Air Quality Modeling Study and would be better defined at a later date.

Closely integrated with the 2002 36 km annual continental US and episodic 12 km episodic central states emissions and air quality modeling will be ongoing project management, technical review, and quality assurance activities performed under the guidance of the CENRAP Project Manager, Administrative Project Manager and Modeling Workgroup. The Modeling Team members will participate with CENRAP management in regular conference calls, as well as ad hoc topical conference calls as needed, and will attend periodic meetings with the CENRAP Modeling Workgroup.

Complementing the data acquisition, modeling input development activities, and project management activities, four other CENRAP 2002 annual Emissions and Air Quality Modeling activities will be performed, consistent with the Quality Assurance Project Plan (QAPP; Morris and Tonnesen, 2004):

### 1.5.1 Data Gatekeepers

The CENRAP Emissions and Air Quality Modeling Team are receiving emissions, meteorological and air quality data from other CENRAP contractors and other sources (e.g., EPA, WRAP, IDNR, etc.) and as a first line of QA, we have defined a Gatekeeper function to assure the data have been received correctly, evaluate the quality of the data, and document the data received. Separate air quality, meteorological and emissions Gatekeepers have been identified whose roles are defined below. In addition, a Data Management Gatekeeper has been defined who will post data, reports and results to the project website and archive all key data generated in the project.

- **Air Quality Data Gatekeeper**. Obtain air quality data as appropriate for model input development and model performance evaluation and assure quality of all air quality data obtained, consistent with QAPP. This gatekeeper will also provide documentation of evaluation database for CMAQ and CAMx for all modeling runs.

- **Meteorological Gatekeeper**. Obtain meteorological data as MM5 output files for 36 km annual 2002 modeling runs and other episode periods at 12 km and perform data quality checks as QAPP together with appropriate documentation of model performance evaluation activities.

- **Emissions Gatekeeper.** Obtain emissions inventory data necessary to support annual 2002 base and future year modeling. Assure quality of all emissions data received, consistent with QAPP and develop all emissions modeling files to support modeling runs for 2002. This gatekeeper will also develop the chemical speciation files and temporal and spatial allocation files necessary to convert annual inventories into hourly and daily emissions modeling files, as appropriate. CENRAP will provide the Emissions Gatekeeper emissions for CENRAP states and the Emissions Gatekeeper will also develop all emissions modeling files for non-CENRAP states to support modeling runs for base-year and future-year base case and emissions strategies as defined by CENRAP.

- **Data Management Gatekeeper**: This gatekeeper will maintain the CENRAP Modeling Website including posting modeling input and output files, reports, interpretation of results, and other documents as requested by CENRAP to support all Phase II tasks. This includes, for example, the storage of model inputs and outputs for annual (and episodic) runs and the transfer (via fire wire or alternative media) of electronic files to CENRAP states, other regional planning organizations, EPA, other contractors, and stakeholders.

### 1.5.2 Emissions QA/QC

Emissions Quality Assurance (QA) and Quality Control (QC) are the single most critical steps in performing air quality modeling studies. Because emissions processing is tedious, time consuming and involves complex manipulation of many different types of large data sets, errors are frequently made in emissions processing and, if rigorous QA measures are not in place, these errors may remain undetected. In the CENRAP 2002 annual Emissions and Air Quality Modeling Study we will expand on the multi-step emissions the QA/QC approach the Modeling Team has developed for WRAP and VISTAS emissions and air quality modeling. This includes the initial emissions QA/QC by the Emissions Gatekeeper described above, as well as QA/AC by the emissions modeler during the processing of emissions and then additional QA/QC by the air quality modeler of the processed model ready emission files. This multi-step process with three separate groups involved in the QA/QC of the emissions is much more likely to catch any errors prior to the air quality model simulations.

Emissions QA/QC that would be performed as part of the emissions modeling will include:

Input Screening Error Checking Algorithms: Although the SMOKE emissions model will be used for emissions processing, additional input error checking algorithms will be used to screen the data and identify potential emission input errors. Additionally, EPA has issued a revised stack QA and augmentation procedures memorandum that will be used to identify and augment any outlying stacks.

SMOKE error messages: SMOKE provides various cautionary or warning messages during the emissions processing. We will redirect the SMOKE output to log files and review the log files for serious error messages. An archive of the log files will be maintained so that the error messages can be reviewed at a later date if necessary.

SMOKE emissions summaries: We will use QA functions built into the SMOKE processing system to provide summaries of processed emissions as daily totals according to species, source category and county and state boundaries. These summaries will then be compared with summary data prepared for the pre-processed emissions, e.g., state and county totals for emissions from the augmented emissions data.

Once the CMAQ-ready emission inputs have been prepared, we will perform additional emissions QA/QC as follows:

Spatial Summary: We will sum the emissions for all layers and for all 24 hours that is used to prepare a PAVE plot showing the daily total emissions spatial distribution. For a 20 day simulation this produces approximately 20 days x 20 species x 5 emissions categories = 2,000 plots. In our base case simulations these plots will be presented as tons per day. The objective of this step is to identify errors in spatial distribution of emissions.

Vertical Profile: For point sources the emissions total for each layer will be summed and plotted to show the vertical distribution of emissions. These plots show the emissions on the x-axis for each model layer on the y-axis. The objective of this step is to identify possible errors in vertical distribution of emissions.

Short Term Temporal Summary: The total domain emissions for each hour will be accumulated and time series plots prepared that display the diurnal variation in total hourly emissions. The objective of this step is to identify errors in temporal profiles.

Long Term Temporal Summary: The total domain emissions for each day will be accumulated and displayed as time series plots that show the daily total emissions across the domain as a function of time. The objective of this step is to identify particular days for which emissions appear to be inconsistent with other days for no reason (e.g., not a weekend) and compare against the general trend.

Control Strategy Spatial Displays: Spatial summary plots of the daily total emissions differences between a control strategy and base case emissions scenarios will be generated. These plots can be used to immediately identify a problem in a control strategy. For example, if a CENRAP state's $SO_2$ control strategy is being analyzed and there are changes in emissions for other pollutants or for $SO_2$ outside of the CENRAP state occurs, problems in emissions processing can be identified prior to the air quality model simulation.

Once the CMAQ-ready emissions have been subjected to the above QA, the CMAQ-to-CAMx emissions processor would be exercised to generate CAMx-ready emission inputs. The CAMx-ready emission inputs would then be subjected to an additional round of QA to assure that identical emissions are contained within the CMAQ-ready and CAMx-ready emission inputs.

### 1.5.3 Meteorology QA/QC

The CENRAP meteorological modeler will have primary responsibility in the QA/QC of the MM5 meteorological fields (Johnson, 2004). However, the CENRAP Emissions and Air Quality Modeling Team will also perform some QA/QC of the meteorological data to assure that it has transferred correctly, to obtain an assessment of the quality of the data and to assist in the interpretation of the air quality modeling results.

The CENRAP Meteorological Gatekeeper will perform the following:

- Analyze the MM5 data to assure it has been transferred correctly;
- Evaluate the MM5 using METSTAT and the surface meteorological network;
- Evaluate upper-air MM5 meteorological estimates by comparison them to upper-air observations and satellite images;
- Compare the CENRAP 2002 MM5 simulation performed with the performance of the WRAP and VISTAS MM5 modeling; and
- Generate the CMAQ-ready and CAMx-ready meteorological inputs using the MCIP2.3 and MM5CAMx processors, respectively.

### 1.5.4 Air Quality Modeling QA/QC

Key aspects of QA for the CMAQ and CAMx input and output data include the following:

- Verification that correct configuration and science options are used in compiling and running each model of the in the CMAQ and CAMx modeling system.
- Verification that correct input data sets are used when running each model.
- Evaluation of CMAQ and CAMx results to verify that model output is reasonable and consistent with general expectations.
- Processing of ambient monitoring data for use in the model performance evaluation.
- Evaluation of the CMAQ and CAMx results against concurrent observations.
- Backup and archiving of critical model input data.

The most critical element in the QA for CMAQ and CAMx simulations is the QA/QC of the meteorological and emissions input files, which is discussed above. The major QA issue specifically associated with the air quality model simulations is verification that the correct science options were specified in the model itself and that the correct input files were used when running the model. For the CMAQ and CAMx modeling we employ a system of naming conventions using environment variables in the compile and run scripts that guarantee that correct inputs and science options are used. We also employ a redundant naming system so that the name of key science options or inputs are included in the name of the CMAQ and CAMx executable program, in the name of the CMAQ and CAMx output files, and in the name of the directory in which the files are located. This is accomplished by using the environment variables in the scripts to specify the names and locations of key input files.

A second key QA procedure is to never "recycle" run scripts, i.e., we always preserve the original runs scripts and directory structure that were used in performing a model simulation.

We will also perform a post-processing QA of the CMAQ and CAMx output files similar to that described for the emissions processing. We will generate animated gif files using PAVE that can be viewed to search for unexpected patterns in the CMAQ and CAMx output files. In the case of model sensitivity studies, the animated gifs will be prepared as difference plots for the sensitivity case minus the base case. Often, errors in the emissions inputs can be discovered by viewing the animated GIFs. Finally, we will produce 24-hour average plots for each day of the CMAQ and CAMx simulations. This provides a summary that can be useful for more quickly comparing various model simulations.

### 1.5.5 Overview of Data Flow and Quality Assurance Process

Figure 1-2 displays an overview of the data flow and quality assurance process in the CENRAP Emissions and Air Quality Modeling study. The CENRAP Modeling Team receives different types of data from various CENRAP participants and contractors and other sources that have performed their own Quality Assurance (QA) and Quality Control (QC). Whenever data are received by the Modeling Team, it is first subjected to a QA check by a Gatekeeper who assess the accuracy and quality of the data and prepares a summary presentation on the QA check. Figure 1-2a lists the Gatekeepers in the Modeling Team for emissions, boundary conditions, meteorological, ozone column (TOMS) and air quality data. If the Gatekeeper identifies any problems with the data, the provider of the data is contacted and asked to correct the data. Once the Gatekeeper has conducted a QA check of the data it is passed on to the modeler who performs their own QA of the data. The data are then used in the modeling and resultant output (e.g., model-ready emissions or meteorological files) are then subjected to another round of QA to assure the integrity of the data is retained.

Once the model-ready inputs have been developed and subjected to QA/QC, the models (e.g., CMAQ and/or CAMx) are applied using Base Case emissions and the modeling results subjected to a model performance evaluation. The model performance evaluation (MPE) represents an extensive QA effort and is one of the most time consuming component of the study. EPA has developed draft guidance for evaluating regional PM and haze models that includes performance goals (EPA, 2001). In addition, the Modeling team has adapted EPA MPE approaches and goals for 1-hour (EPA, 1991) and 8-hour (EPA, 1999) ozone modeling. The CENRAP Modeling Team performs the MPE/QA process using as many different tools and analysis as possible in order to fully understand the accuracy and reliability of the model simulation. As seen in Figure 1-2b, the MPE process in CENRAP is a multi-step process using several different techniques:

UCR Analysis Tools: The University of California at Riverside (UCR) Analysis Tools are run on a Linux platform separately for each monitoring network and for different subregions. Graphics are automatically generated using gnuplot and the software generates the following:

- Tabular statistical measures;
- Time Series Plots; and
- Scatter Plots by allsite_allday, allday_onesite and allsite_oneday.

ENVIRON Analysis Tools: ENVIRON has developed specialized evaluation tools to analyze visibility model performance for the Best and Worst 20% visibility days that are used in visibility projections for the Section 308 SIPs/TIPs. ENVIRON has also developed "Soccer Plots" that displays model performance across networks, episodes, species, models and sensitivity tests and compare them with performance goals. As part of VISTAS, the Georgia DNR has developed "Bugle Plots" that display model performance as a function of observed concentration that have been integrated with ENVIRON's evaluation tools.

CENRAP Model Evaluation Tool: CENRAP has developed a model evaluation tool that includes the observations in a MySQL database for ease of manipulation of the observation database. The CENRAP model evaluation tool can interface with Excel and/or gnuplot to generate the usual set of scatter plots, time series plots, etc.

The evaluation of the CENRAP 2002 CMAQ Base Case simulation will use each of the analysis tools listed above to take advantage of their different descriptive and complimentary nature. The use of multiple model evaluation tools is also a useful QA/QC procedure to assure that errors are not introduced in the model evaluation process.

### 1.5.6 Proposed Model Performance Goals

The issue of model performance goals for PM species is an area of ongoing research and debate. For ozone modeling, EPA has established performance goals for 1-hour ozone normalized bias and gross error of #±15% and #35%, respectively (EPA, 1991). EPA's draft fine particulate modeling guidance notes that performance goals for ozone should be viewed as upper bounds of model performance, that PM models may not be able to always achieve and we should demand better model performance for PM components that make up a larger fraction of the PM mass than those that are minor contributors (EPA, 2001). Measuring PM species is not as precise as ozone monitoring. In fact, the differences in measurement techniques for some species likely exceed the more stringent performance goals, such as those for ozone. For example, recent comparisons of the PM species measurements using the IMPROVE and STN measurement technologies found differences of approximately "20% ($SO_4$) to "50% (EC) (Solomon et al., 2004).

In the CENRAP 2002 CMAQ Base Case modeling, we have adopted three levels of model performance goals for bias and gross error as listed in Table 1-2 that are used to help evaluate model performance. Note that we are not suggesting that these performance goals be generally adopted or that they are the most appropriate goals to use. Rather, we are just using them to frame and put the PM model performance into context and to facilitate model performance intercomparison across episodes, species, models and sensitivity tests.

As noted in EPA's draft PM modeling guidance, less abundant PM species should have less stringent performance goals. Accordingly, we are also using performance goals that are a continuous function of average observed concentrations proposed by Dr. James Boylan at the Georgia Department of Natural Resources that have the following features:

- Asymptotically approaching proposed performance goals or criteria when the mean of the observed concentrations are greater than 2.5 ug/m$^3$.
- Approaching 200% error and "200% bias when the mean of the observed concentrations are extremely small.

Dr. Boylan uses bias/error goals and criteria of ±30%/50% and ±60%/75% and plots bias and error as a function of average observed concentrations. As the mean observed concentration approaches zero the bias performance goal and criteria flare out to ±200% creating a horn shape, hence the name "Bugle Plots".

**Table 1-2.** Model performance goals used in Phase I to help interpret modeling results.

| Fractional Bias | Fractional Error | Comment |
|---|---|---|
| #"15% | #35% | Ozone model performance goal for which PM model performance would be considered good. |
| #"30% | #50% | A level of model performance that we would hope each PM species could meet |
| #"60% | #75% | At or above this level of performance indicates fundamental problems with the modeling system. |

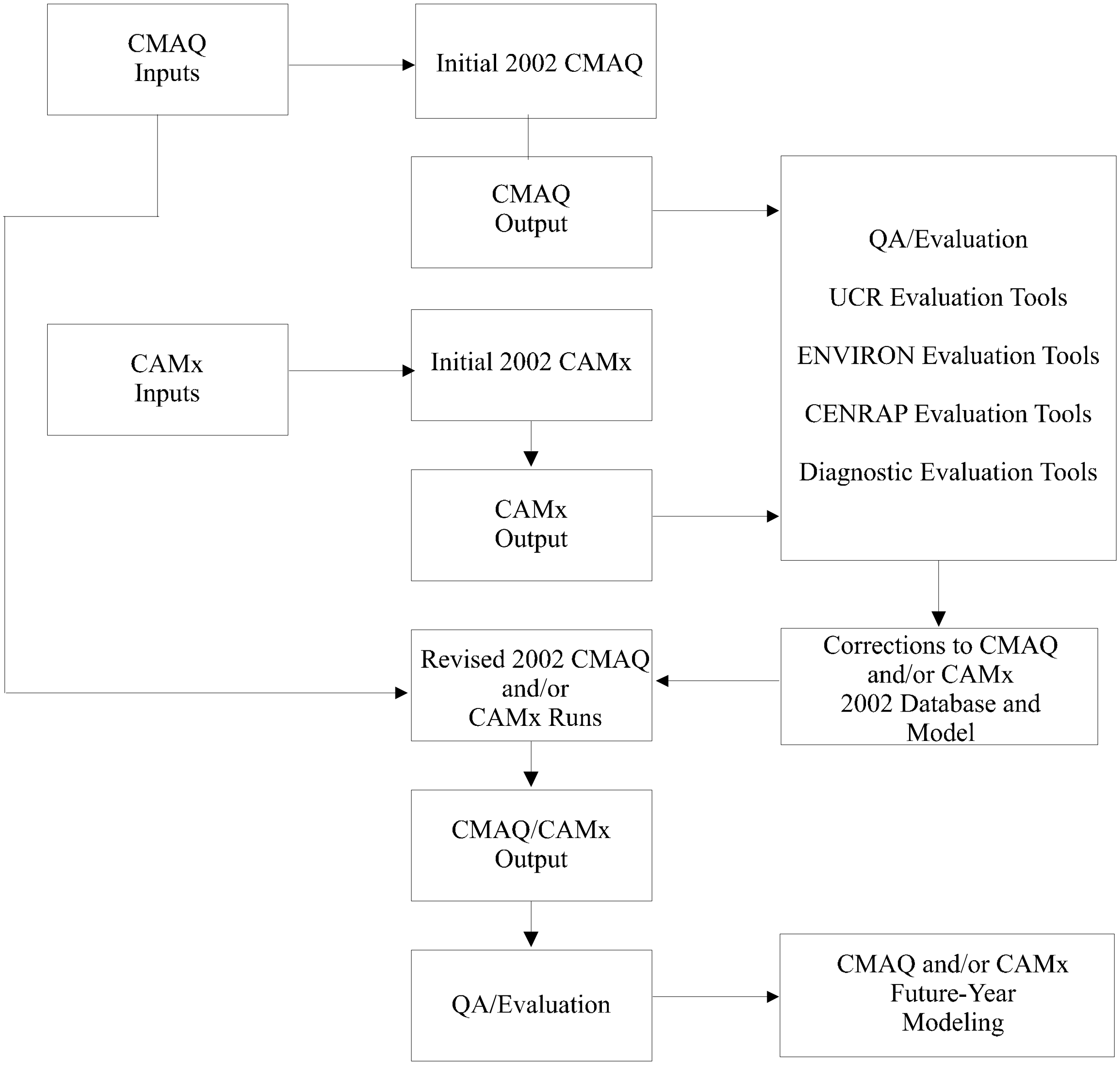


**Figure 1-2a.** Data flow and quality assurance steps in the CENRAP Emissions and Air Quality Modeling.

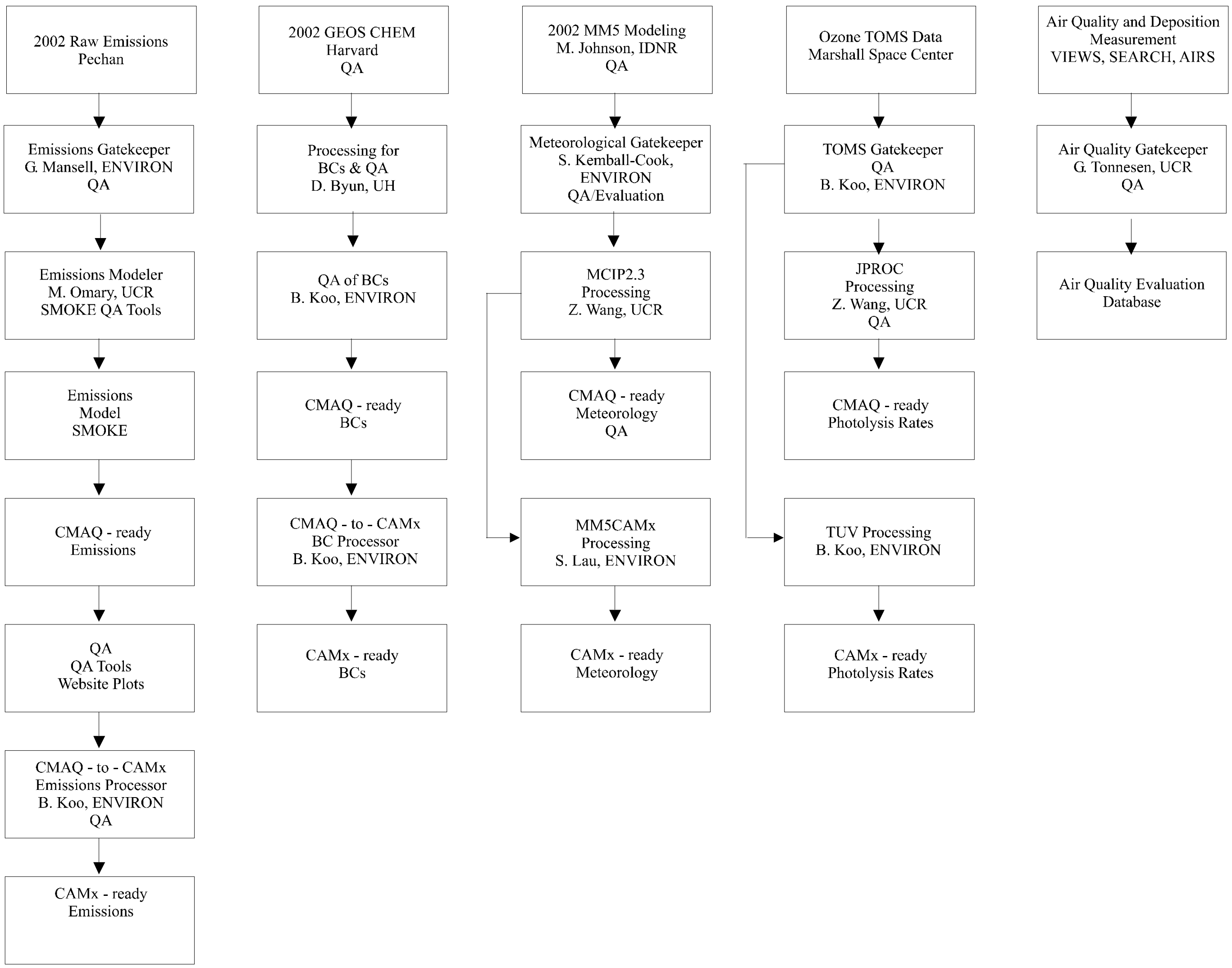


**Figure 1-2b.** Concluded. Data flow and quality assurance steps in the CENRAP Emissions and Air Quality Modeling.

## 1.6 CENRAP Annual Modeling Study Participants and Responsibilities

The CENRAP 2002 annual Emissions and Air Quality Modeling Study is being carried out by ENVIRON and UCR under contract to CENRAP. Table 1-3 lists the contact information of the key participants. Their roles in the work are as follows:

- Annette Sharp, CENRAP Technical Director is the Administrative Project Office for the study;
- Lee Warden of the Oklahoma Department of Environmental Quality is the CENRAP Project Manager;
- Calvin Ku of the Missouri Department of Natural Resources is the CENRAP Alternate Project Manager;
- Kathy Pendleton of the Texas Commission on Environmental Quality is the CENRAP contact for emissions;
- Matthew Johnson of the Iowa Department of Natural Resources is the CENRAP contact for 2002 36 km MM5 meteorology;
- Bret Anderson of EPA Region VII is the CENRAP contact for 12 km MM5 meteorology;
- T.W. Tesche of Alpine Geophysics, LLC is the technical advisor to CENRAP on modeling activities;
- Ralph Morris of ENVIRON is the Project Manager and Co-Principal Investigator for the CENRAP Emissions and Air Quality Modeling ENVIRON/UCR Team;
- Gail Tonnesen of UCR is the Co-Principal Investigator for the CENRAP Emissions and Air Quality Modeling ENVIRON/UCR Team;
- Gerard Mansell of ENVIRON is the Emissions Coordinator for the Emissions and Air Quality Modeling ENVIRON/UCR Team; and
- Mohammed Omary of UCR is the Chief Emissions Modeler for the Emissions and Air Quality Modeling ENVIRON/UCR Team.

**Table 1-3.** Summary of key participants and contact information for the CENRAP 2002 Annual Emissions and Air Quality Modeling Study.

| Person & Role | Affiliation/Address | Contact Information |
|---|---|---|
| Annette Sharp<br>(Administrative Project Manager) | CENRAP<br>10005 S. Pennsylvania, Ste. C<br>Oklahoma City, OK 73159 | (405) 378-7377<br>asharp@censara.org |
| Lee Warden<br>(Technical Project Manager) | Oklahoma DEQ<br>707 N. Robinson<br>Oklahoma City, OK 73102 | (405) 702-4201<br>Lee.Warden@deq.state.ok.us |
| Calvin Ku<br>(Alternate Technical Project Manager) | Missouri DNR<br>P.O. Box 176<br>Jefferson City, MO 65012 | (573) 751-4817<br>calvin.ku@dnr.mo.gov |
| Kathy Pendleton<br>(Emissions Contact) | Texas CEQ<br>12100 Park 35 Circle, MC-164<br>Austin, TX 78753 | (512) 239-1936<br>kpendlet@tceq.state.tx.us |
| Matthew Johnson<br>(36 km MM5 Meteorology Contact) | Iowa DNR<br>7900 Hickman Rd. Ste. 1<br>Urbandale, IA 50322 | (515) 242-5164<br>matthew.johnson@dnr.state.ia.us |
| Bret Anderson<br>(12 km MM5 Meteorology Contact) | EPA Region VII<br>901 North Fifth St.<br>Kansas City, KS 66101 | (913) 551-7862<br>Anderson.bret@epa.gov |
| T.W. Tesche<br>(Advisor on Modeling) | Alpine Geophysics, LLC<br>3479 Reeves Drive<br>Ft. Wright, KY 41017 | (859) 341-7502<br>(Fax) (859) 341-7508<br>twt@ac.net |
| Ralph Morris<br>(Project Manager and Co-Principal Investigator) | ENVIRON<br>101 Rowland Way<br>Novato, CA 94945 | (415) 899-0708<br>(Fax) (415) 899-0707<br>rmorris@environcorp.com |
| Gail Tonnesen<br>(Co-Principal Investigator) | UC Riverside<br>CE-CERT<br>1084 Columbia Avenue<br>Riverside, CA 92507 | (951) 781-5676<br>(Fax) (951) 781-5790<br>tonnesen@cert.ucr.edu |
| **Key ENVIRON Participants** | | |
| Gerry Mansell<br>(Emissions Modeling Coordinator) | ENVIRON<br>101 Rowland Way<br>Novato, CA 94945 | (415) 899-0727<br>gmansell@environcorp.com |
| Bonyoung Koo | ENVIRON | (415) 899-0727<br>bkoo@environcorp.com |
| Edward Tai | ENVIRON | (415) 899-0725<br>etai@environcorp.com |
| Steven Lau | ENVIRON | (415) 899-0739<br>slau@environcorp.com |
| **Key UCR CE-CERT Participants** | | |
| Mohammed Omary<br>(Chief Emissions Modeler) | UC Riverside<br>CE-CERT<br>1084 Columbia Avenue<br>Riverside, CA 92507 | (951) 781-5652<br>omary@cert.ucr.edu |
| Zion Wang | UC Riverside | (951) 781-5655<br>zsw@cert.ucr.edu |
| Chao-Jung Chien | UC Riverside | (951) 781-5666<br>chien@cert.ucr.edu |
| Glen Kaukola | UC Riverside | (951) 781-5630<br>glen@cert.ucr.edu |

## 2.0 MODEL SELECTION

This chapter introduces the regional meteorological, emissions and air quality models to be used in the annual PM/regional haze modeling for CENRAP. The specific science configurations for each modeling system are identified and discussed briefly, where necessary. Although the initial configurations of each modeling system have been selected as the culmination of a review of previous regional haze modeling studies performed in the CENRAP region (e.g., Pitchford et al., 2004; Pun, Chen and Seigneur, 2004; Tonnesen and Morris 2004) as well as elsewhere in the United States (e.g., Morris et al, 2004a; Tonnesen et al., 2003; Baker, 2004), there remains the possibility that certain algorithms and parameter settings may still be updated in the establishment of the final annual 2002 base case simulation and model performance testing. The CENRAP Emissions and Air Quality Modeling Team will remain alert to progressive model code improvements, data base refinements, and emergent analysis procedures throughout the entire activity. Notable limitations of the models relevant to their intended purpose in CENRAP are identified. We conclude with a general overview of the input requirements for each system with more details provided in Section 5.

### 2.1 Recommended Models

Based on the previous CENRAP, WRAP, VISTAS, MRPO, BRAVO, EPA and other work, CENRAP selected the following models for use in modeling particulate matter (PM) and regional haze in the central states:

- **MM5:** The Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) Mesoscale Meteorological Model (MM5) is a nonhydrostatic, prognostic meteorological model routinely used for urban- and regional-scale photochemical, fine particulate, and regional haze regulatory modeling studies.

- **SMOKE**: The Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system is an emissions modeling system that generates hourly gridded speciated emission inputs of mobile, nonroad, area, point, fire and biogenic emission sources for photochemical grid models.

- **CMAQ:** EPA's Models-3/Community Multiscale Air Quality (CMAQ) modeling system is a 'One-Atmosphere' photochemical grid model capable of addressing ozone, particulate matter (PM), visibility and acid deposition at regional scale for periods up to one year.

- **CAMx:** ENVIRON's Comprehensive Air Quality Model with Extensions (CAMx) modeling system is also a state-of-science 'One-Atmosphere' photochemical grid model capable of addressing ozone, particulate matter (PM), visibility and acid deposition at regional scale for periods up to one year.

Application of the MM5 for the 2002 annual modeling on a 36 km grid for the continental US is being performed by the Iowa Department of Natural Resources (IDNR; Johnson, 2004). Details of the 2002 36 km MM5 model application and evaluation procedures being carried out by

IDNR may be found in Johnson (2004). Application of the MM5 model on a 12 km grid covering the Central States for portions of 2002 is being performed by EPA Region VII. For completeness, in this chapter we describe all four regional modeling systems and their intended use in the CENRAP 2002 annual modeling.

## 2.2 MM5 Mesoscale Prognostic Model

Over the past decade, researchers at the Pennsylvania State University (PSU) and the National Center for Atmospheric Research (PSU/NCAR) have collaborated in the refinement and extension of the PSU Mesoscale Meteorological Model leading to the current version of the system, MM5 (ver 3.6, MPP). Originally developed in the 1970s at PSU and first documented by Anthes and Warner (1978), the MM5 modeling system maintains its status as a state-of-the-science model through enhancements provided by a broad user community (e.g., Chen and Dudhia, 2001; Stauffer and Seaman, 1990, 1991; Xiu and Pleim, 2000). The MM5 modeling system is routinely employed in forecasting projects as well as refined investigations of severe weather. Utilization of MM5 within air quality applications is also a common practice. In recent years, the MM5 modeling system has been successfully applied in continental scale annual simulations for the years 1996 (Olerud et al., 2000), 2001 (McNally and Tesche, 2003), and 2002 (Johnson, 2004). Due to its ongoing scientific development worldwide, extensive historical applications, broad user community support, public availability, and established performance record compared with other applications-oriented prognostic models, CENRAP selected the MM5 as the preferred meteorological model. This section provides an overview of the MM5 and its data input requirements.

### 2.2.1 MM5 Overview

The non-hydrostatic MM5 model (Dudhia, 1993; Grell et al., 1994) is a three-dimensional, limited-area, primitive equation, prognostic model that has been used widely in regional air quality model applications (Seaman, 2000). The basic model has been under continuous development, improvement, testing and open peer-review for more than 20 years (Anthes and Warner, 1978; Anthes et al., 1987) and has been used world-wide by hundreds of scientists for a variety of mesoscale studies, including cyclogenesis, polar lows, cold-air damming, coastal fronts, severe thunderstorms, tropical storms, subtropical easterly jets, mesoscale convective complexes, desert mixed layers, urban-scale modeling, air quality studies, frontal weather, lake-effect snows, sea-breezes, orographically induced flows, and operational mesoscale forecasting.

MM5 is based on the prognostic equations for three-dimensional wind components (u, v, and w), temperature (T), water vapor mixing ratio ($q_v$), and the perturbation pressure (p'). Use of a constant reference-state pressure increases the accuracy of the calculations in the vicinity of steep terrain. The model uses an efficient semi-implicit temporal integration scheme and has a nested-grid capability that can use up to ten different domains of arbitrary horizontal and vertical resolution. The interfaces of the nested grids can be either one-way or two-way interactive. The model is also capable of using a hydrostatic option, if desired, for coarse-grid applications.

MM5 uses a terrain-following non-dimensionalized pressure, or "sigma", vertical coordinate similar to that used in many operational and research models. In the non-hydrostatic MM5 (Dudhia, 1993), the sigma levels are defined according to the initial hydrostatically-balanced reference state so that the sigma levels are also time-invariant. The gridded meteorological fields produced by MM5 are directly compatible with the input requirements of 'one atmosphere' air-quality models (e.g., CMAQ and CAMx).

Distinct planetary boundary layer (PBL) parameterizations are available for air-quality applications, both of which represent sub-grid-scale turbulent fluxes of heat, moisture and momentum. These parameterizations employ various surface energy budget equations to estimate ground temperature ($T_g$), based on the insolation, atmospheric path length, water vapor, cloud cover and longwave radiation. The surface physical properties of albedo, roughness length, moisture availability, emissivity and thermal inertia are defined as functions of land-use for numerous categories via a look-up table. One scheme uses a first-order eddy diffusivity formulation for stable and neutral environments and a modified first-order scheme for unstable regimes. The other uses a prognostic equation for the second-order turbulent kinetic energy, while diagnosing the other key boundary layer terms.

Initial and lateral boundary conditions are specified from mesoscale three-dimensional analyses performed at 12-hour intervals on the outermost grid mesh selected by the user. Additional surface fields are analyzed at three-hour intervals. A Cressman-based technique is used to analyze standard surface and radiosonde observations, using the National Meteorological Center's (NMC) spectral analysis as a first guess. The lateral boundary data are introduced into MM5 using a relaxation technique applied in the outermost five rows and columns of the most coarse grid domain.

A major feature of the MM5 is its use of state-of-science methods for Four Dimensional Data Assimilation (FDDA). The theory underlying this approach and details on how it has been applied in a variety of applications throughout the country are described in depth elsewhere (Stauffer and Seaman, 1990, 1991; Seaman et al., 1992, 1997).

Results of detailed performance evaluations of the MM5 modeling system in regulatory air quality application studies have been widely reported in the literature (e.g., Emery et al., 1999; Tesche et al., 2000, 2003) and many have involved comparisons with other prognostic models such as RAMS and SAIMM. The MM5 enjoys a far richer application history in regulatory modeling studies compared with RAMS or other models. Furthermore, in evaluations of these models in over 60 recent regional scale air quality application studies since 1995, we have generally found that MM5 model tends to produce somewhat better photochemical model inputs than alternative models. For these and other reasons, MM5 was selected as the meteorological modeling system for the CENRAP study.

### 2.2.2 MM5 Configuration for CENRAP Annual Modeling

Based on the sensitivity testing carried out by IDNR, LADCO and others, the MM5 (ver 3.63) configuration used by the IDNR modelers in the CENRAP 2002 annual 36 km MM5 modeling consist of the following (see Table 2-1 for more details):

- Nested 36 km grid with 34 vertical layers;
- 12 km grid for episodic modeling;
- For 12 km runs use two way nesting with no feedback;
- Initialization and boundary conditions from Eta analysis fields;
  - Eta 3D and surface analysis data (ds609.2);
  - Not using NCEP global tropospheric SST data (ds083.0) ;
  - Observational enhancement (LITTLE_R)
    - NCEP ADP surface obs (ds464.0)
    - NCEP ADP upper-air obs (ds353.4)
- Pleim-Xiu (P-X) land soil model (LSM);
- Pleim-Chang Asymmetric Convective Mixing (ACM) PBL model;
- Kain-Fritsch 2 cumulus parameterization;
- Mixed phase (Reisner 1) cloud microphysics;
- Raptid Radiative Transfer Model (RRTM) radiation;
- No Shallow Convection (ISHALLO=0);
- Standard 3D FDDA analysis nudging; and
- No surface nudging.

## 2.3 SMOKE Emissions Modeling System

### 2.3.1 SMOKE Overview

The Sparse Matrix Operator Kernel Emissions (SMOKE) Emissions Processing System Prototype was originally developed at MCNC (Coats, 1995; Houyoux and Vukovich, 1999). As with most ‘emissions models’, SMOKE is principally an *emission processing system* and not a true *emissions modeling system* in which emissions estimates are simulated from ‘first principles’. This means that, with the exception of mobile and biogenic sources, its purpose is to provide an efficient, modern tool for converting emissions inventory data into the formatted emission files required by an air quality simulation model. For mobile sources, SMOKE actually simulates emissions rates based on input mobile-source activity data, emission factors and sometimes output from transportation travel-demand models.

SMOKE was originally designed to allow emissions data processing methods to utilize emergent high-performance-computing (HPC) as applied to sparse-matrix algorithms. Indeed, SMOKE is the fastest emissions processing tool currently available to the air quality modeling community. The sparse matrix approach utilized throughout SMOKE permits both rapid and flexible processing of emissions data. The processing is rapid because SMOKE utilizes a series of matrix calculations instead of less efficient algorithms used in previous systems. The processing is flexible because the processing steps of temporal projection, controls, chemical speciation, temporal allocation, and spatial allocation have been separated into independent operations wherever possible. The results from these steps are merged together at a final stage of processing.

SMOKE supports area, mobile, fire and point source emission processing and also includes biogenic emissions modeling through a rewrite of the Biogenic Emission Inventory System, version 3 (BEIS3) (see, http://www.epa.gov/ttn/chief/software.html#pcbeis). SMOKE has been available since 1996, and it has been used for emissions processing in a number of regional air quality modeling applications. In 1998 and 1999, SMOKE was redesigned and improved with the support of the U.S. Environmental Protection Agency (EPA), for use with EPA's Models-3/CMAQ (http://www.epa.gov/asmdnerl/models3). The primary purposes of the SMOKE redesign were support of: (a) emissions processing with user-selected chemical mechanisms and (b) emissions processing for reactivity assessments.

SMOKE contains a number of major features that make it an attractive component of the CENRAP modeling system (Seppanen, 2003). The model supports a variety of input formats from other emissions processing systems and models including the Inventory Data Analyzer (IDA), Emissions Modeling System—2003 (EMS-2003), and the Emissions Preprocessor System 2.x (EPS2.x). It supports both gridded and county total land use scheme for biogenic emissions modeling. Although not necessary in CENRAP, SMOKE can accommodate emissions files from up to 10 countries and any pollutant can be processed by the system.

Recent *computational improvements* to SMOKE include: (a) enhanced disk space requirements compared with other emissions processing software, (b) run-time memory allocation, eliminating any need to recompile the programs for different inventories, grids, or chemical mechanisms, and (c) updated I/O API libraries. A number of *science features* have been incorporated into the "current" version of SMOKE (ver. 2.0) including: (a) any chemical mechanism can be used to partition pollutants to model species, as long as the appropriate input data are supplied, (b) integration with the MOBILE6.2 on-road mobile source emissions model including link based processing, (c) support of plume-in-grid (PiG) processing, (d) integration of the BEIS3 emissions factors in SMOKE.

A new version of SMOKE (ver.2.1) has just been (October 2004) released (www.cmascenter.org). As with any new release of a model, care must be taken before unilaterally switching from a working version of the model that is already set up to the new just released version. This is particularly important for emissions models that perform extensive data processing and manipulation in multiple steps. If SMOKE ver.2.1 can be easily substituted for SMOKE ver2.0 (e.g., plug and play) then it may be considered for the CENRAP annual modeling.

Notable features of SMOKE from an *applications* standpoint include: (a) improved control strategy input formats and designs, (b) control strategies can include changes in the reactivity of emitted pollutants, a useful capability, for example, when a solvent is changed in an industrial process, (c) no third party software is required to run SMOKE, although some input file preparation may require other software, (d) fewer SMOKE programs than the SMOKE prototype because programs were combined where possible to be used for multiple source categories, (e) integration with Models-3 file formats and settings, (f) improved data file formats, (g) support of various air quality model emissions input formats (e.g., CMAQ, MAQSIP, UAM-

IV, UAM-V, REMSAD and CAMx), (h) enhanced quality assurance pre- and post-processing, (h) fully integrated with Models-3, which will provide the SMOKE Tool for SMOKE input file preparation, (i) enhanced treatment of growth and control factors, (j) improved emissions reporting and QA capabilities, and (k) improved temporal allocation.

Continuing model development activities with SMOKE now occur out of the University of North Carolina (UNC) Carolina Environmental Program (CEP). SMOKE ver2.0 was released on 30 Sept '03, which is the version we are currently planning to employ for the CENRAP annual modeling. Note that SMOKE ver 2.1 has just been released and if set up in time will be considered for use in the CENRAP annual modeling. However, SMOKE ver 2.1 needs to be set up and evaluated prior to any commitment to its use. The SMOKE executables, scripts and databases may be downloaded through the Community Modeling and Analysis (CMAS) center's Model Clearinghouse at http://www.cmascenter.org/modelclear.shtml. The SMOKE user's guide is available online at the main SMOKE website, http://www.cep.unc.edu/empd/products/smoke.

### 2.3.2 SMOKE Configuration for CENRAP Annual Modeling

As an emissions processing system, SMOKE has far fewer 'science configuration' options compared with the MM5 and CMAQ models. For a thorough characterization of the methods that will be used to exercise the SMOKE system for the annual 2002 emissions processing, see section 5.2, "Development of Emissions Model Inputs and Resultant Inventories". Table 2-1 summarizes the version of the SMOKE system to be used and the sources of data to be employed in constructing the required modeling inventories.

## 2.4 CMAQ Modeling System

### 2.4.1 CMAQ Overview

For more than a decade, EPA has been developing the Models-3 Community Multiscale Air Quality (CMAQ) modeling system with the overarching aim of producing a 'One-Atmosphere' air quality modeling system capable of addressing ozone, particulate matter (PM), visibility and acid deposition within a common platform (Dennis, et al., 1996; Byun et al., 1998a; Byun and Ching, 1999, Pleim et al., 2003). The original justification for the Models-3 development emerged from the challenges posed by the 1990 Clean Air Act Amendments and EPA's desire to develop an advanced modeling framework for 'holistic' environmental modeling utilizing state-of-science representations of atmospheric processes in a high performance computing environment (Ching, et al., 1998). EPA completed the initial stage of development with Models-3 and released the Community Multi-Scale Air Quality model (CMAQ) in mid-1999 as the initial operating science model under the Models-3 framework (Byun et al., 1998b). The most recent rendition is CMAQ version 4.4, publicly released October 2004 and is the version to be used in the CENRAP annual modeling.

CMAQ consists of a core Chemical Transport Model (CTM) and several pre-processors including the Meteorological-Chemistry Interface Processor (MCIP), initial and boundary conditions processors (ICON and BCON) and a photolysis rates processor (JPROC). EPA is continuing to improve and develop new modules for the CMAQ model and typically provides a new release each year. In the past EPA has also provides patches for CMAQ as errors are discovered and corrected. More recently EPA has funded the Community Modeling and Analysis Systems (CMAS) center to support the coordination, update and distribution of the Models-3 system (www.cmascenter.org).

A number of features in CMAQ's theoretical formulation and technical implementation make the model well-suited for annual PM modeling. In CMAQ, the modal approach has been adapted to dynamically represent the PM size distribution using three log-normal modes (2 fine and 1 coarse). Transfer of mass between the aerosol and gas phases is assumed to be in equilibrium and all secondary aerosol (sulfate, nitrate, SOA) is assumed to be in the fine modes. The thermodynamics of inorganic aerosol composition are treated using the ISORROPIA module. Aerosol composition is coupled to mass transfer between the aerosol and gas phases. For aqueous phase chemistry, the RADM model is currently employed. This scheme includes oxidation of $SO_2$ to sulfate by ozone, hydrogen peroxide, oxygen catalyzed by metals and radicals. The impact of clouds on the PM size distribution is treated empirically. For wet deposition processes, CMAQ uses the RADM/RPM approach. Particle dry deposition is included as well. CMAQ contains three options for treating secondary organic aerosol (SOA), latest being the Secondary Organic Aerosol Model (SORGAM) that was updated in August 2003 to be an reversible semi-volatile scheme whereby VOCs can be converted to condensable gases that can then form SOA and then evaporate back into condensable gases depending on atmospheric conditions.

The newest features implemented in the latest CMAQ (ver 4.4 released October 2004) are described in the release notes available on the CMAS Center website (www.cmascenter.org). Table 2-3 highlights the major options in CMAQ (ver 4.4) for different processes and compares them with the recently released CAMx (ver 4.10s) model in Table 2-4, which is discussed later in this chapter.

### 2.4.2 CMAQ Configuration for CENRAP Annual Modeling

In this section we identify the main science options we recommend for the CENRAP annual PM modeling with CMAQ. In particular, we propose to run CMAQ (ver 4.4) with the base configuration as shown in Table 2-3. The model would be set up and exercised on the 36 km grid continental US Inter-RPO modeling domain that is also used by WRAP and VISTAS. For the 12 km episodic modeling, CMAQ will be set up on a 12 km domain cover the central states whose definition is to be determine using one-way nesting. That is, boundary conditions for the 12 km grid simulation are extracted from the 36 km run using the CMAQ BCON processor. A total of 19 vertical layers would be implemented, extending up to a region top of 100 mb (approximately 15 km AGL).

The PPM advection solver would be used along with the spatially varying (Smagorinsky) horizontal diffusion approach and K-theory for vertical diffusion. MM5 meteorological output based on the Pleim-Xiu Land-Surface Model (LSM) and the ACM planetary boundary layer (PBL) scheme will be used (see Table 2-1) and the recently (October 2004) updated CMAQ

Meteorological-Chemistry Interface Processor (MCIP2.3) would process the MM5 data using the "pass through" option. The CB4 gas-phase, RADM aqueous-phase, and AERO3/ISORROPIA aerosol chemistry schemes are recommended for use in the initial CMAQ 2002 modeling. Treatment of reversible secondary organic aerosols would be simulated by the SORGAM implementation in CMAQ (ver 4.4).

## 2.5 CAMx Modeling System

### 2.5.1 CAMx Overview

The Comprehensive Model with Extensions (CAMx) modeling system is a publicly available (www.camx.com) three-dimensional multi-scale photochemical/aerosol grid modeling system that is developed and maintained by ENVIRON International Corporation. CAMx was developed with all new code during the late 1990s using modern and modular coding practices. This has made the model an ideal platform for the extension to treat a variety of air quality issues including ozone, particulate matter (PM), visibility, acid deposition, and air toxics. The flexible CAMx framework has also made it a convenient and robust host model for the implementation of a variety of mass balance and sensitivity analysis techniques including Process Analysis (IRR and IPR), Decoupled Direct Method (DDM), and the Ozone Source Apportionment Technology (OSAT). Designed originally to address multiscale ozone issues from the urban- to regional-scale, CAMx has been widely used in recent years by a variety regulatory agencies for 1-hr and 8-hr ozone and PM10 SIP modeling studies as well as by several RPOs for regional haze modeling. Key attributes of the CAMx model include the following:

- Two-way grid nesting that supports multi-levels of fully interactive grid nesting (e.g., 36/12/4/1.33 km);
- CB4 or SAPRC99 Chemical Mechanisms;
- Two chemical solvers, the CAMx Chemical Mechanism Compiler (CMC) Fast Solver or the highly accurate Implicit Explicit Hybrid (IEH) solver;
- Multiple numerical algorithms for horizontal transport including the Piecewise Parabolic Method (PPM), Bott, and Smolarkiewicz advection solvers;
- Subgrid-scale Plume-in-Grid (PiG) algorithm to treat the near-source plume dynamics and chemistry from large NOx point source plumes;
- Ability to interface with a variety of meteorological models including the MM5 and RAMS prognostic hydrostatic meteorological models and the CALMET diagnostic meteorological model (others also compatible);
- The Ozone Source Apportionment Technology (OSAT) ozone apportionment technique that identifies the ozone contribution due to geographic source regions and source categories (e.g., mobile, point, biogenic, etc.); and

- The Decoupled Direct Method (DDM) sensitivity method is implemented for emissions and IC/BC to obtain first-order sensitivity coefficients for all gas-phase species.

- Treatment of particulate matter (PM) using an empirical aerosol thermodynamics algorithm.

Culminating extensive model development efforts at ENVIRON and other participating groups, the CAMx (ver 4.10s) code was released in the autumn of 2004 as a truly "One-Atmosphere' models that rigorously integrates the gas-phase ozone chemistry with the simulation of primary and secondary fine and course particulate aerosols. This extension of CAMx to treat PM involved the addition of several science modules to represent important physical processes for aerosols. Noteworthy among these are:

- Two separate treatments of particulate matter (PM), Mechanism 4 (M4) "one-atmosphere" treatment uses two size sections and science modules comparable to CMAQ (e.g., RADM aqueous-phase chemistry and ISORROPIA equilibrium) and a multi-section "full-science" approach using aerosol modules developed at Carnegie Mellon University (CMU).

- Size distribution is represented using the Multi-component Aerosol Dynamics Model (MADM), which uses a sectional approach to represent the aerosol particle size distribution (Pilinis et al., 2000). MADM treats the effects of condensation/evaporation, coagulation and nucleation upon the particle size distribution.

- Inorganic aerosol thermodynamics can be represented using ISORROPIA (Nenes et al, 1998; 1999) equilibrium approach within MADM, or a fully dynamic or hybrid approach can also be used.

- Secondary organic aerosol thermodynamics are represented using the semi-volatile scheme of Strader and co-workers (1999).

- Aqueous-phase chemical reactions are modeled either using the RADM module (like CMAQ) or the Variable Size-Resolution Model (VRSM) of Fahey and Pandis (2001), which automatically determine whether water droplets can be represented by a single 'bulk' droplet-size mode or whether it is necessary to use fine and coarse droplet-size modes to account for the different pH effects on sulfate formation.

CAMx (ver 4.10s) provides two key options to users interested in simulating PM. For CPU-efficient annual PM modeling applications, CAMx may be run using Mechanism 4 (M4) with only two size sections (fine and coarse) and the efficient RADM bulk aqueous-phase module (as used in CMAQ). Alternatively, more rigorous aerosol simulations (perhaps for shorter episode) may be addressed using the version that treats N-size sections (N is typically 10) and the rigorous, but computationally-extensive CMU multi-section aqueous-phase chemistry module.

A PM Source Apportionment Technology (PSAT) has recently been added to CAMx and extensively tested and evaluated. It is currently being used by the MRPO for their BART analysis and may be useful to CENRAP as well. It is currently available on request and will be publicly released on the CAMx website (www.CAMx.com) with the next version of CAMx.

### 2.5.2 CAMx Configuration for CENRAP Annual Modeling

We recommend exercising CAMx (ver 4.10s) using similar science options as CMAQ. However, in some instances, the CMAQ and CAMx model development teams chose different options for characterizing physical and chemical processes, or for implementing the governing equations on modern parallel computers. In these cases, we will utilize the science configurations embodied in the current release of CAMx.

Table 2-4 lists the main CAMx configurations recommended for the CENRAP annual modeling. The latest version of CAMx (ver 4.10s or newer) will be employed and the model will be set up and exercised on the same 36 and 12 km grids as CMAQ. However, for the 12 km grid episodic simulations CAMx would be run using two-way grid nesting instead of the one-way nesting used by CMAQ. The base configuration of CAMx would use 19 vertical layers up to 100mb (~15 km AGL) that exactly match those used by CMAQ. The PPM advection solver would be used along with the spatially varying (Smagorinsky) horizontal diffusion approach. Vertical diffusion in CAMx would be modeled by K-theory. The MM5 simulation using the Pleim-Xiu Land-Surface Model (LSM) and the ACM Planetary Boundary Layer (PBL) scheme would be used in the CAMx base configuration using the MM5CAMx processor that is similar to the CMAQ MCIP2.3 “pass through” option of the MM5 data invoked. CAMx would be exercised with the CB4 gas-phase, RADM aqueous-phase, and CMU/ISORROPIA aerosol chemistry schemes. The SOAP secondary organic aerosol scheme would be used for the base configuration in CAMx.

Note that it may be desirable to exercise CAMx using its “full-science” configuration for selected periods to investigative scientific issues that may be of interest to CENRAP such as:

- The full sectional approach could be used to determine whether allowing secondary PM to grow into the coarse mode affects the model estimates;

- Model could be exercised with chemical active Sea Salt emissions, this could be important for fine particulate and visibility at key coastal sites in the CENRAP domain (e.g., Breton Wilderness Area), especially when looking at clean days or natural background; and

- The full sectional aqueous-phase chemistry module may be important for sulfate formation.

## 2.6 Advantages in Operating Multiple Models

EPA's guidance on model selection for $PM_{2.5}$ SIPs and Regional Haze "reasonable progress demonstrations" do not identify a preferred photochemical grid modeling system, recognizing that at present there is "no single model which has been extensively tested and shown to be clearly superior or easier to use than several alternatives" (EPA, 2001, pg. 169.) The agency recommends that models used for $PM_{2.5}$ SIPS or RH reasonable progress requirements should meet the requirements for alternative models. The CMAQ, CMAQ-MADRID, CMAQ-AIM and CAMx modeling systems all meet these requirements.

We believe that there is potentially significant value in including multiple modeling systems in the CENRAP annual modeling analysis. Our testing and comparisons of the CMAQ and CAMx models for WRAP, VISTAS and other recent $PM_{2.5}$/regional haze applications demonstrates that the models are capable of producing results of comparable accuracy and reliability and having results from both models has many benefits, such as:

- **Diagnosis:** To serve as an efficient diagnostic tool addressing model performance issues that may arise in the establishment of the model annual 2002 and episodic base cases. CMAQ and CAMx both include Process Analysis that can help diagnose model performance. CAMx's suite of diagnostic probing tools plus it's flexi-nesting algorithms make it an attractive tool for assisting in the diagnosis of model performance;

- **Model Evaluation Corroboration:** To provide corroboration of the base case model performance evaluation exercises to be performed with the two models and help identify any compensatory errors in the modeling systems;

- **Emissions Control Response Corroboration:** To provide corroboration of the response of a modeling system to generic and specific future year emissions changes on modeled gas-phase and particulate aerosol concentrations and resultant regional haze impacts;

- **Quantification of Model Uncertainty:** To provide one estimate of the range of uncertainty in the annual and episodic base case simulations, and in the estimate of $PM_{2.5}$ and visibility reductions associated with future emissions change scenarios;

- **Alternative Science:** CAMx and CMAQ contain alternative science algorithms that may elucidate model performance issues with one model or the other or provide an alternative approach for simulating aerosols.

- **Consistency with Other RPOs:** The Midwest RPO (MRPO) may end up using CAMx for their regional haze modeling, whereas the Western (WRAP) and Southeastern (VISTAS) RPOs are currently using CMAQ. As sources in the MRPO, WRAP and VISTAS likely influence visibility at Class I areas in CENRAP and vice versa, having results from a both models would be useful for reconciling any differences.

- **Backup Contingency:** To provide a 'backstop' model in the event that unforeseen difficulties with one model occur.

The benefits of employing a pair of complimentary state-of-science air quality models are thus quite significant and well worth the extra effort. Especially considering that the same MM5 output (through MCIP2.3 and MM5CAMx) and SMOKE output and CMAQ IC/BC files (through CMAQ-to-CAMx emissions and IC/BC converters) can be used to operate CMAQ and CAMx without performing any additional meteorological or emissions modeling.

## 2.7 Model Limitations

All mathematical models possess inherent limitations owing to the necessary simplifications and approximations made in formulating the governing equations, implementing them for numerical solution on fast computers, and in supplying them with input data sets and parameters that are themselves approximations of the full state of the atmosphere and emissions processes. Below, we list some of the more important limitations of the various modeling systems to be employed by CENRAP.

### 2.7.1 MM5

MM5 many different physics options that can drastically alter the predicted meteorological fields. MM5 meteorological estimates are particularly sensitive to the choice of Land Soil Model (LSM) and Planetary Boundary Layer (PBL) model. There are numerous limitations in the MM5 with the LSM and PBL treatment being some of the most important. The MM5 Pleim-Xiu/ACM LSM/PBL physic options selected by CENRAP frequently predicts very low PBL heights that can appear as "holes" in the spatial distribution of PBL heights that don't appear physically realistic and may affect air quality modeling. Although the 2002 annual MM5 model performance in the CENRAP region mostly met performance benchmarks, there were some concerns raised and, in particular, the overstatement of precipitation amounts has been raised as a major concern (Baker, 2004b). Concerns have also been raised concerning the MM5 performance over the western third of the US (Johnson, 2004). The many limitations in MM5 have spawned the development of a new meteorological model, the Weather Research Forecast (WRF) model. However, the WRF model will not be used or tested in the CENRAP modeling.

### 2.7.2 SMOKE

In WRAP and VISTAS a number of undocumented features of SMOKE necessitated re-runs of the emissions processing software to overcome errors and/or ambiguities in source documentation and QA reporting. It is unclear whether similar conditions will be encountered with the SMOKE ver2.0 (or ver2.1) release to be used in CENRAP. Features are continuing to be developed in the SMOKE emissions model. As it is not as mature as some other emission models (e.g., EMS, EPS, etc.) it does not include as many features. We will keep abreast of SMOKE development activities to identify new features that will assist in the CENRAP emissions modeling.

### 2.7.3 CMAQ

Like all air quality models, a major limitation of CMAQ is the emissions, meteorological and IC/BC inputs. Key science limitations in the model itself include the nitrate formation chemistry and the secondary organic aerosol (SOA) module. The CENRAP preliminary modeling found the CMAQ nitrate performance suspect with winter overestimations and summer underestimations (Pun, Chen and Seigneur, 2004; Tonnesen and Morris, 2004). The CENRAP preliminary modeling also found the performance for Organic Carbon (OC) to be less than ideal; much of the OC performance problems is due to deficiencies in the CMAQ SOA module that fails to account for several known processes important to SOA (e.g., polymerization). Other science limitations in the current version of CMAQ include inadequate treatment of sea salt and the assumption that all secondary PM is in the fine mode. Lack of any two-way grid nesting limits the ability of the model to properly resolve point source plumes or urban photochemistry and their effects on more distant Class I areas without a prohibitive number of grid cells. Another limitation of CMAQ is the computational requirements, including the need of excessive disk space.

### 2.7.4 CAMx

The model inputs are also a major limitation in CAMx and CAMx shares many of the formulation deficiencies of CMAQ. Nitrate formation chemistry is also a major limitation, as evident by the preliminary CENRAP modeling. Although CAMx has some more advanced science modules available, such as the VSRM aqueous-phase and MADM dynamic aerosol modules, for annual modeling these modules are too computationally expensive to use.

## 2.8 Model Input Requirements

Each of the CENRAP modeling system components have significant data base requirements. These data needs fall into two categories: those required for model setup and operation, and those required for model evaluation testing. Below, we identify the main input data base requirements for the meteorological, emissions, and air quality models.

### 2.8.1 MM5

The databases required to set up, exercise, and evaluate the MM5 model for the annual 2002 episode consist of various fixed and variable inputs.

- Topography: High resolution (e.g., 30 sec to 5 min) topographic information derived from the Geophysical Data Center global data sets from the National Center for Atmospheric Research (NCAR) terrain databases are available for prescribing terrain elevations throughout the 36 km and 12 km grid domain.

- Vegetation Type and Land Use: Vegetation type and land use information on the 36 km grid may be developed using the NCAR/PSU 10 min. (~18.5 km) databases while for the 12 km grids, the United States Geological Survey (USGS) data are available.

- Atmospheric Data: Initial and boundary conditions to the MM5 may be developed from operationally analyzed fields derived from the National Center for Environmental Predictions (NCEP) ETA (40 km resolution) following the procedures outlined by Stauffer and Seaman (1990). These 3-hr synoptic-scale initialization data the horizontal wind components (u and v), temperature (T), and relative humidity (RH) at the standard pressure levels, plus sea-level pressure (SLP) and ground temperature ($T_g$). Here, $T_g$ represents surface temperature over land and sea-surface temperature over water.

- Water Temperature: Water temperatures required on both 36 km and 12 km grids can be derived from the ETA skin temperature variable. These temperatures are bi-linearly interpolated to each model domain and, where necessary, filtered to smooth out irregularities.

- Clouds and Precipitation: While the non-hydrostatic MM5 treats cloud formation and precipitation directly through explicit resolved-scale and parameterized sub-grid scale processes, the model does not require precipitation or cloud input. The potential for precipitation and cloud formation enters through the thermodynamic and cloud processes formulations in the model. The only precipitation-related input required is the initial mixing ratio field that is developed from the NWS and NMC data sets previously discussed.

- Multi-Scale FDDA: The standard "multi-scale" data assimilation strategy to be used on the 36 km and 12 km grids will objectively analyzed three-dimensional fields produced every 3-hr from the NWS rawinsonde wind, temperature, and mixing ratio data, and similar analyses generated every three hours from the available NWS surface data.

### 2.8.2 SMOKE

The databases required to set up and operate SMOKE for the 2002 annual simulation are as follows:

- Area Source emissions in IDA format
- NonRoad source emissions in IDA format
- Stationary Point Source emissions in IDA format
- CEM emissions, day specific for 2002
- Wildfire, prescribed burns and agricultural burning emissions, day specific for 2002
- On-road Motor Vehicle VMT and activity data
- MOBILE6.2 input parameters

Also required for annual modeling are data files specific for:

- Temporal allocation
- Spatial allocation
- Speciation

Chapter 5 discusses the SMOKE data input requirements and data sources in detail.

### 2.8.3 CMAQ

As described in more detail in Chapter 5, the CMAQ Chemical Transport Model (CTM) requires the following inputs:

- Three-dimensional hourly meteorological fields that will be generated by the CMAQ MCIP2.3 processing of the MM5 output;
- Three-dimensional hourly emissions generated by SMOKE;
- Initial conditions and boundary conditions (IC/BC);
- Topographic information;
- Land use categories; and
- Photolysis rates generated by the CMAQ JPROC processor.

### 2.8.4 CAMx

CAMx model inputs include (see Chapter 5):

- Three-dimensional hourly meteorological fields generated by MM5CAMx processing of the MM5 output;
- Two-dimensional low-level (surface layer) emissions and elevated point source emissions generated by the CMAQ-to-CAMx emissions processor.
- IC/BC inputs generated by the CMAQ-to-CAMx IC/BC processors;
- Photolysis rates look up table;
- Albedo/Haze/Ozone Column input file;
- Land use and topography

**Table 2-1.** MM5 meteorological model configuration for CENRAP 2002 Annual modeling (Johnson, 2004).

| Science Options | Configuration | Details/Comments |
|---|---|---|
| Model Code | MM5 version 3.63 | Grell et al., 1994 |
| Horizontal Grid Mesh | 36 km (12 km for episodes) | |
| 36 km grid | 165 x 129 cells | |
| 12 km grid | TBD | |
| Vertical Grid Mesh | 34 layers | Vertically varying; sigma pressure coord. |
| Grid Interaction | No Feedback | IFEED=0 |
| Initialization | Eta first guess fields/LittleR | |
| Boundary Conditions | Eta first guess fields/LittleR | |
| Microphysics | Reisner I Mixed Ice | Look up table |
| Cumulus Scheme | Kain-Fritsch 2 | On 36 and 12 km Grids |
| Planetary Boundary Layer | ACM PBL | |
| Radiation | RRTM | |
| Vegetation Data | USGS | 24 Category Scheme |
| Land Surface Model | Pleim-Xiu Land Surface Model (LSM) | |
| Shallow Convection | None | |
| Sea Surface Temperature | Eta Skin | Spatially varying |
| Thermal Roughness | Garratt | |
| Snow Cover Effects | None | |
| 4D Data Assimilation | Analysis Nudging on 36 and 12 | |
| Surface Nudging | None | |
| Integration Time Step | 90 seconds | |
| Simulation Periods | Annual 2002 for 36 km | 12 km episodic only |
| Platform | Linux Cluster | Done at IDNR |

**Table 2-2.** SMOKE emissions model configuration for CENRAP Annual modeling.

| Emissions Component | Configuration | Details/Comments |
|---|---|---|
| Emissions Model | SMOKE ver 2.0 | SMOKE ver 2.1 may be used if operational in time |
| Horizontal Grid Mesh | 36/12 km | |
| 36 km grid | 148 x 112 cells | RPO Unified Grid |
| 12 km grid | TBD | |
| Area Source Emissions | CENRAP Domain: CENRAP State 2002 EI | Updated '02 developed by CENRAP states |
| | Other States: '02 NEI augmented with other 2002 | Generated from EPA NEI02 v.1 and RPO interaction |
| | Mexico/Canada Emissions: | Same as used in WRAP |
| On-Road Mobile Sources | CENRAP Domain: CENRAP VMT data | Updated '02 developed by CENRAP states (Reid et al., 2004) |
| | Other States: EPA '02 NEI augmented with other 2002 | Generated from EPA NEI02 v.1 and RPO interaction |
| | Mexico/Canada Emissions: | 2000 Canadian Inventory? |
| Point Sources | CENRAP Domain: CENRAP State 2002 EI | Updated '02 developed by CENRAP states and stakeholders |
| | Other States: EPA '02 NEI augmented with other 2002 | Generated from EPA NEI02 v.1 and RPO interaction |
| | Mexico/Canada Emissions: | 2000 Canadian Inventory? |
| Off-Road Mobile Sources | CENRAP Domain: CENRAP State 2002 EI | Updated '02 developed by CENRAP states |
| | Other States: EPA '02 NEI augmented with other 2002 | Generated from EPA NEI02 v.1 and RPO interaction |
| | Mexico/Canada Emissions: | 2000 Canadian Inventory? |
| Biogenic Sources | SMOKE BEIS-3 | BELD3 vegetative database |
| Temporal Adjustments | Seasonal, day, hour | Based on latest collected information and CEM-based profiles |
| Chemical Speciation | Revised CB4 Chemical Speciation | Updated January 2004 |
| Gridding | Revised EPA Spatial Surrogates Used | Gridding of surrogates from http://www.epa.gov/ttn/chief/emch/spatial/ |
| Growth and Controls | CENRAP Contractor TBD | Base Cases defined by CENRAP Workgroups |
| Quality Assurance | QA Tools in SMOKE 2.0 | Follow QAPP |
| Simulation Periods | Annual 2002 for 36 km | Episodic periods at 12 km TBD |

**Table 2-3.** CMAQ air quality model configuration for CENRAP annual modeling.

| Science Options | Configuration | Details/Comments |
|---|---|---|
| Model Code | CMAQ (ver 4.4) | Available at: www.cmascenter.org |
| Horizontal Grid Mesh | 36 km annual/12 km episodic | 36 km covering continental U.S; 12 km TBD |
| 36 km grid | 148 x 112 cells | RPO National Grid |
| 12 km grid | TBD | |
| Vertical Grid Mesh | 19 Layers | First 17 layers sync'd w/ MM5 |
| Grid Interaction | One-way nesting | |
| Initial Conditions | ~15 days full spin-up | Separately run 4 quarters of 2002 |
| Boundary Conditions | 2002 GEOS-CHEM day-specific | 2002 GEOS-CHEM day specific 3-hourly |
| Emissions | | |
| Baseline Emissions Processing | See SMOKE (Ver 2.0) model configuration | MM5 Meteorology input to SMOKE, CMAQ |
| Sub-grid-scale Plumes | No Plume-in-Grid (PinG) | PinG for PM not available yet |
| Chemistry | | |
| Gas Phase Chemistry | CBM-IV | |
| Aerosol Chemistry | AE3/ISORROPIA | |
| Secondary Organic Aerosols | Secondary Organic Aerosol Model (SORGAM) | Schell et al., (2001) |
| Cloud Chemistry | RADM-type aqueous chemistry | Includes subgrid cloud processes |
| N2O5 Reaction Probability | 0.01 – 0.001 | |
| Meteorological Processor | MCIP ver 2.3 | Includes dry deposition and snow cover updates |
| Horizontal Transport | | |
| Eddy Diffusivity Scheme | K-theory with Kh grid size dependence | Multiscale Smagorinsky (1963) approach |
| Vertical Transport | | |
| Eddy Diffusivity Scheme | K-theory | |
| Diffusivity Lower Limit | Kzmin = 1.0 | Run MCIP2, 3 with Kz-min.=0.1 |
| Deposition Scheme | M3dry | Directly linked to Pleim-Xiu Land Surface Model parameters |
| Numerics | | |
| Gas Phase Chemistry Solver | Euler Backward Iterative (EBI) solver | |
| Horizontal Advection Scheme | Piecewise Parabolic Method (PPM) scheme | |
| Simulation Periods | Annual 2002 for 36 km | Episodic periods at 12 km TBD |
| Integration Time Step | TBD | 15 minute coupling time step |

**Table 2-4** CAMx air quality model configuration for CENRAP annual modeling.

| Science Options | Configuration | Details |
|---|---|---|
| Model Code | CAMx (ver 4.10s) | Available at: www.camx.com |
| Horizontal Grid Mesh | 36 km annual/12 km episodic | 36 km covering continental U.S; 12 km TBD |
| 36 km grid | 148 x 112 cells | |
| 12 km grid | TBD | |
| Vertical Grid Mesh | 19 Layers | 17 Layers sync'd w/ MM5 |
| Grid Interaction | Two-way nesting | |
| Initial Conditions | ~15 days full spin-up | Separately run 4 quarters of 2002 |
| Boundary Conditions | 2002 GEOS-CHEM day-specific | 2002 GEOS-CHEM day specific 3-hourly |
| Emissions | | |
| Baseline Emissions Processing | See SMOKE (Ver 2.0) model configuraton | MM5 Meteorology input to SMOKE, CAMx |
| Sub-grid-scale Plumes | No Plume-in-Grid (PinG) | PinG for PM not available yet |
| Chemistry | | |
| Gas Phase Chemistry | CBM-IV | with Isoprene updates |
| Aerosol Chemistry | ISORROPIA equilibrium | Dynamic and hybrid also available |
| Secondary Organic Aerosols | SOAP | |
| Cloud Chemistry | RADM-type aqueous chemistry | Alternative is CMU multi-section aqueous chemistry |
| N2O5 Reaction Probability | None | |
| Meteorological Processor | MM5CAMx | |
| Horizontal Transport | | |
| Eddy Diffusivity Scheme | K-theory with Kh grid size dependence | |
| Vertical Transport | | |
| Eddy Diffusivity Scheme | K-Theory | |
| Diffusivity Lower Limit | Kzmin = 1.0 (or other) | Run MM5CAMx with Kz-min=1.0 |
| Planetary Boundary Layer | No Patch | |
| Deposition Scheme | Wesely | |
| Numerics | | |
| Gas Phase Chemistry Solver | CMC Fast Solver | |
| Horizontal Advection Scheme | Piecewise Parabolic Method (PPM) scheme | |
| Simulation Periods | Annual 2002 at 36 km | Episodic periods at 12 km TBD |
| Integration Time Step | Wind speed dependent | |

## 3.0 EPISODE SELECTION

This chapter provides a brief overview of reasons for the selection of the 2002 annual period for the CENRAP regional haze modeling.

### 3.1 Overview of EPA Guidance

EPA's current draft guidance on PM2.5/Regional Haze modeling (EPA, 2001) identifies specific goals to consider when selecting one or more episodes for use in demonstrating reasonable progress in attaining the regional haze NAAQS. However, since there is much in common with the goals for selecting episodes for annual and episodic PM2.5 attainment demonstrations as well as regional haze, EPA's guidance addresses all three in a common document. More recently, EPA has published an updated summary of PM2.5 and Regional Haze Modeling Guidance (Timin, 2002) that serves, in some respects, as in interim placeholder until the final guidance is issued as part of the PM2.5/regional haze NAAQS implementation process that is expected during 2004.

EPA recommends that episode selection derive from three principal criteria:

- A variety of meteorological conditions should be covered, this includes the types of meteorological conditions that produce the Worst 20% and Best 20% visibility days at Class I areas in the CENRAP States;

- To the extent possible, the modeling data base should include days for which extensive data bases (i.e. beyond routine aerometric and emissions monitoring) are available; and

- Sufficient days should be available such that relative reduction factors (RRFs) can be based on several (i.e., $\geq$ 15) days.

For regional haze modeling, the guidance goes further by suggesting that the preferred approach is to model a full, *representative* year (EPA, 2001, pg. 188). Moreover, the required RRF values should be based on model results averaged over the 20% worst and 20% best visibility days determined for each Class I are based on monitoring data from the 2000 – 2004 baseline period. More recent EPA guidance (Timin, 2002) suggests that states should model at least 10 worst and 10 best visibility days at each Class 1 area. EPA also lists several 'other considerations' to bear in mind when choosing potential PM/regional haze episodes including: (a) choose periods which have already been modeled, (b) choose periods which are drawn from the years upon which the current design values are based, (c) include weekend days among those chosen, and (d) choose modeling periods that meet as many episode selection criteria as possible in the maximum number of nonattainment or Class I areas as possible.

### 3.2 Selection of CY 2002 For CENRAP Annual Modeling

Due to limited available resources CENRAP, at least initially, is limited to modeling a single Calendar Year (CY). The RHR uses the five-year baseline of 2000-2004 as the starting point for

projecting future-year visibility. Thus, the modeling year should be selected from this five-year baseline period. The 2002 calendar year was selected for the following reasons:

- Based on available information, appears to be a fairly typical year in terms of meteorology for the future-year baseline of 2000-2004;
- 2003 and 2004 appeared to be colder and wetter than typical in the eastern US;
- The enhanced IMPROVE and IMPROVE Protocol and Supersites PM monitoring data were fully operational by 2002; and
- 2002 is being considered by the other RPOs.

# 4.0 MODELING DOMAINS AND DATA AVAILABILITY

This chapter summarizes the model domain definitions for the CENRAP annual modeling including the model domain, resolution, map projections and nesting schemes for high resolution sub-domains.

## 4.1 Horizontal Modeling Domain

The 36 km continental US horizontal domain for each of the models will be identical to those used by WRAP and VISTAS. The CMAQ and CAMx air quality modeling domain is nested in the MM5 domain. The selection of the MM5 domain is described by Johnson (2004). Figure 4-1 shows the MM5 horizontal domain as the outer most, blue grid. Also shown in Figure 4-1 is the CMAQ and CAMx 36 km domain nested in the MM5 domain. To achieve finer spatial resolution in the CENRAP states we will also use a nested high resolution grid with a 12 km grid resolution whose definition will be determined later.

Both MM5 and CMAQ will employ the Regional Planning Organization (RPO) unified grid definition for the 36 km continental domain for the CENRAP annual modeling. The RPO unified grid consists of a Lambert-Conformal map projection using the map projections parameters listed in Table 4-1.

**Table 4-1.** RPO Unified grid definition.

| PARAMETER | VALUE |
|---|---|
| projection | Lambert-Conformal |
| alpha | 33 degrees |
| beta | 45 degrees |
| x center | 97 degrees |
| y center | 40 degrees |

The MM5 36 km grid include 164 cells in the east-west dimension and by 128 cells in the north-south dimension. The CMAQ/CAMx 36 km grid include 148 cells in the east-west dimension and 112 cells in the north-south dimension. Because the MM5 model is also nested in the Eta model, there is a possibility of boundary effects near the MM5 boundary that occur as the Eta meteorological variables are being simulated by MM5 and must come into dynamic balance with MM5's algorithms. Thus, a larger MM5 domain was selected to provide a buffer of 8 to 9 grid cells around each boundary of the CMAQ/CAMx 36 km domain. This is designed to eliminate any errors in the meteorology from boundary effects in the MM5 simulation at the interface of the MM5 and Eta models. The buffer region used here exceeds the EPA suggestion of at least 5 grid cell buffer at each boundary.

Table 4-2 lists the number of rows and columns and the definition of the X and Y origin (i.e., the southwest corner) for the 36 km for both MM5 and CMAQ/CAMx. Note that the CMAQ/CAMx grid is rotated 90 degrees relative to the MM5 grid, so rows and columns are reversed. In Table 4-2 "Dot" refers to the grid mesh defined at the vertices of the grid cells while "cross" refers to the grid mesh defined by the grid cell centers. Thus, the dimension of the dot mesh is equal to the cross mesh plus one. Finally, we note that the grid definition for the SMOKE emissions model, CMAQ Meteorology Chemistry Interface Processor (MCIP), CMAQ Chemical Transport Model (CCTM), MM5CAMx processor and CAMx model are identical.

**Table 4-2.** Grid definitions for MM5 and CMAQ/CAMx.

| MODEL | COLUMNS DOT(CROSS) | ROWS DOT(CROSS) | XORIGIN | YORIGIN |
|---|---|---|---|---|
| MM5 36km | 129 (128) | 165 (164) | -2952000 | -2304000 |
| CMAQ/CAMx 36km | 149 (148) | 113 (112) | -2736000 | -2088000 |

## 4.2 Vertical Modeling Domain

The CMAQ and CAMx vertical structure is primarily defined by the vertical grid used in the MM5 modeling. The MM5 model employed a terrain following coordinate system defined by pressure, using 34 layers that extend from the surface to the 100 mb. Table 4-3 list the layer definitions for both MM5 and for CMAQ and CAMx. We will use the exactly same vertical layer structure in CAMx as in CMAQ. A layer averaging scheme is adopted for CMAQ/CAMx to reduce the computational cost of the CMAQ and CAMx simulations. The effects of layer averaging were evaluated by WRAP and VISTAS and found to have a relatively minor effect on the model performance metrics when both the 34 layer and a 19 layer CMAQ model simulations were compared to ambient monitoring data (Morris et al., 2004a).

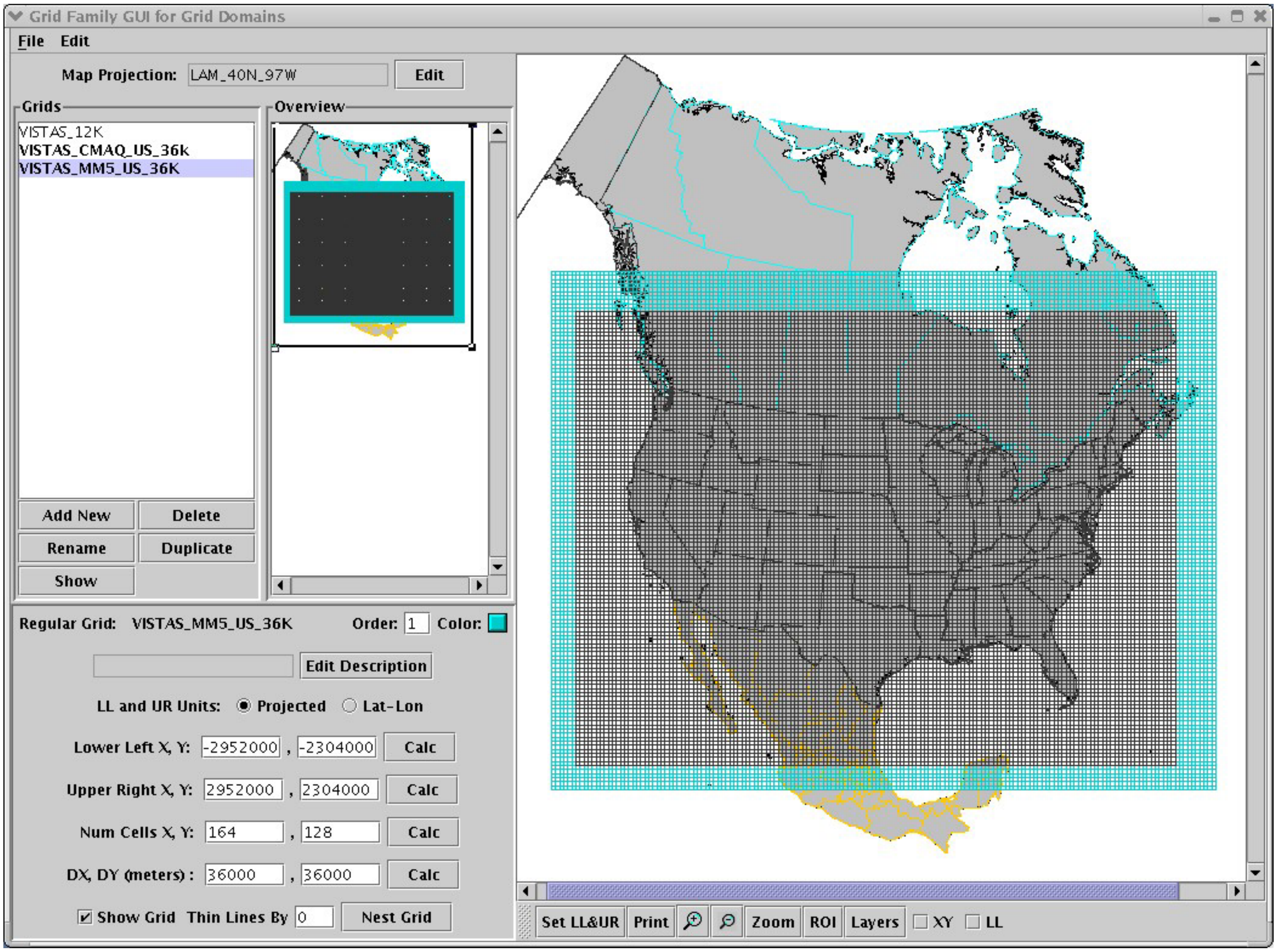


**Figure 4-1**. Nesting of 36-km CMAQ/CAMx grid in the MM5 36-km grid.

**Table 4-3.** Vertical layer definition for MM5 simulations (left most columns), and approach for reducing CMAQ/CAMx layers by collapsing multiple MM5 layers (right columns).

| MM5 | | | | | CMAQ 19L | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Layer | Sigma | Pres(mb) | Height(m | Depth(m) | Layer | Sigma | Pres(mb) | Height(m) | Depth(m |
| **34** | **0.000** | **100** | **14662** | **1841** | **19** | **0.000** | **100** | **14662** | **6536** |
| 33 | 0.050 | 145 | 12822 | 1466 | | 0.050 | 145 | | |
| 32 | 0.100 | 190 | 11356 | 1228 | | 0.100 | 190 | | |
| 31 | 0.150 | 235 | 10127 | 1062 | | 0.150 | 235 | | |
| 30 | 0.200 | 280 | 9066 | 939 | | 0.200 | 280 | | |
| **29** | **0.250** | **325** | **8127** | **843** | **18** | **0.250** | **325** | **8127** | **2966** |
| 28 | 0.300 | 370 | 7284 | 767 | | 0.300 | 370 | | |
| 27 | 0.350 | 415 | 6517 | 704 | | 0.350 | 415 | | |
| 26 | 0.400 | 460 | 5812 | 652 | | 0.400 | 460 | | |
| **25** | **0.450** | **505** | **5160** | **607** | **17** | **0.450** | **505** | **5160** | **1712** |
| 24 | 0.500 | 550 | 4553 | 569 | | 0.500 | 550 | | |
| 23 | 0.550 | 595 | 3984 | 536 | | 0.550 | 595 | | |
| **22** | **0.600** | **640** | **3448** | **506** | **16** | **0.600** | **640** | **3448** | **986** |
| 21 | 0.650 | 685 | 2942 | 480 | | 0.650 | 685 | | |
| **20** | **0.700** | **730** | **2462** | **367** | **15** | **0.700** | **730** | **2462** | **633** |
| 19 | 0.740 | 766 | 2095 | 266 | | 0.740 | 766 | | |
| **18** | **0.770** | **793** | **1828** | **259** | **14** | **0.770** | **793** | **1828** | **428** |
| 17 | 0.800 | 820 | 1569 | 169 | | 0.800 | 820 | | |
| **16** | **0.820** | **838** | **1400** | **166** | **13** | **0.820** | **838** | **1400** | **329** |
| 15 | 0.840 | 856 | 1235 | 163 | | 0.840 | 856 | | |
| **14** | **0.860** | **874** | **1071** | **160** | **12** | **0.860** | **874** | **1071** | **160** |
| 13 | 0.880 | 892 | 911 | 158 | **11** | **0.880** | **892** | **911** | **158** |
| **12** | **0.900** | **910** | **753** | **78** | **10** | **0.900** | **910** | **753** | **155** |
| 11 | 0.910 | 919 | 675 | 77 | | 0.910 | 919 | | |
| **10** | **0.920** | **928** | **598** | **77** | **9** | **0.920** | **928** | **598** | **153** |
| 9 | 0.930 | 937 | 521 | 76 | | 0.930 | 937 | | |
| **8** | **0.940** | **946** | **445** | **76** | **8** | **0.940** | **946** | **445** | **76** |
| **7** | **0.950** | **955** | **369** | **75** | **7** | **0.950** | **955** | **369** | **75** |
| **6** | **0.960** | **964** | **294** | **74** | **6** | **0.960** | **964** | **294** | **74** |
| **5** | **0.970** | **973** | **220** | **74** | **5** | **0.970** | **973** | **220** | **74** |
| **4** | **0.980** | **982** | **146** | **37** | **4** | **0.980** | **982** | **146** | **37** |
| **3** | **0.985** | **986.5** | **109** | **37** | **3** | **0.985** | **986.5** | **109** | **37** |
| **2** | **0.990** | **991** | **73** | **36** | **2** | **0.990** | **991** | **73** | **36** |
| **1** | **0.995** | **995.5** | **36** | **36** | **1** | **0.995** | **995.5** | **36** | **36** |
| **0** | **1.000** | **1000 0 0** | | **0** | **0** | **1.000** | **1000 0 0** | | |

## 4.3 Data Availability

The CMAQ and CAMx modeling systems require emissions, meteorological, initial and boundary condition (IC/BC) and ozone column data for defining the inputs.

### 4.3.1 Emissions Data

The base year emissions inventory for CENRAP annual modeling and CENRAP states will be founded on revised 2002 emissions developed by CENRAP emission inventory contractor (Strait, Roe and Vukovich, 2004). Emissions for the non-CENRAP states, Mexican and Canadian will be based on the latest available inventories that are being used by WRAP and VISTAS. For purposes of air quality model validation, actual 2002 calendar year emissions for Electrical Generating Units (EGU) and fire activity will be used as available. Whereas for strategy and future year emission runs, "typical year" emissions for these categories will be processed for 2002 and the future years. A 2002 CENRAP state emission inventory is expected by the end of November 2004 and will be used in the model performance evaluation

As necessary, all emissions will be converted to Inventory Data Analyzer (IDA) formatted versions and the data will be processed for air quality modeling using Version 2.0 (or Version 2.1) of the Sparse Matrix Operating Kernel Emissions (SMOKE) model. Included in these runs will be the temporal and speciation profiles and cross-reference data provided with the version 2.0 (or 2.1) release of the model augmented with any recommended and approved emission profile data provided by the emissions inventory contractor, obtained from EPA, or prepared by the Study Team prior to initial emissions modeling. Spatial allocation of the emissions will be based on profiles and spatial allocation factors developed for the National RPO grid. Additional description of emissions processing is described in Chapter 5 and emissions QA is described in Chapter 6.

### 4.3.2 Air Quality

Data from ambient monitoring networks for both gas and aerosol species are used in the model performance evaluation. Table 4-4 summarizes ambient monitoring networks. Data have been compiled for all networks except the PAMS and PM Supersites. Figure 4-2 displays the locations of monitoring sites in and near the CENRAP States.

**Table 4-4.** Overview of ambient data monitoring networks.

| Monitoring Network | Chemical Species Measured | Sampling Period | Data Availability/Source |
|---|---|---|---|
| The Interagency Monitoring of Protected Visual Environments (**IMPROVE**) | Speciated PM25 and PM10 (see species mappings) | 1 in 3 days; 24 hr average | http://vista.cira.colostate.edu/improve/Data/IMPROVE/improve_data.htm |
| Clean Air Status and Trends Network (**CASTNET**) | Speciated PM25, Ozone (see species mappings) | Approximately 1-week average | http://www.epa.gov/castnet/data.html |
| National Atmospheric Deposition Program (**NADP**) | Wet deposition (hydrogen (acidity as pH), sulfate, nitrate, ammonium, chloride, and base cations (such as calcium, magnesium, potassium and sodium)), Mercury | 1-week average | http://nadp.sws.uiuc.edu/ |
| Air Quality System (**AQS**) Aka Aerometric Information Retrieval System (**AIRS**) | CO, NO2, O3, SO2, PM25, PM10, Pb | Typically hourly average | http://www.epa.gov/air/data/ |
| Speciation Trends Network (**STN**) | Speciated PM | 24-hour average | http://www.epa.gov/ttn/amtic/amticpm.html |
| Southeastern Aerosol Research and Characterization (**SEARCH**)<br><br>(Southeastern US only) | 24-hr PM25 (FRM Mass, OC, BC, SO4, NO3, NH4, Elem.); 24-hr PM coarse (SO4, NO3, NH4, elements); Hourly PM2.5 (Mass, SO4, NO3, NH4, EC, TC); Hourly gases (O3, NO, NO2, NOy, HNO3, SO2, CO) | Hourly or 24-hour average, depending on parameter. | Electric Power Research Institute (EPRI), Southern Company, and other companies.<br><br>http://www.atmospheric-research.com |
| EPA Particulate Matter Supersites<br><br>(Includes St. Louis in the CENRAP region) | Speciated PM25 | | http://www.epa.gov/ttn/amtic/supersites.html |
| Photochemical Assessment Monitoring Stations (**PAMS**) | Varies for each of 4 station types. | | http://www.epa.gov/ttn/amtic/pamsmain.html |
| National Park Service Gaseous Pollutant Monitoring Network | Acid deposition (Dry; SO4, NO3, HNO3, NH4, SO2), O3, meteorological data | Hourly | http://www2.nature.nps.gov/ard/gas/netdata1.htm |

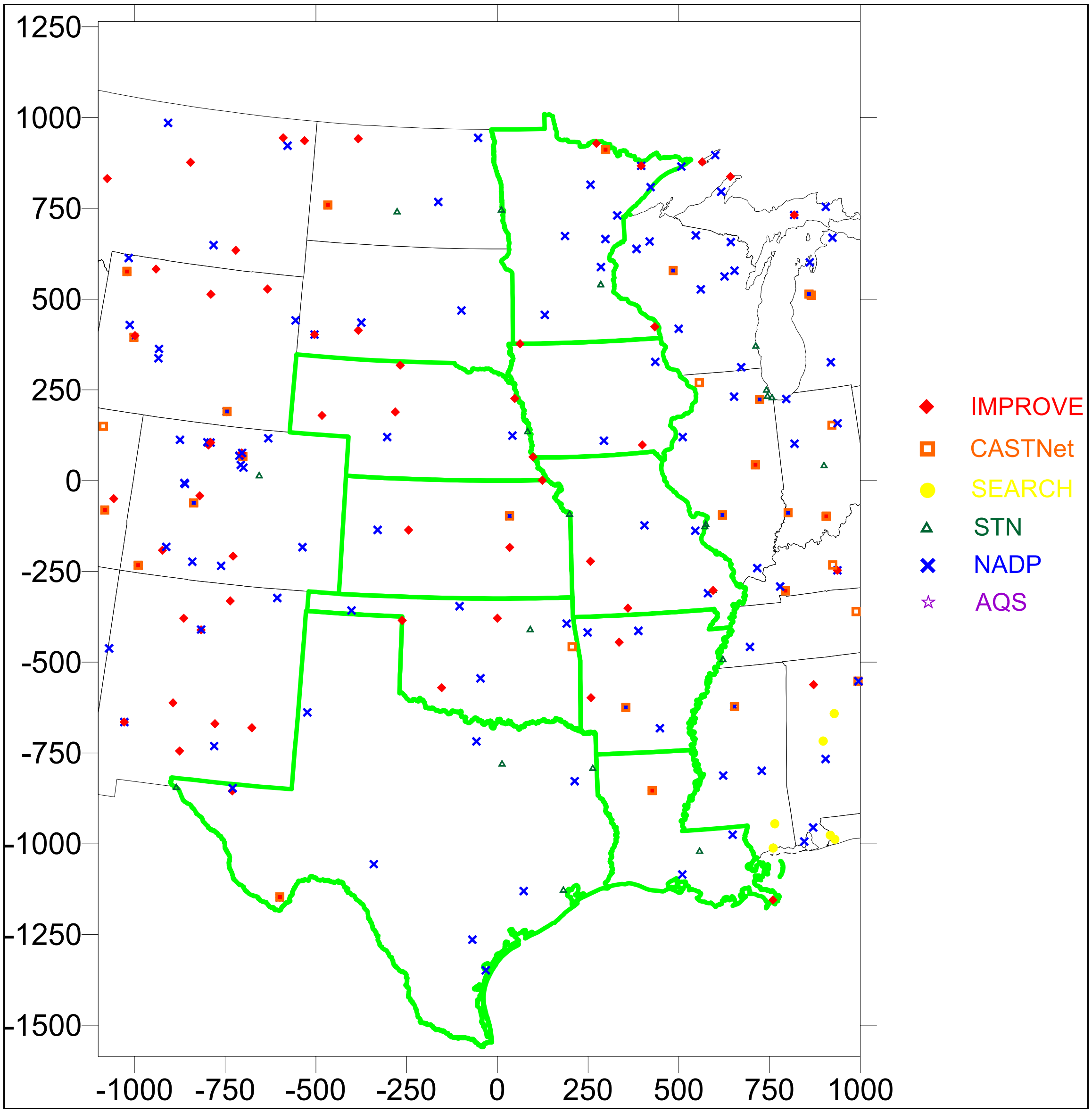


**Figure 4-2a.** Locations of IMPROVE, CASTNet, SEARCH, STN and NADP monitoring sites in and near the CENRAP States.

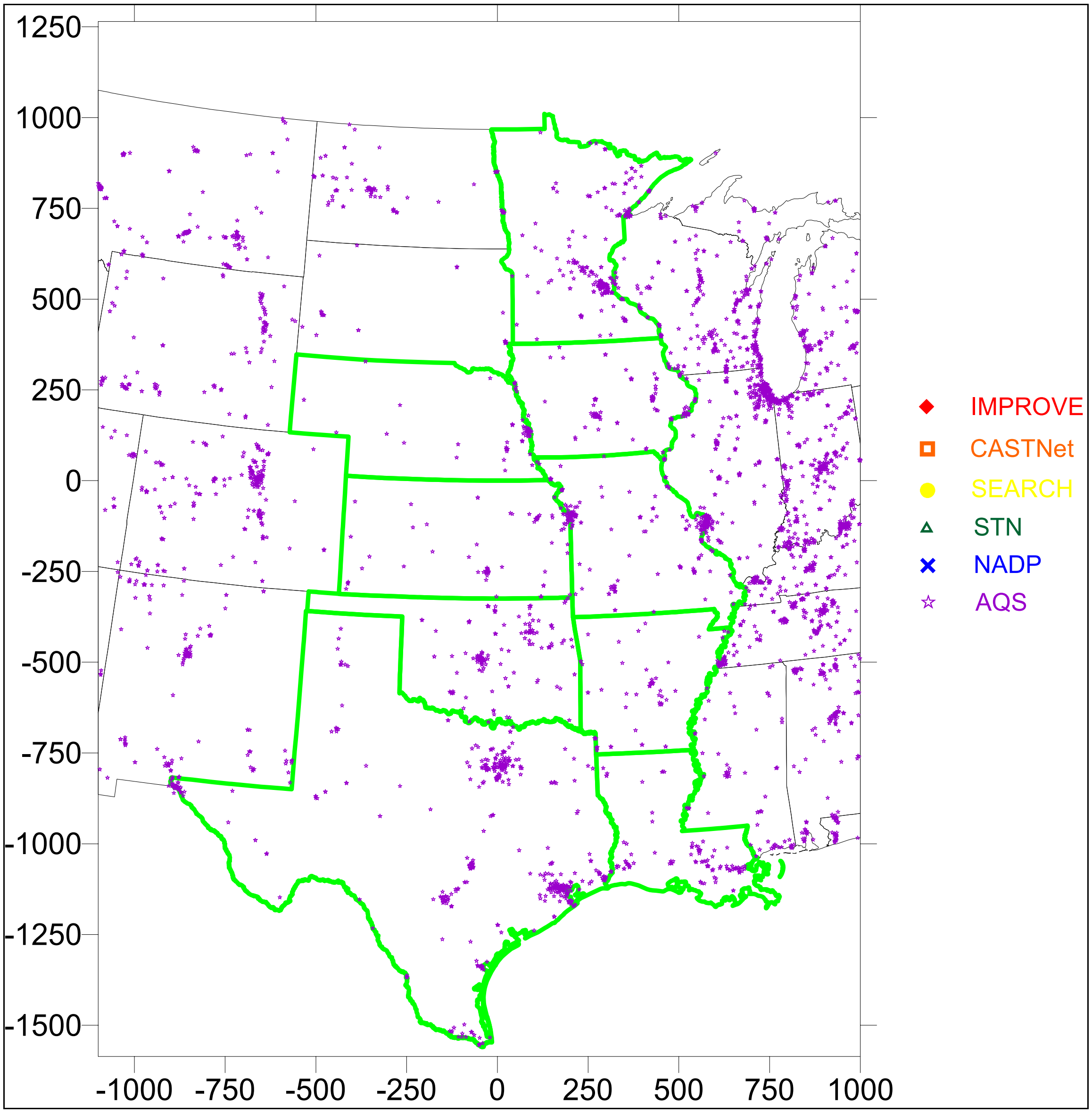


**Figure 4-2b.** Locations of AQS monitoring sites in and near the CENRAP States.

### 4.3.3 Ozone Column Data

Additional data used in the air quality modeling include the Total Ozone Mapping Spectrometer (TOMS). TOMS data is available for 24-hour average and is obtained from http://toms.gsfc.nasa.gov/eptoms/ep.html. The TOMS data is used in the CMAQ (JPROC) and CAMx (TUV) radiation model to calculate photolysis rates.

### 4.3.4 Meteorological Data

Meteorological data are being generated using the MM5 prognostic meteorological model by the Iowa Department of Natural Resources (IDNR for the 2002 36 km continental US and by EPA Region VII for the 12 km episodic modeling). IDNR is operating the MM5 at 5-day increments for 2002 on the 36 km a ~15 day spin up period for the end of December 2001. Details on the CENRAP 2002 36 km MM5 modeling can be found in Johnson (2004). The Region VII 12 km episodic modeling is ongoing.

### 4.3.5 Initial and Boundary Conditions Data

The CMAQ default Initial Concentrations (ICs) will be used for both CMAQ and CAMx along with a ~15 day spin up period to eliminate any significant influence of the ICs.

The CMAQ and CAMx Boundary Conditions (BCs) will be based on results from a 2002 GEOS-CHEM global clime model simulation. The 2002 GEOS-CHEM model output has been processed to define day-specific high time resolved (i.e., 3-hourly) CMAQ and CAMx BCs for 2002.

## 5.0 MODEL INPUT PREPARATION PROCEDURES

In this section we describe the procedures to be used to develop the CMAQ and CAMx model inputs for the CENRAP 2002 annual modeling on the 36 km continental US grid and potentially additional ~4 week episodic simulations on a 12 km grid covering the central states. The development of the CMAQ and CAMx meteorological and emissions inputs are discussed first followed by the science options to be used by CMAQ and CAMx. The procedures for developing the initial and boundary conditions and photolysis rates inputs are then discussed along with the model application procedures.

## 5.1 Meteorological Inputs to Emissions and Air Quality Models

The emissions and air quality models require certain meteorological input data including wind fields, estimates of turbulent eddy dispersion, humidity, temperature, clouds, and actinic flux. Spatially gridded and hourly varying meteorological data are needed to estimate biogenic, mobile source emissions, and plume-rise for large, elevated point sources. Meteorological data are needed to drive chemical transport models for solving atmospheric diffusion and chemistry equations for model species. Because observed data are not available for the full gridded model domain, numerical meteorological models are used to provide these inputs.

The National Center for Atmospheric Research (NCAR)/Pennsylvania State University (PSU) Fifth-Generation Mesoscale Model (MM5) (v3.63) is being used by the Iowa Department of Natural Resources (IDNR) for CENRAP to simulate meteorology at a 36-km resolution for calendar year 2002 (plus the end of 2001) over the entire continental United States and including portions of Canada, Mexico, and the Atlantic and Pacific Oceans. MM5 is also being applied over the central states using a 12 km resolution grid for portions of the 2002 calendar year. The MM5 is a three-dimensional prognostic meteorological model that is used not only for meteorology studies but also for air quality studies. Some of the physics used in the simulation include nonhydrostatic dynamics; four-dimensional data assimilation of wind, temperature, and mixing ratio; explicit treatment of moisture; cumulus cloud parameterization; vertical mixing of momentum in the mixed layer; PBL process parameterization; atmospheric radiation; sea ice treatment; and snow cover (see Chapter 2 for more details).

After the MM5 simulation is completed, the MM5 output files are transferred to the CENRAP Emissions and Air Quality Modeling Team and analyzed by the Meteorological Gatekeeper. The Meteorological Gatekeeper performs two main roles. To provide an independent evaluation of the 2002 MM5 simulation that also serves to determine whether the MM5 data have been transferred correctly from the CENRAP. And to process the 2002 MM5 output using Version 2.3 of the Models-3 CMAQ Meteorological-Chemical Interface Processor (MCIP) and the MM5CAMx processor to generate meteorological fields that will be used for emissions processing and air quality simulations.

### 5.1.1 MCIP Reformatting Methodology

The Models-3 Community Multiscale Air Quality (CMAQ) modeling system is designed to simulate multiscale (urban and regional) and multi-pollutant (oxidants, acid deposition, and particles) air quality problems. But before running the CMAQ Chemical Transport Model (CCTM), the MM5 generated meteorological data must be pre-processed and converted to Models-3 consistent data structures. MCIP version 2.3 will be used to preprocess the MM5 meteorological output. The "pass through" option in MCIP will be used in the CENRAP annual modeling. One of MCIP's functions is to translate meteorological parameters from the output of the Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) Mesoscale Modeling System Generation 5 (MM5) to the Models-3 input/output applications program interface (I/O API format) which is required for operation of Models-3 CMAQ processors. Some other necessary parameters not available from the meteorological model are estimated with appropriate diagnostic algorithms in the program. The key functions of MCIP include:

1. Reading in meteorological model output files
2. Extraction of meteorological data for CTM window domain
3. Interpolation of coarse meteorological model output for finer grid
4. Collapsing of meteorological profile data if coarser vertical resolution data is requested
5. Computation or passing through surface and PBL parameters
6. Diagnosing of cloud parameters
7. Computation of species-specific dry deposition velocities
8. Generation of coordinate dependent meteorological data for the generalized coordinate CCTM simulation
9. Output meteorological data in Models-3 I/O API format

The MCIP processor transforms the data into I/O API format while also calculating several new data fields (e.g. low, middle, and high cloud fractions) that are not readily available in the raw MM5 output. It also interpolates temperature and wind speed to observation height (1.5m and 10m, respectively). The MCIP processor culls a minimum of six cells about the domain periphery to minimize edge effects in the MM5 simulation. MCIP can be used to further reduce the rows or columns in the MM5 data so that the domain definition for the MCIP output files precisely matches the domain used in the air quality modeling. MCIP also allows MM5 layers to be "collapsed" (i.e., some layers can be aggregated). When feasible it is desirable to use the same layer structure in the air quality model as in the MM5 to prevent errors associated with aggregating layer data and to maintain consistency between data produced by the meteorological model and those used by the chemistry-transport model. However, due to computational costs associated with using large number of vertical layers, vertical layer collapsing is typically used to reduce the total number of layers used by the CCTM. In the CENRAP annual modeling we will collapse from 34 layers in MM5 output into 19 layers for the CMAQ air quality simulations. The first 8 layers of CMAQ, up to approximately 450 m AGL, will match the MM5 vertical layer structure exactly. The region top for CMAQ is the same as used by MM5, 100mb (approximately 15 km AGL). The 36 km analysis domain contains 148 columns, 112 rows, and 19 layers. The definition of the horizontal extend of the 12 km domain is to be determined, but 19 layers will be used. More details on the CMAQ modeling domain definitions are provided in Chapter 4 with the vertical layer structure of MM5 and MCIP/CMAQ shown in Table 4-3.

### 5.1.2 Products of the CMAQ Meteorological Input Development Process

The meteorological input development process produces three two-dimensional and four three-dimensional daily meteorological and geophysical output data in the Models-3 I/O API format. These CCTM-ready meteorological input files are used in both emissions processing and the CCTM simulations. The met fields are 36 km and 12 km horizontal resolution on a Lambert Conformal Projection (LCP) coordinate system with 19 vertical sigma layers extending from the surface to the 100 mb pressure level. The data files include three-dimensional gridded fields of u- and v-wind components, vertical velocity, temperatures, Jacobian, Jacobian weighted air density, total air density, water vapor, cloud water content, rain water content, ice and snow mixing ratio, layer heights, and vertical exchange coefficients. Two-dimensional gridded fields of latitude and longitude, squared map-scaled factor, surface temperatures and pressures, 1.5 and 10 meter temperature, planetary boundary heights, rainfall, total cloud fraction, snow cover, deposition velocities, u* and w*, surface roughness length, as well as dominant land use category are also developed.

Table 5-1 shows the configuration to be used in MCIP version 2.2 for processing the 2002 MM5 output to produce CCTM-ready meteorology input files.

**Table 5-1.** MCIP V2.3 configuration used In the CENRAP annual modeling.

| Module or option | Values or setting | Additional Information |
|---|---|---|
| PBL value computation option | 1 | Use PBL value from input meteorology |
| Radiation fields | 1 | Use radiation fields from input meteorology |
| Dry deposition option | 2 | Use Models-3 (Pleim) dry deposition routine |
| Output interval | 60 | Unit is in minutes |
| Vertical layer structure | 19 layers | See Chapter 4 |

### 5.1.3 MM5 Reformatting Methodology

MM5 CAMx serves the same purpose as MCIP in the CAMx modeling system. MM5CAMx will be exercised using the same layer structure as MCIP. Two sets of vertical turbulent diffusivity files will be generated:

- Use of the O'Brien scheme (OB70).
- Use of the CMAQ scheme.

MM5CAMx will be operated initially with a 0.1 $m^2/s$ minimum $K_V$ (Kz_min) value, however the CAMx-ready $K_V$ files may be updated to a 1.0 $m^2/s$ Kz_min to be consistent with CMAQ.

### 5.1.4 Treatment of Minimum $K_V$

The minimum $K_V$ value (Kz_min) is an area of ongoing investigation by the CMAQ and CAMx developers. EPA initially recommended a 1.0 $m^2/s$ Kz_min for CMAQ modeling, but for their ozone forecasting EPA is using Kz_min values of 0.1 to 2.0 $m^2/s$ depending on the amount of urban land use present. Thus, to maximize flexibility we will process the 2002 MM5 data using MCIP and MM5CAMx using a 0.1 $m^2/s$ Kz_min and then impose other Kz_min values (e.g., 1.0 $m^2/s$) in the model-ready files.

## 5.2 Development of Emissions Model Inputs and Resultant Inventories

The base year emissions inventory for the CENRP annual modeling will be founded on revised 2002 emissions developed by CENRAP emission inventory (EI) contractors in IDA SMOKE format (Strait, Roe and Vukovich, 2004). These revised emissions were reviewed by CENRAP stakeholders and considered complete in late 2004. The emission inventory data are being reviewed and quality assured by the CENRAP EI Contractor and will be available by the end of November 2004.

Non-CENRAP state emissions are based on inventories obtained by the Study Team and determined to be representative of the 2002 episode year. Western State base year point and area source emissions for 2002 were provided by the WRAP RPO. Additionally, an inventory of point source resolved agricultural, prescribed, and wild fire emissions were provided by WRAP and will be utilized in the modeling. The South Eastern states emissions inventory of 2002 were provided by VISTAS RPO.

For the remaining U.S. portion of the domain, point source projections were based on EPA's 2001 modeling platform inventories and area source, nonroad mobile, and fire emissions were based on EPA's preliminary 2002 NEI. VMT and MOBILE input files were taken from EPA's preliminary 2002 FTP site and used in running MOBILE6 nationally.

Mexican and Canadian emissions are to be based on the latest available inventories obtainable by the Study Team in formats lending themselves to emissions modeling.

For purposes of air quality model validation, actual 2002 calendar year emissions for EGU and fire activity will be used, while during strategy and future year emission runs, "typical year" emissions for these categories will be processed.

A final 2002 CENRAP state emission inventory is expected by November 2004 and will be used in the final model performance demonstration and configuration expected to begin in early 2005. Non-CENRAP state emissions are expected to be based on State submittals to EPA under the CERR and will be augmented with additional data provided by RPO, State, and international sources. Air quality modeling will use actual 2002 calendar year emissions for EGU and fires, while "typical year" emissions for these categories will be processed during the strategy runs.

These emissions will then be converted to Inventory Data Analyzer (IDA) formatted versions and the data will be processed for air quality modeling using Version 2.0 (or Version

2.1) of the Sparse Matrix Operating Kernel Emissions (SMOKE) model. Included in these runs will be the temporal, spatial, and speciation profiles and cross-reference data currently provided with the 2.0 release of the model augmented with any recommended and approved emission profile data provided by the emissions inventory contractor or obtained from EPA prior to initial emissions modeling. The processing will be adjusted for each run to account for the specific air quality model (AQM) input required by CMAQ.

### 5.2.1 Emissions Modeling Methodology

Emissions inventory development for photochemical modeling must address several source categories including: (a) stationary point sources, (b) area sources, (c) on-road mobile sources, (d) non-road mobile sources, and (e) biogenic sources. For this analysis, these estimates must be developed to support the episode that is being modeled (i.e., the historical base year when the episode actually occurred; 2002).

Development of an emissions inventory customized for the CENRAP region requires a merging of: (a) the most recent *pertinent* regional inventory and (b) available high-resolution, locale-specific emissions estimated by local, state, and regional agencies in the CENRAP region. Local air regulatory and transportation planning agencies are generally the best sources of domain specific activity and control factors to use in developing the base year emissions. Often, these local emissions data sets come from a variety of sources, frequently in different formats.

Contacts with CENRAP's emission inventory contractors and the U.S. EPA will be established and formal requests made for inventory corrections, updates and ancillary data pertinent to the modeling of emissions in their jurisdictions. Where feasible and consistent with project resources and schedule, these updated data sets will be acquired and will be used to create day-specific modeling inventories specific to the CENRAP domain for the base year episodes to be modeled.

### 5.2.2 Set-up of SMOKE Over the CENRAP Domain

SMOKE will be configured to generate point, area, nonroad, highway, and biogenic source emissions. In addition, certain subcategories, such as fires and EGUs will be maintained in separate source category files in order to allow maximum flexibility in producing alternate strategies. Settings for each of the source categories are discussed in relevant sections below. With the exception of biogenic and highway mobile source emissions that are generated using the, respectively, BEIS and MOBILE6 modules in SMOKE, pre-computed annual emissions will be processed using the month, day, and hour specific temporal profiles of the SMOKE model.

To produce an emissions inventory to support annual modeling, representative time periods will be selected and modeled. Area, nonroad, and point sources will be modeled as a block of Thursday, Friday, Saturday, Sunday, Monday, one per month (total of 60 days modeled). On-Road motor vehicles will be represented by an entire single week for each month. These selection criteria allow for the representation of day-of-the-week variability in the on-road motor vehicles, and model a representation of the meteorological variability in each month.

Holidays will be modeled as Sundays. A list of modeled holidays is provided in Table 5-2. The biogenic emissions will be modeled on a day specific basis (365 days).

**Table 5-2.** SMOKE modeled holidays.

| Date | Julian Day | Holiday Description |
|---|---|---|
| January 1, 2002 | 2002001 | New Year's Day |
| March 29, 2002 | 2002089 | Good Friday |
| May 27, 2002 | 2002147 | Memorial Day |
| July 4, 2002 | 2002185 | July 4th |
| September 2, 2002 | 2002245 | Labor Day |
| November 28, 2002 | 2002332 | Thanksgiving Thurs |
| November 29, 2002 | 2002333 | Thanksgiving Fri |
| December 24, 2002 | 2002358 | Christmas Eve |
| December 25, 2002 | 2002359 | Christmas Day |

Population will be used as a gridding default for all source categories when the assigned surrogate would cause SMOKE to drop emissions. This can be a case when the county-level emission inventories are prepared using surrogates other than those available for modeling purposes.

The domain for the air quality modeling will be based on the EPA's 36-km national CMAQ domain, illustrated in Figure 5-1 below (details on the modeling domains are provided in Chapter 4).

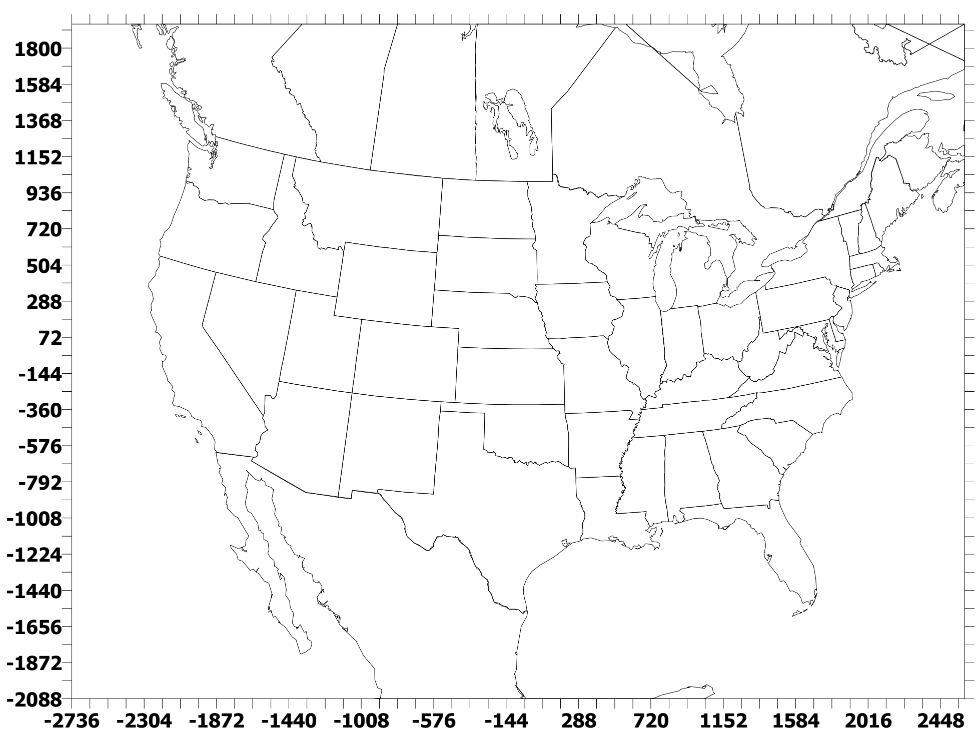


**Figure 5-1.** EPA 36-km National CMAQ domain.

The parameters for the SMOKE runs are as follows:

Episodes: 2002 Calendar Base Year.
Future Years: To be determined (most likely 2018).

Output Time Zone: Greenwich Mean Time (zone 0)

Projection: Lambert Conformal with Alpha=33, Beta=45, Gamma=-97, and center at (Lon -97, Lat 40).

Domain:
- 36 Kilometer Grid: Origin at (-2736, -2088) kilometers with 148 rows by 112 columns and 36-km square grid cells.
- 12 Kilometer Grid: To be determined.

Layer Structure: The CMAQ layer structure will include 19 layers, with specific layer positions defined in the meteorology files (see Chapter 4).

CMAQ Model Species: The CMAQ initial configuration will be for the CB-IV chemical mechanism with PM. The model species in the emission input files will be: CO, NO, NO2, ALD2, ETH, FORM, ISOP, NR, OLE, PAR, TERPB, TOL, XYL, NH3, SO2, SULF, PEC, PMFINE, PNO3, POA, PSO4, and PMC.

Meteorology Data: Daily (25-hour). SMOKE requires the following five types of MCIP outputs: (1) Grid cross 2-d, (2) Grid cross 3-d, (3) Met cross 2-d, (4) Met cross 3-d, and (5), Met dot 3-d. These files need to match the grid projection and overlap with the emissions modeling region but can be larger in the horizontal directions than the modeling region shown in Figure 5-1. Therefore, the data files for the 36 Kilometer grid domain will be at least 90 columns by 132 rows.

Elevated Sources: All sources will be treated by SMOKE as potentially elevated. No plume-in-grid sources will be modeled. Wildfire and some prescribed fire emissions will be handled as point sources.

Producing 365 day-specific input files for all source categories places a burden on available computing facilities, data management systems, and would adversely affect the Phase II schedule. Selecting representative model days for some or all of the source categories reduces the processing and file handling requirements to a more manageable level, and in most cases does not compromise the accuracy of the emissions files.

Other current or recent projects undertaken by the EPA, WRAP, VISTAS and LADCo have used a selection approach for all of the source categories (except biogenics) that use a representative weekday/Saturday/Sunday either for each month or each season to model all of the emissions files. In an attempt to better represent the level of temporal and spatial detail available for each source category, we have developed a more detailed strategy.

Biogenic emissions will be modeled for each episode day, using the daily meteorology. Point sources, including CEM and fire emissions will be modeled for each episode day to take advantage of the available day-specific emissions (if available) and meteorology. Area sources, including non-road mobile and dust emissions, with the exception of windblown dust emissions, do not utilize meteorological data, and are temporally allocated by monthly, daily and hourly profiles. Reviewing these profiles indicate that maximum temporal definition can be achieved by selecting representative Monday, weekday, Saturday, Sunday, for each month. Holidays will be treated as Sunday.

Motor vehicle emissions are influenced by meteorological variability, but the processing requirements for daily motor vehicle emissions were determined to be prohibitive under the current schedule. Instead of calculating the emissions factors as a function of the hourly meteorological data, a weekly average of these data, and hence weekly emission factors, will be calculated.

### 5.2.3 Development of Point Source Emissions

Stack parameters are often more important to the reliability of the air quality modeling results than the emissions rates themselves. Stack parameter data are frequently incorrect, especially in some of the current regional modeling inventories and careful QA is required to assure that the point source emissions are properly located both horizontally and vertically on the modeling grid. SMOKE has a number of built-in QA procedures designed to catch missing or out-of-range stack parameters. These procedures will be invoked in the processing of the point source data.

Depending on the emissions input files from CENRAP or its contractor, for the initial baseline modeling, we will be separating the point source emissions into EGU and non-EGU categories. The non-EGU category will not be using any day or hour-specific emissions. All non-EGU point source emissions will be temporally allocated to month, day, and hours using annual emissions and source category code (SCC) based allocation factors. These factors will be based on the cross-reference and profile data supplied with the SMOKE 2.0 version and will be supplemented with relevant data provided to the study team by CENRAP or its contractors.

For EGU sources with EPA reported CEM data, or with hourly emissions provided by stakeholders, actual hourly data will be used. For those sources where EPA CEM data are utilized, NOx, SO2, and heat input-based hour-specific profiles will be developed and applied to NOx, SO2, and all other emissions, respectively. This will ensure that the annual emission values provided by the EI contractor are maintained, but distributed using hourly to annual profiles. For sources providing hour-specific data and where they were approved by the State in which they operated, those data were substituted for EPA CEM-based emissions and distributions.

To temporally allocate the remaining EGU point sources, the NOx, SO2, and heat input data were collected from the 2002 Continuous Emissions Monitoring (CEM) datasets, and used to develop unit-level temporal distributions. The hour, day of week, and monthly specific temporal profiles will be used in conjunction with the EI supplied emissions data to calculate hourly EGU emissions by unit.

Off-shore point sources for the Gulf of Mexico region will be based on the latest emissions data developed by the Minerals Management Service (MMS). These data include emissions from all platforms in the Gulf of Mexico for calendar year 2000. The data will be processed and re-formatted for input to the SMOKE modeling system.

All point sources will be spatially allocated in the domain based on the stationary source geographic coordinates. If a point source is missing its latitude/longitude coordinates, the source will be placed in the center of its respective county.

### 5.2.4 Development of Area and Non-Road Source Emissions

All non-road mobile and area source emissions, except ammonia emissions (see below), will be temporally allocated to month, day of the week, and hours using annual emissions and source category code (SCC) based allocation factors. These factors will be based on the cross-reference and profile data supplied with the SMOKE 2.0 version and will be supplemented with relevant data provided to the study team by CENRAP or its contractors and other RPO's. Area and non-road sources will be spatially allocated in the domain based on SCC-based spatial allocation factor files. If an area or non-road source SCC does not have an existing cross-reference profile assigned to it, the county-level emissions will be allocated by population density in the respective county.

Off-shore area emission sources in the Gulf of Mexico region will be based on the latest emissions data developed by MMS. These data include emissions from all non-platforms activities in the Gulf of Mexico including support and survey vessels, helicopters, commercial marine vessels, and exploration and development activities. The data will be processed and re-formatted for input to the SMOKE modeling system.

If needed, a crustal PM transport factor will be applied to fugitive dust emission sources that have been identified in U.S. EPA modeling to have only a portion of its mass transportable from the source of the emission generation. The EPA's studies indicate that 60 to 90 percent of PM emissions from fugitive dust sources do not reach an elevated level necessary to be transported or modeled in an episodic simulation. For this reason, the CENRAP emissions contractor should apply county-specific fugitive dust emissions transport factors to these sources to adjust PM emissions accordingly. If the fugitive dust emissions adjustment is done by the CENRAP emissions contractor prior to receipt by the emissions and air quality Modeling Team then no additional adjustment will be made in the emissions modeling. (for adjustment factors see: http://www.epa.gov/ttn/chief/emch/invent/statusfugdustemissions_082203.pdf).

Ammonia Emissions

Ammonia emissions for CENRAP states have been developed by the EI contractor and include emissions from livestock, fertilizer application, natural soils, mobile sources, point sources and other miscellaneous sources. These estimate were developed primarily using the Carnegie Mellon University (CMU) NH3 model. County-level emissions estimates were developed along with associated temporal and spatial allocation profiles necessary for emissions processing with SMOKE. These data are currently being reviewed and quality assured by the CENRAP EI contractor and will be available by the end of November 2004. The study team will incorporate these emissions estimate for the CENRAP states in to the regional modeling inventories.

For all non-CENRAP states, the project team will apply the ammonia emission inventory modeling system recently developed for WRAP. The model treats all major sources of ammonia emissions (livestock, fertilizer application, natural soils, domestics sources and wild animals). The remaining ammonia emissions source categories are based on the latest 2002 inventories used for the WRAP. The WRAP ammonia model will be run using the latest MM5 meteorological data for all non-CENRAP states. The model generates hourly gridded emissions data using gridded meteorological data to apply various adjustments to emission factors and temporal allocation factors. Therefore, SMOKE is not required for the generation of these emissions estimates, although these emissions are processed through SMOKE in order to merge these source categories with the remaining area source emission estimates to obtain gridded model-ready data files.

### 5.2.5 Development of On-Road Mobile Source Emissions

The MOBILE6 module of SMOKE will be used to develop the base year on-road mobile source emissions estimates for CO, NOx, PM, and VOC emissions. The MOBILE6 parameters, vehicle fleet descriptions, and VMT estimates will be combined with gridded, episode-specific temperature data to calculate the gridded, temporalized emission estimates. Of note, whereas the on-network emissions estimates are spatially allocated based on link location and subsequently summed to the grid cell level, the off-network emissions estimates are spatially allocated based on a combination of the FHWA version 2.0 highway networks and population. For the CENRAP 36/12 km modeling, no link-based data will be used. The MOBILE6 emissions factors are based on episode-specific temperatures predicted by the meteorological model. Further, the MOBILE6 emissions factors model accounts for the following:

- Weekly average minimum/maximum temperatures;
- Facility speeds;
- Locale-specific inspection/maintenance (I/M) control programs, if any;
- Adjustments for running losses;
- Splitting of evaporative and exhaust emissions into separate source categories;
- VMT, fleet turnover, and changes in fuel composition and Reid vapor pressure (RVP).

The primary input to MOBILE6 is the MOBILE shell file. The MOBILE shell contains the various options (e.g. type of inspection and maintenance program in effect, type of oxygenated fuel program in effect, alternative vehicle mix profiles, RVP of in-use fuel, operating mode) that direct the calculation of the MOBILE6 emissions factors.

### 5.2.6 Development of Biogenic Source Emissions

A revised version of a commonly used biogenic emissions model, the Biogenic Emissions Inventory System (BEIS), has recently been developed and tested by EPA over two separate modeling domains/episodes. This version of the model (BEIS-3, v1.2) contains several changes over BEIS-2, including the following:

- Vegetation input data -- are now based on a 1-km Biogenic Emissions Landuse Database (BELD3) vegetation data base,
- Emission factors – many updates including some recent NARSTO modifications,
- Environmental algorithm -- includes a sunlit/shaded leaf solar radiation model.

For this particular application of BEIS-3, version 1.2 as currently incorporated in the SMOKE processor will be used.

The BELD-3 landuse data on a Lambert conformal grid at 1-km resolution have already been developed, are available, and will be used to estimate biogenic emissions in this study. The BEIS model also requires as input hourly, gridded temperature and solar radiation data to estimate biogenic emissions, and these data will be derived from the MM5 predictions.

### 5.2.7 Wildfires, Prescribed Burns, Agricultural Burns. Wind Blown Dust and Sea Salt Source Emissions

Agricultural Burns

Emissions from agricultural burns in CENRAP States have been developed by the CENRAP EI contractor as area sources with annual emissions. These emissions will be temporally and spatially allocated using the temporal profiles, and cross-reference files developed by the EI contractor. Spatial allocation surrogates, based on the MM5 modeling domain have been developed by the contractor at a 12km spatial resolution. The study team will reprocess these surrogates for the unified RPO domain used for air quality modeling at a spatial resolution of 36km.

Wildfires and Prescribed Burns

Wildfires and prescribed burn emissions will be handled separately from the standard area source input files. Wildfire emissions currently are not available for CENRAP states. WRAP and VISTAS have developed their own wild fire emissions as point sources. These emissions will be used for modeling on the unified RPO domain. An effort is being made to develop national wild file emissions. When these emissions are available, they will be used for CENRAP states.

Prescribed burn emissions for the CENRAP States have been received from the CENRAP EI contractor as both point and area sources. Burn emissions represented as area sources will be treated similar to agricultural burning emissions using spatial allocation surrogates. For the point source emissions, the study team will try to resolve the issue of temporal and vertical allocation of these files. The SMOKE 2.0 can model fire plume rise if provided with the following variables:

PTOP – Top of the fire plume profile (meters above ground level)
PBOT – Bottom of the fire plume profile (meters above ground level)
Lay1 – The percent of the emissions entrained in the first modeling layer

The WRAP Fire Emissions Joint Forum Emissions Inventory Report (FEJF, 2002) has documented an approach to calculating these plume descriptors. In this method, the fires are assigned to one of 5 size categories, based on the total burn acreage, and the biomass fuel loading. These categories are then used to calculate representative hourly plume profiles. These profiles are then used by SMOKE2.0 to distribute the vertical emissions from the fires. To successfully model fires as elevated point sources, the data provided by the EI contractor will need to include both the day, or days, on which the fire occurs, and a spatial identifier of the fire location. At a minimum, a latitude and longitude of the fire location can be used, although a polygonal coverage would be preferable.

In addition, wildfire and prescribed burn data, including emissions estimates and plume rise distributions, will be obtained from the WRAP contractor and used to supplement the inventory for the WRAP states.

Windblown Dust

The study team has developed wind blown dust emissions for the entire domain for $PM_{10}$ and $PM_{2.5}$. The dust emissions from agricultural lands have adjustment factors applied for the WRAP and CENRAP states only. The emissions for the rest of the states need to be adjusted for agricultural lands. The emissions model developed by the project team will be run using the MM5 meteorology from CENRAP. Emissions from wind erosion of natural geogenic sources (SCCs 2730100000 [total] and 2730100001 [dust devils]) will be excluded from the resulting modeling files using a 100 percent reduction in the control packets and replaced with the modeled estimates

Sea Salt

CMAQ currently treats sea salt as an inert PM species. That is, the sea salt is not allowed to chemically interact with other species, such as producing particulate sodium nitrate. There are plans to update CMAQ to have chemically active sea salt, but it is unclear whether such an update will occur during the CENRAP annual modeling. Accordingly, the initial modeling will be conducted without any sea salt emissions. If CMAQ is updated to treat chemically active Sea Salt, or if CAMx is run using its full-science options, then Sea Salt emissions will be generated using appropriate procedures.

### 5.2.8 Speciation and Reformatting of Emissions

SMOKE will be run to speciate the emissions estimates according to the requirements of the Carbon Bond Mechanism version four (CBM-IV, CB-IV or CB4). The SMOKE model will also reformat the emissions estimates for use in CMAQ modeling. For each model-ready emissions inventory, SMOKE will produce at a minimum five (5) separate air quality model-ready files: low-level point source, area source, elevated point source, mobile source, and biogenics. Other source categories, such as EGU and fire emissions may also be handled as separate air quality model-ready files.

### 5.2.9 Development of Modeling Inventories

The emissions inventories developed for the CENRAP modeling can be grouped into three distinct types: 1) 2002 annual inventories, 2) Future Year episodic inventories, and 3) Alternate Base Year Episodic Inventories. In all cases, the Study Team expects to receive the emissions inventory data for the CENRAP states from the Emissions Inventory contractor, add non-CENRAP states and Canadian and Mexican data acquired from alternate EPA and/or RPO sources, and produce the CMAQ ready emissions files.

#### 5.2.9.1 2002 Annual Inventories

The initial 2002 annual base inventory will be produced for the entire 2002 year. The initial inventory will consist of all of the required source categories: area, nonroad, point, motor vehicle and biogenics.

A revised 2002 annual inventory will be developed in the event that updated, or revised, emission inventory data area received from CENRAP. This may require re-production of all or some of the source categories. A "typical" inventory will be produced, based on the revised 2002, and replacing the revised actual emissions for EGUs and fires with "typical" EGU and fire emissions, or any other source emissions as provided by the EI contractor. The Study Team will also produce future year scenario annual inventories, based on the 2002 meteorology. The composition of these future year scenarios is yet to be determined.

#### 5.2.9.2 Future Year Episodic Inventories

Utilizing emissions data provided by the EI contractors, the Study Team will model alternate future-year sensitivity and control strategy scenarios. The specific episodes and emissions scenarios are as yet to be determined. This may require re-modeling of all, or some, of the source categories, with the exception of the biogenics.

### 5.2.10 Products of the Emissions Inventory Development Process

In addition to the CMAQ-ready input files generated for each hour of the days modeled in the 2002 base year annual run, a number of quality assurance (QA) files will be prepared and used to check for gross errors in the emissions inputs. Importing the model-ready emissions into PAVE and looking at both the spatial and temporal distribution of the emission provides insight into the quality and accuracy of the emissions inputs.

- Visualizing the model-ready emissions with the scale of the plots set to a very low value, we can determine whether there are areas omitted from the raw inventory or if emissions sources are erroneously located in water cells.

- Spot-check the holiday emissions files to confirm that they are temporally allocated like Sundays.

- Producing pie charts emission summaries that highlight the contribution of each emissions source component (e.g. nonroad mobile).

- Normalizing the emissions by population for each state will illustrate where the inventories may be deficient and provide a reality check of the inventories.

- Spot-check vertical allocation of point sources using PAVE.

We will use state inventory summaries prepared prior to the emissions processing to compare against SMOKE output report totals generated after each major step of the emissions generation process.

To check the chemical speciation of the emissions to CB-IV terms and the vertical allocation of the emissions, we will compare reports generated with SMOKE reports to target these specific areas of the processing. For speciation, we will compare the inventory import state totals versus the same state totals with the speciation matrix applied.

For checking the vertical allocation of the emissions, we will create reports by source, hour, and layer for randomly selected states in the domain. We will create these reports for a representative weekday in each of the episodes for each of these selected states.

The quantitative QA analyses often reveal significant deficiencies in the input data or the model setup. It may become necessary to tailor these procedures to track down the source of each major problem. As such, we can only outline the basic quantitative QA steps that we will perform in an attempt to reveal the underlying problems with the inventories or processing. Following are some of the reports that may be generated to review the processed emissions:

- State and county totals from inventory for each source category

- State and county totals after spatial allocation for each source category

- State and county totals by day after temporal allocation for each source category for representative days

- State and county totals by model species after chemical speciation for each source category

- State and county model-ready totals (after spatial allocation, temporal allocation, and chemical speciation) for each source category and for all source categories combined

- If elevated source selection is chosen by user, the report indicating which sources have been selected as elevated and plume-in-grid will be included

- Totals by source category code (SCC) from the inventory for area, mobile, and point sources

- Totals by state and SCC from the inventory for area, mobile, and point sources

- Totals by county and SCC from the inventory for area, mobile, and point sources
- Totals by SCC and spatial surrogates code for area and mobile sources
- Totals by speciation profile code for area, mobile, and point sources
- Totals by speciation profile code and SCC for area, mobile, and point sources
- Totals by monthly temporal profile code for area, mobile, and point sources
- Totals by monthly temporal profile code and SCC for area, mobile, and point sources
- Totals by weekly temporal profile code for area, mobile, and point sources
- Totals by weekly temporal profile code and SCC for area, mobile, and point sources
- Totals by diurnal temporal profile code for area, mobile, and point sources
- Totals by diurnal temporal profile code and SCC for area, mobile, and point sources
- PAVE plots of gridded inventory pollutants for all pollutants for area, mobile, and point sources

### 5.2.11 Preparation of CAMx Emission Inputs

Once the CAMx-ready 3D emission inputs are generated, the CMAQ-to-CAMx emissions converter would be used to generate the CAMx-ready emission inputs for 2002. This converter performs the following:

- Reads the 3D CMAQ-ready I/O AP2 emission inputs.
- Maps CMAQ species to CAMx species and performs unit conversions.
- Writes out a 2D surface-layer CAMx-ready low-level emission input file.
- For each (i,j,k) cell that has non-zero emissions, writes out a CAMx-ready elevated point source input file.

## 5.3 CMAQ Modeling Methodology

### 5.3.1 CMAQ Science Configuration

This section described the model configuration and science options to be used in the CENRAP annual modeling effort. The recommendations are based on testing and model evaluations of several models or model configurations carried out in BRAVO (Pitchford, 2004), CENRAP preliminary modeling (Pun, Chen and Seigneur, 2004; Tonnesen and Morris, 2004), VISTAS modeling (Morris et al., 2004), MRPO modeling (Baker, 2004) and WRAP modeling (Tonnesen, 2003). Table 5-3 summarizes the proposed configuration for CMAQ. The latest version of CMAQ is currently Version 4.4 that was released October 2004 and is currently proposed for use in the CENRAP annual modeling. However, if EPA releases an updated version of CMAQ in time, CENRAP would likely switch to the latest version at that time.

In the CMAQ base configuration we will run the 36 km for the 2002 annual period and run the 12 km grid for the selected episode(s) using one-way grid nesting where the boundary conditions for the 12 km grid simulation are extracted from the 36 km run using the CMAQ BCON processor. The base configuration of CMAQ will use 19 vertical layers up to a region top of 100 mb (approximately 15 km AGL) (see Table 4-3).

The PPM advection solver would be used along with the spatially varying (Smagorinsky) horizontal diffusion approach. K-theory will be used for vertical diffusion. We will initially run MCIP2.3 specifying a minimum eddy diffusion constant (Kz_min) of 0.1 $m^2/s$. Note that in the past the CMAQ default Kz_min value is 1.0 $m^2/s$. However, EPA has reported improved CMAQ performance using a variable Kz_min value in their ozone forecasting that ranges from 0.1 $m^2/s$ to 2.0 $m^2/s$ depending on the fraction of urban land use in the grid cell (0% to 100%). By processing the MM5 data with MCIP2.3 using a 0.1 $m^2/s$ Kz_min we can easily generate CMAQ Kz input files with a 1.0 $m^2/s$ Kz_min value.

The MCIP2.3 will be used to process the MM5 data using the "pass through" option.

The AERO3/ISORROPIA aerosol chemistry scheme will be used for inorganic aerosol thermodynamics.

The CB-IV gas-phase chemical mechanism is selected for the CENRAP annual modeling. VISTAS did extensive evaluation of CMAQ using the CB-IV and SAPRC chemical mechanisms and found that they produced similar PM model performance and major responses to emission reductions. Thus, given that CB-IV runs faster and uses less disk space that SAPRC and there are more uncertainties in the other components in the model and model inputs than in the gas-phase photochemistry, the more efficient CB-IV chemical mechanism was selected.

**Table 5-3.** Proposed CENRAP annual model configuration for the CMAQ.

| Model Option | CMAQ |
|---|---|
| Model Version | Version 4.4 (October 2004) |
| Horizontal Resolution | 36/12 km |
| No. Vertical Layers | NZ = 19 |
| Horizontal Advection | PPM |
| Vertical Advection | PPM |
| Horizontal Diffusion | Spatially Varying |
| Vertical Diffusion | $K_V$ (Eddy Diffusion) |
| MM5 Configuration | Pleim-Xiu/ACM |
| MM5 Processing | MCIP2.2 Pass Through |
| Gas-Phase Chemistry | CB4 |
| Gas-Phase Chemistry Solver | EBI |
| Secondary Organic Aerosol | SORGAM |
| Aqueous-Phase Chemistry | RADM |
| Aerosol Chemistry | AE3/ISORROPIA |
| Dry Deposition | Pleim-Xiu |
| Plume-in-Grid | Off |
| Initial Concentrations | CMAQ Default |
| Boundary Conditions | 3-Hourly 2002 GEOS-CHEM |
| Emissions | CENRAP States 2002<br>Other States 2002 from RPOs |

### 5.3.2 Spin-Up Initialization

For the 2002 annual CMAQ modeling, the model will be exercised separately for four quarters. The 2002 MM5 modeling started on December 16, 2001 at 12Z. Thus, allowing for 12 hours of spin up of the MM5 model, CMAQ will be initialized at 00Z on December 17, 2001. This results in a 14 day spin up period for CMAQ and the first quarter run segment of 2002. For the other quarter run segments of 2002, CMAQ will be initialized with a 15 day spin up period.

### 5.3.3 Boundary Conditions

Harvard University has been contracted by the RPOs to perform a 2002 GEOS-CHEM global climate model simulation. VISTAS has processed the 2002 GEOS-CHEM model output and generated day-specific 3-hourly boundary conditions (BCs) for the 36 km Inter-RPO grid in the CMAQ BCON format. The CENRAP 2002 annual modeling will use the 2002 GEOS-CHEM day-specific BCs.

### 5.3.4 Photolysis Rates

Several chemical reactions in the atmosphere are initiated by the photodissociation of various trace gases. To accurately represent the complex chemical transformations in the atmosphere, accurate estimates of these photodissociation rates must be made. The Models-3 CMAQ system includes the JPROC processor, which calculates a table of clear-sky photolysis rates (or J-values) for a specific date. JPROC uses default values for total aerosol loading and provides the option to use default column O3 data or to use TOMS data for total column O3. Previously, in the Phase I modeling we used TOMS data, and we will continue to use TOMS data for Phase II. There are a few days in 2002 for which TOMS data is not available, and we will used default column O3 data or previous days data for those days.

JPROC produces a "look-up" table provides the photolysis rates as a function of latitude, altitude, and time (in terms of the number of hours of deviation from local noon, or hour angle). In the current CMAQ implementation, the J-values are calculated for six latitudinal bands (10º, 20º, 30º, 40º, 50º, and 60º N), seven altitudes (0 km, 1 km, 2 km, 3 km, 4 km, 5 km, and 10 km), and hourly values up to "8 hours of deviation from local noon. During model calculations, photolysis rates for each model grid cell are estimated by first interpolating the clear-sky photolysis rates from the look-up table using the grid cell latitude, altitude, and hour angle, followed by applying a cloud correction factor.

The photolysis rates input file must be prepared as separate look-up tables for each simulation day. The modeling team has already prepared scripts to automate the production of photolysis rate files for each day of the annual simulation. Photolysis files are ASCII files, and these will be visually checked for selected days to verify that photolysis are within the expected ranges.

## 5.4 CAMx Modeling Methodology

This section described the model configuration and science options to be used in the CENRAP CAMx annual modeling effort. Table 5-4 summarizes the proposed configuration for CAMx. The latest version of CAMx is currently Version 4.10s that was released August 2004 and is currently proposed for use in the CENRAP annual modeling. However, a more appropriate updated version of CAMx is available, CENRAP would likely switch to the latest version.

### 5.4.1 CAMx Science Components

In the CAMx base configuration we will run the 36 km for the 2002 annual period and run the 12 km grid for the selected episode(s) using two-way grid nesting. The base configuration of CAMx will use 19 vertical layers up to a region top of 100 mb (approximately 15 km AGL) (see Table 4-3).

The PPM advection solver would be used along with the spatially varying (Smagorinsky) horizontal diffusion approach. K-theory will be used for vertical diffusion. Two sets of CAMx vertical diffusivity inputs will be generated using MM5CAMx: (1) one using the O'Brien scheme; and (2) the other using the Kv scheme in CMAQ. We will initially run MM5CAMx specifying a minimum eddy diffusion constant (Kz_min) of 0.1 $m^2/s$. As part of the CAMx modeling system there is a utility that produces enhanced minimum Kz (Kz_min) values near the surface to account for increased mixing due to roughness and the urban heat island. The selection of the Kz_min approach will be based on the latest thinking and sensitivity tests.

The CAMx 4.10s Mechanism 4 (M4) Course/Fine approach will be used for the CENRAP annual modeling that assumes all secondary PM is fine.

The CB-IV gas-phase chemical mechanism is selected for the CENRAP annual modeling. The RADM aqueous-phase chemistry and SOAP SOA module will also be used.

**Table 5-4.** Proposed CENRAP annual model configuration for the CAMx.

| Model Option | CMAQ |
|---|---|
| Model Version | Version 4.10s (August 2004) |
| Horizontal Resolution | 36/12 km |
| No. Vertical Layers | NZ = 19 |
| Horizontal Advection | PPM |
| Vertical Advection | PPM |
| Horizontal Diffusion | Spatially Varying |
| Vertical Diffusion | $K_V$ (OB70 and CMAQ) |
| MM5 Configuration | Pleim-Xiu/ACM |
| MM5 Processing | MM5CAMx |
| Gas-Phase Chemistry | CB4 |
| Gas-Phase Chemistry Solver | CMC |
| Secondary Organic Aerosol | SOAP |
| Aqueous-Phase Chemistry | RADM |
| Aerosol Chemistry | ISORROPIA |
| Dry Deposition | Wesley |
| Plume-in-Grid | Off |
| Initial Concentrations | CMAQ Default |
| Boundary Conditions | 3-Hourly 2002 GEOS-CHEM |
| Emissions | CENRAP States 2002<br>Other States 2002 from RPOs |

### 5.4.2 Spin-Up Initialization

For the 2002 annual CAMx modeling, the model will be exercised separately for four quarters using a ~15 day initialization period the same as CMAQ.

### 5.4.3 Boundary Conditions

CAMx boundary conditions would be day-specific 3-hourly and based on the 2002 GEOS-CHEM global climate model simulation. The CMAQ-to-CAMx BC processor would be used to processor the CMAQ BCON files for input into CAMx.

### 5.4.4 Photolysis Rates

The TUV photolysis rates processor would be used to generate the photolysis rates input file for CAMx. TOMS ozone data would be used to develop the CAMx Albedo/Haze/Ozone input file for 2002.

# 6.0 QUALITY ASSURANCE PLAN

In this section we discuss the quality assurance procedures that will be used in the CENRAP 2002 annual modeling. More details on the quality assurance and quality control (QA/QC) procedures to be used are described in the Quality Assurance Project Plan (QAPP; Morris and Tonnesen, 2004).

## 6.1 Quality Assurance Objectives

In December 2002, the USEPA publish extensive guidance on developing a Quality Assurance Project Plan (QAPP) for modeling studies (EPA, 2002). The objective of a QAPP is to ensure that a modeling study is scientifically sound, robust, and defensible. The new EPA guidance suggests that a QAPP should include the following elements :

- a systematic planning process including identification of assessments and related performance criteria;
- peer reviewed theory and equations;
- a carefully designed life-cycle development process that minimizes errors;
- clear documentation of assumptions, theory, and parameterization that is detailed enough so others can fully understand the model output;
- input data and parameters that are accurate and appropriate for the problem;
- output data that can be used to help inform decision making; and
- documentation of any changes from the original quality assurance plan.

Moreover, the EPA guidance specifies that different levels of QAPP may be required depending on the intended application of the model, with a modeling study designed for regulatory purposes requiring the highest level of quality assurance.

The QAPP also provides a valuable resource for project management. It can be used to document data sources and assumptions used in the modeling study, and it can be used to guide project personnel through the data processing and model application process to ensure that choices are consistent with the project objectives.

The guidance document also addresses model development, coding and selection of models, and model performance requirements. For the CENRAP 2002 annual modeling we are using existing EPA sponsored models (SMOKE and CMAQ) and a model developed by ENVIRON (CAMx) and have no current plans for model development activities. Thus, our QAPP focuses primarily on documenting data sources and QA of data processing performed by the model team. In addition, because no official EPA guidance currently exists for visibility model performance, a major objective of our QAPP will be to propose and define model performance evaluation procedures. QA objectives for specific aspects of the project are discussed below, and these will be incorporated into a QAPP that conforms to the EPA guidance document for modeling studies.

## 6.2 Emissions Model Inputs and Outputs

Emissions Quality Assurance (QA) and Quality Control (QC) are the single most critical step in performing air quality modeling studies. Because emissions processing is tedious, time consuming and involves complex manipulation of many different types of large data sets, errors are frequently made in emissions processing and, if rigorous QA measures are not in place, these errors may remain undetected.

As part of the CENRAP annual modeling effort, an "Emissions Gatekeeper" function will be implemented. The Study Team envisions the role of this Gatekeeper as one to perform quality assurance activities on the following emission inventory (EI) data:

(1) EI data obtained from the CENRAP emissions inventory contractors and other sources; and
(2) The emission inventory to be used for modeling outside of the States in the CENRAP region.

Specifically, the Emissions Gatekeeper will review the content and format of the provided emission inventories ensuring an appropriate appraisal of the emissions data and estimates for the CENRAP States. Other tasks will include any additional translation from mass emissions files into the emissions modeling input file structure necessary for modeling. The Study Team will supplement these activities with QA checks on the intermediate and model output files using internal and public domain visualization and diagnostic packages.

We propose a multi-step emissions QA/QC approach that involves several staff to QA/QC the emissions as they are processed. This includes the initial emissions QA/QC by the Emissions Gatekeeper described above, as well as QA/QC by the Emissions Modeler during the processing of emissions and then additional QA/QC by the air quality modeler of the processed model ready emission files. This multistep process with three separate groups involved in the QA/QC of the emissions is much more likely to catch any errors prior to the air quality model simulations.

### 6.2.1 Emissions Modeling QA/QC

<u>Input Screening Error Checking Algorithms</u>: Although the SMOKE emissions model will be used for emissions processing, some of additional input error checking algorithms will be used to screen the data and identify potential emission input errors. Additionally, EPA has issued a revised stack QA and augmentation procedures memorandum that will be used to identify and augment any outlying stacks.

<u>SMOKE error messages</u>: SMOKE provides various cautionary or warning messages during the emissions processing. We will redirect the SMOKE output to log files and review the log files for serious error messages. An archive of the log files will be maintained so that the error messages can be reviewed at a later date if necessary.

<u>SMOKE emissions summaries</u>: We will use QA functions built into the SMOKE processing system to provide summaries of processed emissions as daily totals according to species, source category and county and state boundaries. These summaries will then be compared with summary data prepared for the pre-processed emissions, e.g., state and county totals for emissions from the augmented emissions data.

### 6.2.2 QA of the Model-Ready Emissions Impacts

The goal of the post-processed emissions summary QA is to detect possible errors in the final, model-ready binary emissions files by preparing summary plots that characterize spatial and temporal patterns in the emissions data. This step is designed to catch errors that may be missed in the internal SMOKE QA procedures. We will use a QA/QC post-processing program that read the CMAQ-ready I/O API emissions file formats for each of the major source categories (mobile, area, point, biogenic, fire) and produce the following plots.

<u>Spatial Summary</u>: We will sum the emissions for all layers and for all 24 hours that is used to prepare a PAVE plot showing the daily total emissions spatial distribution. For a 20 day simulation this produces approximately 20 days x 20 species x 5 emissions categories = 2,000 plots. In our base case simulations these plots will be presented as tons per day. The objective of this step is to identify errors in spatial distribution of emissions.

<u>Vertical Profile</u>: For point sources the emissions total for each layer will be summed and plotted to show the vertical distribution of emissions. These plots show the emissions on the x-axis for each model layer on the y-axis. The objective of this step is to identify possible errors in vertical distribution of emissions.

<u>Short Term Temporal Summary</u>: The total domain emissions for each hour will be accumulated and time series plots prepared that display the diurnal variation in total hourly emissions. The objective of this step is to identify errors in temporal profiles.

<u>Long Term Temporal Summary</u>: The total domain emissions for each day will be accumulated and displayed as time series plots that show the daily total emissions across the domain as a function of time. The objective of this step is to identify particular days for which emissions appear to be inconsistent with other days for no reason (e.g., not a weekend) and compare against the general trend.

<u>Control Strategy Spatial Displays</u>: Spatial summary plots of the daily total emissions differences between a control strategy and base case emissions scenarios will be generated. These plots can be used to immediately identify a problem in a control strategy.

## 6.3 Meteorological Model Outputs

As part of the CENRAP annual modeling QA effort, a “Meteorological Gatekeeper” function will be implemented. The task of the Gatekeeper is to provide an independent review and quality assurance of the meteorological modeling and related data sets developed by

CENRAP and subsequently by the emissions and air quality modeling teams. This Gatekeeper QA review serves two specific purposes: (a) to ensure that any potential problems with the data sets (should they exist) are identified and corrected in a timely manner, and (b) to provide the study team with information to support ongoing CMAQ and CAMx model performance testing and sensitivity analyses. In the case of meteorology, the Gatekeeper's independent QA analysis of the MM5 meteorological data sets serves to provide direct assistance to the emissions and air quality modeling team as it undertakes to ratify the SMOKE model outputs and to diagnose the CMAQ and CAMx model performance and sensitivity analyses.

In addition to having personal responsibility for the quality and chain of custody of the meteorological data sets supplied by CENRAP, the Meteorological Gatekeeper will be responsible for ensuring and maintaining the integrity of the data files uploaded to the project website. This website, hosted by UCR, serves as the repository of data for the ENVIRON/UCR modeling centers and for the CENRAP Modeling Workgroup participants. In performing the Gatekeeper quality assurance activity, one of the first steps is to conduct an independent operational evaluation on the MM5 model results at 36 km and 12 km grid scale. This evaluation covers surface and aloft wind direction, temperature, mixing ratio, precipitation, and planetary boundary layer (PBL) depths on a continental scale (36 km) and subregional scale (12 km) basis. The Gatekeeper will also perform supplemental, ad hoc analysis of pertinent MM5 fields (e.g., PBL depths) where that might be useful to the emissions and air quality modeling teams. Another task of the Gatekeeper will be to exercise the Meteorological Chemistry Interface Processor (MCIP) version 2.3 and MM5CAMx processor to produce binary input files for the CMAQ and CAMx air quality models, respectively.

In summary, the quality assurance plan for the meteorological data will include the following elements:

- Upon receiving the MM5 output files from CENRAP, we will verify the integrity of the file transfer (e.g., no missing and/or corrupted files);

- We will process the 2002 MM5 data using the MCIP2.3 and MM5CAMx processors to generate 2002 model-ready meteorological inputs for CMAQ and CAMx, respectively.

- We will create horizontal and vertical plots of temperature, pressure, precipitation, modeled flow patterns, PBL heights, etc. to assess whether the MCIP output fields are reasonable;

- The CENRAP 2002 MM5 simulation will be evaluated using the same surface observations, subdomains and procedures as used to evaluate the VISTAS and WRAP 2002 MM5 simulation as an independent QA and evaluation of the database; and

- We will make slected plots available on the CENRAP website for viewing and download.

## 6.4 Air Quality Model Inputs and Outputs

Key aspects of QA for the CMAQ input and output data include the following:

- Verification that correct configuration and science options are used in compiling and running each model of the in the CMAQ modeling system, where these include the MCIP, JPROC, ICON, BCON and the CCTM.
- Verification that correct configuration and science options are used in compiling and running each model of the in the CAMx modeling system, where these include the MM5CAMx, TUV, CMAQ-to-CAMx IC, BC and emissions processors and other processors.
- Verification that correct input data sets are used when running each model.
- Evaluation of CMAQ and CAMx results to verify that model output is reasonable and consistent with general expectations.
- Processing of ambient monitoring data for use in the model performance evaluation.
- Evaluation of the CMAQ and CAMx results against concurrent observations.
- Backup and archiving of critical model input data.

The most critical element in the QA plan for CMAQ and CAMx simulations is the QA/QC of the meteorological and emissions input files. The major QA issue specifically associated with the air quality model simulations is verification that the correct science options were specified in the model itself and that the correct input files were used when running the model. For the CMAQ and CAMx modeling we employ a system of naming conventions using environment variables in the compile and run scripts that guarantee that correct inputs and science options are used. We also employ a redundant naming system so that the name of key science options or inputs are included in the name of CMAQ and CAMx executable program, in the name of the CMAQ and CAMx output files, and in the name of the directory in which the files are located. This is accomplished by using the environment variables in the scripts to specify the names and locations of key input files. For example, if a model simulation is performed using the CB4 mechanism, all compile and run scripts contain the variable definition "$MECH = CB4", and this variable is hard coded into the script for the executable name, the output file name, and the output directory name. This procedure produces long file/directory names but it effectively prevents mistakes or makes mistakes readily apparent if they do occur.

A second key QA procedure is to never "recycle" run scripts, i.e., we always preserve the original runs scripts and directory structure that were used in performing a model simulation. For example, if we perform simulation with the SAPRC mechanism, instead of editing the original scripts to specify "$MECH = SAPRC" we will create a parallel directory structure with a new set of scripts to perform the SAPRC simulations. This provides a permanent archive of the scripts that were used in performing model simulations. In addition, output from the model simulation will be directed to a log file that provides a record of input file names, warning messages etc that will be archived.

We will also perform a post-processing QA of the CMAQ and CAMx output files similar to that described for the emissions processing. We will generate animated gif files using PAVE that can be viewed to search for unexpected patterns in the CMAQ and CAMx output files. In the case of model sensitivity studies, the animated gifs will be prepared as difference plots for the sensitivity case minus the base case. Often, errors in the emissions inputs can be discovered by viewing the animated GIFs. Finally, we will produce 24 hour average plots for each day of the CMAQ simulations. This provides a summary that can be useful for more quickly comparing various model simulations.

# 7.0 MODEL PERFORMANCE EVALUATION

## 7.1 OVERVIEW

The CENRAP 2002 annual modeling work is broken down into six Tasks for the current fiscal Year (FY1) that ends April 30, 2005 that focuses on the development of an annual modeling database for 2002 with an attendant model performance evaluation, sensitivity analysis as well as modeling a future-years base case. And work in a second fiscal year (FY2) that ends April 30, 2006 that focuses on future-year control strategy evaluation and sensitivity analysis. The ultimate objective of the FY1 CENRAP modeling work effort is to obtain a properly evaluated operational emissions and air quality modeling system for the 2002 annual calendar year so that control strategy evaluation can begin by May 1, 2005. Consistent with the spirit of a modeling protocol for regulatory decision-making, this document lays out the 'roadmap' for achieving an adequately tested modeling system for regulatory usage. But, obviously, this does not mean that every analysis identified in this chapter will be carried out or is indeed even possible given the CENRAP annual modeling schedule and resources, the existing aerometric data bases, and present technology constraints. The roadmap guides the way to the desired destination – in this case, an evaluated, operational PM/regional haze modeling system – but does not commit the driver to exploring every side street and back country road along the way. Indeed, one expectation of the CENRAP Modeling Workgroup is a close working relationship with the Modeling Team to ensure that the available resources and schedule are applied most efficiently in reaching the aforementioned goal.

This chapter describes a range of model testing methodologies *potentially* available to the CENRAP Emissions and Air Quality Modeling study team in its efforts to adequately evaluate the performance of the CMAQ and CAMx air quality modeling systems for the 2002 annual period. Since one cannot know at this juncture the specific performance problems that may arise in the initial 2002 CMAQ and CMAQ base case simulations or the "final" year 2002 evaluation , we set forth in this chapter a broad range of methods and techniques that *may* be brought to bear in examining CMAQ and CAMx model performance. We identify the core operational evaluation procedures, recommended in EPA (2001) guidance that will be performed in the model performance evaluation. We also describe a broad range of additional performance testing methods that may be worth considering, if necessary. Implementation of one or several of these various techniques would have to be performed under separate funding. However, our base effort model performance evaluation is intended to provide a robust assessment of the operational ability of the CMAQ and CAMx models to predict fine particulate and visibility at Class I areas and other monitoring sites in the CENRAP region.

Clearly, not all of the supplemental evaluative techniques identified in this chapter will ultimately be performed. There are three reasons for this:

- The CENRAP Emissions and Air Quality Modeling available resources places clear limits on the resources available to perform model evaluation analyses. Accordingly, some evaluation steps, while desirable, simply may not be possible given current funding levels;

- The CENRAP timeline places stringent schedule demands on the model evaluation. A number of the model performance evaluation methods introduced in this chapter (e.g. Weight of Evidence analyses, diagnostic testing with individual measurement networks, PM indicator species and ratios analyses, use of aircraft data) could very likely require more time to carry out given their quasi-research nature. Since CENRAP annual modeling study is not a model research and development effort, but rather an operational evaluation of existing modeling systems for regulatory decision-making, some interesting, but time consuming analyses simply may not be possible given the present schedule; and

- To conform to the EPA PM guidance documents requirements for PM model testing, it may not be necessary to conduct many of the diagnostic and Weight of Evidence tests identified in this protocol. Indeed, an adequate evaluation of the CENRAP annual modeling system may be possible through straightforward application of the core operational performance evaluation procedures identified in EPA's 2001 draft guidance.

At a minimum, the evaluation of the CMAQ and CAMx modeling systems for the CENRAP annual 2002 simulations will be consistent with EPA's draft guidance on PM model testing. This guidance essentially calls for an operational evaluation of the model focusing on a specific set of gas phase and aerosol chemical species and a suite of statistical metrics for quantifying model response over the annual cycle. The emphasis is on assessing: (a) How accurately the model predicts observed concentrations? and, (b) How accurately does the model predict responses of predicted air quality to changes in inputs? States are encouraged to utilize the evaluation procedures set forth in the earlier 1991 guidance document (EPA, 1991) for gas phase species and the newer (2001) guidance of PM species. Thus, in carrying out the initial operational evaluation and the subsequent final evaluation, we will implement the suggested EPA performance testing methodologies for the key gas phase and aerosol species. Since these methods are explicitly presented in EPA's guidance document, there is no need to repeat them here.

Subject to the availability of time and currently unallocated resources, the CENRAP annual evaluation will also attempt to employ other testing methods beyond those in the EPA guidance document. However, the level of this effort without additional resources will depend on how smoothly the integration of other data (e.g., emissions and meteorological) are introduced into the Emissions and Air Quality Modeling Study. For example, if emissions are not in an adequate form usable for SMOKE emissions modeling, then current budget resources may have to be reallocated from model performance to fixing the emissions. This discussion is not intended to circumvent a full evaluation of the modeling systems, rather to recognize the very real resource limitations and if resources are diverted to other activities without additional funding, then work is dropped on the back end that usually includes limitations on the model performance evaluation.

We conclude by again emphasizing that most important goal of the CENRAP CMAQ and CAMx evaluation is to determine whether the aggregate modeling systems (model codes plus input data sets and observational data for testing) offers sufficiently reliable and accurate results that public decision-makers may have reasonable confidence in using the model to help choose between alternative regional haze reduction scenarios. If the CMAQ and CAMx model evaluation, as outlined in this chapter, provides sufficient evidence that the modeling systems are operating reliably and in conformance with measurements and scientific expectations, then specific justifications explaining why the model is acceptable for developing regional haze strategies will be offered in the

2002 CMAQ and CAMx Modeling Summary Report. Conversely, should the evaluation determine that the modeling systems suffers from important flaws or errors that undermine its reliability or use, these findings will also be documented, together with recommendations regarding the use of alternate methods, steps to improve the model and/or data base, or other approaches.

## 7.2 Context for the CENRAP Model Evaluation

We begin the discussion of the CENRAP annual modeling evaluation methodology by reviewing how the CMAQ and CAMx model output is used in regional haze applications to project changes in visibility (this issue is discussed in more detail in Chapter 8). When designing a model performance evaluation, it is important to understand how the modeling results will ultimately be used. EPA has published two versions of draft guidance for fine particulate and regional haze modeling (EPA, 2000; 2001), utilizing a Fine Particulate Guidance Workgroup to provide technical input in the development of both documents[1]. More recently, EPA has provided an informal update on the PM/regional haze modeling guidance (Timin, 2002) and conducted a PM model evaluation workshop (see, for example, Timin, 2004; Boylan, 2004) shedding additional light on what the final guidance document might contain.

A key concept in EPA's guidance for addressing regional haze issues is that the modeling results should be used in a relative sense to scale or roll back the observed particulate matter (PM) concentrations from which light extinction is estimated. Adopting the recommendations from the Federal Land Managers' Air Quality Related Values Workgroup (FLAG), estimates of visibility are obtained from observed fine particulate concentrations using the following IMPROVE reconstructed mass extinction equation (FLAG, 2000):

$$b_{ext} = 3\{f(RH)[(NH_4)_2SO_4]\} + 3\{f(RH)[NH_4NO_3]\} + 4\{f'(RH)[OC]\} + 10[EC] + 1[IP] + 0.6[CM] + b_{rayleigh}$$

where:

$b_{ext}$ is the estimated extinction coefficient ($Mm^{-1}$);
$[SO_4]$ is the sulfate concentration assumed to be ammonium sulfate;
$[NO_3]$ is the particulate nitrate concentration assumed to be ammonium nitrate;
[OC] is the organic carbon concentration;
[EC] is the elemental carbon concentration;
[IP] is the inorganic primary fine particulate (< 2.5 :) concentration excluding primary sulfates and nitrates;
[CM] is the coarse particulate (> 2.5 : and < 10 :) concentration;
$b_{rayleigh}$ is the light-scattering due to Rayleigh scattering (assumed to be 10 $Mm^{-1}$);
f(RH) is a relative humidity adjustment factor for the sulfate and nitrates; and
f'(RH) is a relative humidity adjustment factor for OC that is assumed to be 1.0.

[1] Members of the CENRAP modeling team participated on the EPA fine particulate modeling work group over the two-year span of its activities.

The relative humidity (RH) values used in the above equation are monthly- specific and Class I area-specific values based on a long term average (EPA, 2003a).

The regional haze rule expresses reasonable progress in terms of changes in the deciview ()dV) from the current to future year conditions. Deciview (dV) is expressed as the natural logarithm of the extinction coefficient ($b_{ext}$) to Rayleigh scattering:

$$\text{Deciview} = 10 \ln(b_{ext}/10)$$

$$)\text{dV} = 10 \ln[(b_{ext})_{future} / (b_{ext})_{base}]$$

In this framework, changes in visibility in terms of percentage changes in extinction over a natural "clean" background visibility rather than changes in deciview. Changes in deciview or extinction are essentially mathematically identical with a 1 change in deciview being approximately equivalent to a 10% change in extinction. Both visibility parameters will be calculated in the CENRAP annual model applications.

The CENRAP 2002 CMAQ and CAMx model testing will concentrate on an operational evaluation of those model predictions that are most necessary for estimating visibility (e.g., $SO_{4,}$ $NO_{3,}$ OC, EC, IP and CM). Where feasible and supported by sufficient measurement data, we will also evaluate the modeling system for its ability to accurately estimate gas-phase oxidant and precursor/product species since correct, unbiased simulation of gas-phase photochemistry is a necessary element of reliable regional haze predictions. This evaluation will be carried out across the full CENRAP domain for the entire year and also on other subdomains (e.g., CENRAP-N, CENRAP-S, WRAP, MRPO, VISTAS and MANE-VU) and month-by-month basis to help build confidence that the modeling system is operating correctly. With this context in mind, we next turn to the philosophy of the model evaluation process.

## 7.3 Multi-Layered Model Testing Process

EPA's "Draft Guidance for Demonstrating Attainment of Air Quality Goals for $PM_{2.5}$ and Regional Haze" (EPA, 2001) affirms the recommendations of numerous modeling scientists over the past decade (see, for example, Dennis et al., 1990; Tesche et al., 1990, 1994; Seigneur et al., 1998, 2000; Russell and Dennis, 2000; Arnold et al., 2003; Boylan et al., 2003; Tonnesen, 2003) that a comprehensive, multi-layered approach to model performance testing should be performed, consisting of the four components: operational, diagnostic, mechanistic (or scientific) and probabilistic. As applied to regional PM/visibility models, this multi-layered framework may be viewed conceptually as follows:

> **<u>Operational Evaluation</u>**: Tests the ability of the model to estimate PM concentrations (both fine and coarse) and the components at $PM_{10}$ and $PM_{2.5}$ including the quantities used to characterize visibility (i.e., sulfate, nitrate, ammonium, organic carbon, elemental carbon, $PM_{2.5}$, and $PM_{10}$). This evaluation examines whether the measurements are properly represented by the model predictions but does not necessarily ensure that the model is getting "the right answer for the right reason";

> **<u>Diagnostic Evaluation:</u>** Tests the ability of the model to predict visibility and extinction, PM chemical composition including PM precursors (e.g., SOx, NOx, and $NH_3$) and associated oxidants (e.g., ozone and nitric acid); PM size distribution; temporal variation; spatial variation; mass fluxes; and components of light extinction (i.e., scattering and absorption);

> **<u>Mechanistic Evaluation</u>:** Tests the ability of the model to predict the response of PM and visibility to changes in variables such as emissions and meteorology; and

> **<u>Probabilistic Evaluation:</u>** Takes into account the uncertainties associated with the model predictions and observations of PM and visibility.

Within the constraints of the CENRAP annual modeling schedule and budget resources, the CENRAP Emissions and Air Quality Modeling effort will attempt to include elements of each of these components. The operational evaluation will obviously receive the greatest attention since this is the primarily thrust of EPA's 2001 PM guidance. However, we will consider, where feasible and appropriate, diagnostic and mechanistic tests (e.g., use of probing tools, indicator species and ratios, aloft model evaluations, urban vs. rural performance analyses), and traditional sensitivity simulations to explore uncertainty. The scope of these additional diagnostic and mechanistic tests will be shaped by additional resources and the timing of when such analyses are commissioned relative to the CENRAP schedule.

Before discussing the types of testing procedures available for the above evaluation components, we first identify the surface and aloft data sets that are available to support these comparisons.

## 7.4 Development of Consistent Evaluation Data Sets

### 7.4.1 Surface Measurements

The ground-level model evaluation database will be developed using several routine and research-grade databases. The first is the routine gas-phase concentration measurements for ozone, NO, $NO_2$ and CO archived in EPA's Aerometric Information Retrieval System (AIRS/AQS) database. Other sources of information come from the various PM monitoring networks in the U.S., with particular emphasis in the central U.S. These include the: (a) Interagency Monitoring of Protected Visual Environments (IMPROVE), (b) Clean Air Status and Trends Network (CASTNET), (c) Southeastern Aerosol Research and Characterization (SEARCH), (d) EPA $PM_{2.5}$ and $PM_{10}$ Mass Networks (EPA-FRM), (e) EPA Speciation Trends Network (STN); (f) National Acid Deposition Network (NADP) and (g) EPA Supersites (EPA-SPEC) networks. Typically, these networks provide ozone, other gas phase precursors and product species, PM, and visibility measurements.

As an example, the IMPROVE network gives daily (24-hour) average mass concentrations every 3 days for $SO_4$, $NO_3$, organic carbon (OC), elemental carbon (EC), soil (IP), CM, $PM_{2.5}$ and $PM_{10}$. These data are available at approximately 11 sites in the CENRAP states. In addition, hourly values of light extinction and deciview are available at several of these sites. The SEARCH network provides 24-hour as well as continuous (hourly) speciated measurements of $PM_{2.5}$ components and other specifics from 8 stations all located in the southeastern US outside of the CENRAP states (Hansen et al., 2003). We will use data from these and the other observational databases listed in Table 7-1, supplemented with the routine AIRS/AQS data, as appropriate, for CMAQ and CAMx model performance testing.

Care must be taken in selecting data that are representative of regional concentrations and the prediction of regional haze at the sensitive Class I areas within the CENRAP states. For example, the criteria and other pollutants in the AIRS/AQS database are typically urban-oriented. Thus, would not expect CMAQ or CAMx predictions for these species using a 36 km (or even 12 km) grid mesh to be ideally suited to simulating urban-scale concentrations. With other data sets (e.g., SEARCH) it may be possible to reproduce the urban-to-rural gradients in $PM_{2.5}$ components. While finer grid meshes (e.g., 4 km) might be more appropriate, high-resolution modeling at 4 km scale on an annual time frame for the CENRAP region is prohibitive given current computer technology and model run times. Accordingly, some selectivity is needed in assembling pertinent measurement sites for comparisons with regional-scale model predictions and the interpretation of the modeled and measured comparisons must be made with care.

Another important consideration is that different PM monitoring networks may use different measurement approaches that "measure" different amounts of the same species that are also different from the modeled species. For example, the IMPROVE network only speciates $PM_{2.5}$ so any sulfate or nitrate in the coarse mode ($PM_{2.5\text{-}10}$) is included in the CM species. The CMAQ and CAMx models will be evaluated separately for each network. While the CENRAP annual modeling study is clearly not aimed at new model development or algorithm refinement, information discerned from model performance across the various networks should be useful in later model refinement activities and network design improvements. Finally, the mapping of the modeled species to the monitored data will also have to be performed in a consistent fashion.

### 7.4.2 Aloft Measurements

In recent years, the use of instrument aircraft in support of regulatory monitoring and research programs has become much more commonplace. Indeed, in the upper Midwest, the Lake Michigan Air Directors Consortium (LADCo) has been centrally involved in aircraft programs to support model development and applications studies for seventeen (17) years, beginning with pioneering flights in 1987. Supplementing the long-term sampling performed by LADCo in the Midwest, there have been other occasional intensive airborne sampling campaigns throughout the eastern U.S. (e.g., the 1999 SOS field program which provided aloft data for our evaluation of CMAQ for the July '99 episode), that have produced very useful information for air quality model performance testing. Fortunately, during CY-2002, there were at least two mature airborne field programs underway in the eastern U.S. One was centered over the Midwest, the other on the mid-Atlantic coast. A brief characterization of these potentially valuable CMAQ and CAMx model evaluation data sets is given here.

During 2002, the Wisconsin Department of Natural Resources (WDNR) and the Midwest RPO (MRPO) (who funded the Jacko aircraft) collaborated on the support of airborne sampling using two aircraft that, along with ground-based measurements, provided a 3-dimensional representation of air pollution concentrations across the upper Midwest with some flight paths extending south to include the Mammoth Cave, KY and Dolly Sods, WV Class I areas. The goal of the WDNR/MRPO flights was to collect aloft air quality and meteorological data to support model evaluation and data analyses. The aircraft flights were aimed at: (1) characterizing high fine particle and ozone episodes, (2) characterizing air quality over the Class I areas in the upper Midwest (Isle Royale National Park and Seney National Wildlife Refuge in northern Michigan) on both clean and hazy days, and (3) characterizing urban areas in the Midwest.

As indicated in Table 7-2, airborne sampling was performed over a broad region of the Midwest from 1 June to 22 November. Lasting 3-5 hours, the WDNR and Jacko aircraft sampled a variety of aerometric parameters (depending upon the flight and aircraft) including wind speed, wind direction temperature, dew point, relative humidity, pressure, $O_3$, NO, $NO_2$, $NO_x$, $NO_y$, speciated VOCs, carbonyls, $HNO_3$, $NH_3$, Hg, $SO_4$, OC, EC, $PM_{2.5}$, and light scattering (Neph). Still photographs documenting visibility were also collected. Presently, the full WDNR/MRPO aircraft database, from the first flights in 1987 to the recent sampling in 2003 is being aggregated into a master data base archive.

At the University of Maryland, researchers have been using ground-based monitors, radiosondes, profilers, and instrumented aircraft to make observations each year since 1992. Parameters measured included meteorology; selected trace gases; fine particulate chemistry, microphysics and optical properties across broad regions of the middle Atlantic coast. During 2002, the University Research Foundation's Aztec-F aircraft instrument suite included $O_3$, NO, CO, $SO_2$ samplers, as well as a $NO_2$ closed-path tunable diode laser system, and a differential GPS-based meteorology (T, RH) and horizontal wind (*u* and *v* horizontal components) data system. Aztec-F flights were made from 23 May to 3 October, typically lasting 3 hours.

The volume of aircraft information available for CMAQ/CAMx performance testing during 2002 is quite significant. Historically, aircraft data sets have been used only sporadically in evaluating model performance aloft. Based on current resource and schedule constraints, we do not plan on evaluating the CENRAP CMAQ/CAMx 2002 runs using aircraft data.

## 7.5 Model Evaluation Tools

This section introduces the various statistical measures, graphical tools, and related analytical procedures that have proven useful over the years in evaluating grid-based chemical transport models. Many of the methodologies mentioned below are being utilized to one degree or another in WRAP and VISTAS. Where appropriate, they will also be used in the CENRAP 2002 annual 36 km and episodic 12 km evaluation of the CMAQ and CAMx modeling systems. However, while we plan on calculating a rich variety of statistical performance metrics, only a very limited subset of these measures will actually be relied upon to form judgments concerning model acceptability and in the final reporting.

The current methodologies and statistical metrics may be augmented, as necessary and appropriate, by new measures that may become available during the course of the CENRAP modeling. The CENRAP evaluation will employ similar methods and evaluation tools used in WRAP, VISTAS and MRPO only focused on the CENRAP states and vicinity, but the evaluation will remain open to emergent methods for testing model performance where necessary and consistent with the project schedule and resources.

### 7.5.1 Statistical Performance Metrics

EPA's 2001 PM and regional haze guidance suggests a suite of metrics for use in evaluating model performance. The standard set of statistical performance measures suggested by EPA for evaluating fine particulate models includes: (a) normalized bias; (b) normalized gross (unsigned) error; (c) fractional bias; (d) fractional gross error; and (e) fractional bias in standard deviations. These measures are subsumed within the list of metrics that are calculated on a routine basis using the UCR and CENRAP model evaluation tools (these are identified in Table 7-3). The UCR evaluation software will generate these statistical measures for each model simulation performed for each analysis region (see below). In parallel, the CENRAP Evaluation Tool will be used to generate complimentary statistical measures. From past regional PM model evaluations we have found the fractional bias and fractional error to be the most useful summary measures and we will focus mainly upon them in the CENRAP modeling, but not to the exclusion of others that are found to yield discriminating power. For ozone and other gas phase species (NO, $NO_2$, $SO_2$) we will include use the traditional statistical measures (EPA, 1991, 1999).

Typically, the statistical metrics are calculated at each monitoring site across the full computational domain for all simulation days. In the CENRAP CMAQ/CAMx evaluation, we will stratify the performance statistics across relevant space and time scales. As part of the operational evaluation, the gas-phase and aerosol statistical measures shown in Table 7-3 will be computed for the full 36 km and 12 km domains, as well as for the individual RPOs (including CENRAP of course) and on other subdomains as appropriate. Temporally, we will compute the statistical measures for the appropriate averaging times: 1 hr for ozone, and gas-phase precursors such as NO, $NO_2$, CO, $SO_2$; 8-hr for ozone, and 24 hr for sulfate, nitrate, PM and other aerosol species. These results will then be averaged over annual, monthly, and seasonal periods for display, further analysis, and reporting. Should it become necessary as part of model performance diagnosis, we will consider aggregating the statistics in other ways, e.g., (a) day vs. night, (b) weekday vs. weekend, (c) precipitation vs. non-precipitation days, (d) month of the year, and (e) the 20% haziest/cleanest days, in order to help elucidate model performance problems. Absent performance difficulties, these supplemental time/space analyses would only be considered if additional resources are made available. In subregional performance testing, the focus would likely be on the Class I areas and sites where enhanced monitoring (e.g., St. Louis Supersite) is available.

As part of the operational evaluation, the metrics defined in Table 7-3 will be calculated for each gas phase species and each fine particulate species in the extinction equation as well as separately for $SO_4$, $NO_3$ and ammonium ($NH_4$) on both the 36 km and 12 km domains. In any diagnostic evaluations that are performed, we will examine the model's ability to estimate the gaseous species listed above from EPA's guidance (EPA, 2001). However, in reality ambient gaseous species in 2002 are principally available for ozone, $NO_2$, $SO_2$, and CO. Since most of the gaseous air quality monitors are located in urban settings, these data sets will be of somewhat restricted value in evaluating the 36/12 km regional-scale CMAQ predictions. We would consider the merits of assembling a rural-oriented gaseous species observational model performance evaluation database for use with CMAQ/CAMx. This database could be composited with measurement from rural sites from the AQS, CASTNet, NPS, STN, and SEARCH networks. If additional resources are made available, this rural vs. urban scale operational evaluation would be conducted to augment the more traditional operational evaluation using all valid data sets.

### 7.5.2 Graphical Representations

The CENRAP annual modeling CMAQ and CAMx operational air quality model evaluation will utilize numerous graphical displays to facilitate quantitative and qualitative comparisons between CMAQ/CAMx predictions and measurements. Together with the statistical metrics listed in Table 7-3, the graphical procedures are intended to help: (a) identify obviously flawed model simulations, (b) guide the implementation of performance improvements in the 2002 model input files in a logical, defensible manner, and (c) to help elucidate the similarities and differences between the alternative CMAQ/CAMx simulations. These graphical tools are intended to depict the model's ability to predict the observed fine particulate and gaseous species concentrations.

The core graphical displays to be considered for use in the CENRAP annual modeling include the following:

- Scatter plots of predicted and observed concentrations;
- Time series plots at monitoring locations;
- Spatial maps of ground-level gas-phase and particulate concentration maps (i.e., tile plots);
- Bias and error stratified by concentration (Bugle Plots);
- Bias and error stratified by time (e.g., Soccer Plots); and
- Separate displays of above by monitoring network, subregions and time.

These graphical displays will be generated, were appropriate for the full annual cycle as well as for monthly and seasonal periods. The displays will be generated with a consistent suite of products including the UCR analysis tools, CENRAP evaluation tool and ENVIRON evaluation software.

### 7.5.3 Probing Tools and Allied Methods

The CMAQ/CAMx operational model evaluation in the current fiscal year (FY1 through April 30, 2005) will employ routine operational evaluation methods and standard statistical metrics (Table 7-3) and graphical displays to support the assessment of whether the models are shown to perform with sufficient accuracy and reliably for its intended purpose. Ideally, this operational evaluation will confirm that the modeling systems are performing consistent with its scientific formulation, technical implementation, and at a level that is at least as reliable as other current state-of-science methods. Should unforeseen model performance problems arise in the 2002 Base Case model simulations, it may be necessary to draw into the evaluation supplemental diagnostic tools to aid in model testing. These diagnostic techniques are loosely referred to as "probing tools". The actual need for their use, if any, can only be determined once the initial 2002 CMAQ/CAMx operational evaluation is completed. Should such diagnostic methods actually be needed, their usage will require additional resources. Below, we identify the types of probing tools that could be brought to bear to enhance the currently planned CENRAP operational evaluation of the CMAQ and CAMx models.

Current 'One-Atmosphere" models, such as CMAQ and CAMx, have been outfitted with a number of "probing tools" that have proven to be very useful in testing and improving model performance and in evaluating emissions control strategies. Among the probing tools available in one or both models are: (a) ozone source apportionment technology (OSAT) and PM source apportionment technology (PSAT) algorithms, (b) process analysis (PA), and (c) the decoupled direct method (DDM) sensitivity analysis. CENRAP may choose to evaluate these tools as part of their FY2 (May 1, 2005 through April 30, 2006) modeling exercise, or with the allocation of further resources under the current FY1 period.

**Source Apportionment Technology:** CAMx contains a suite of "source attribution" methods. One such method is Ozone Source Apportionment Technology (OSAT). OSAT tracks ozone formation based on how groups of ozone precursors contributed to ozone formation. Thus, OSAT decides whether ozone formation is NOx or VOC limited in each grid cell at each time step, and bases ozone attributions on the relative amounts of the limiting precursor from different sources that are present in that grid cell at that time step. These incremental ozone attributions are integrated throughout the model run. The method is generally applicable and has been widely used to aid model diagnosis in the performance testing phase and to guide control strategy formulations as well.

A new PM Source Apportionment Technology (PSAT) has been implemented in CAMx funded by the MRPO that has been fully tested and evaluated. A Tagged Species Source Apportionment (TSSA) approach has also been implemented in CMAQ and is undergoing further testing.

**Decoupled Direct Method (DDM):** Various forms of the Decoupled Direct Method (DDM) have been installed in CMAQ and CAMx, based on the original work of Dunker and co-workers (Dunker, 1981; 1984; Dunker et al., 2002) and researchers at Georgia Institute of Technology (GIT). In general, the DDM method: (a) calculates first order sensitivities dC/dP where C is a concentration output and P an input parameter[2], (b) promotes accuracy by using consistent numerical methods and the same time steps for concentrations and sensitivities, (c) optimizes the code for efficiency, but not at expense of accuracy, and (d) calculates sensitivities with respect to parameters representing pollutant sources – emissions, BCs and ICs. Finally, the DDM provides a flexible and powerful user interface for defining various sensitivities including:

- Emissions resolved by geographic area.
- Emissions resolved by source category.
- BCs optionally resolved by boundary edge (N, S, E, W, Top).
- All sensitivities available relative to sources of individual species (NO, PAR, etc.) or species group (VOC, NOx or ALL).
- Simultaneously calculate sensitivities to many initial condition, boundary condition and emissions parameters.

In recent comparisons between CAMx DDM sensitivities and brute-force sensitivities (calculated from +/- 20% perturbations) Dunker et al., (2002a,b) reported that sensitivities of ozone with respect to area source NOx and VOC emissions were calculated and results indicated that the

[2] Recent research by Prof. Russell and coworkers at GIT has led to the extension of the CMAQ DDM method to include second order sensitivity coefficients (see, Hakami et al., 2003).

agreement between DDM and brute force sensitivities is excellent. DDM implementation into CMAQ is reported by Kumar (2003).

**Process Analysis (PA):** Photochemical air quality model simulations are usually evaluated primarily in terms of their ability to simulate observed $O_3$ data. There is an increasing awareness that chemical mechanisms, and air quality models must also be evaluated in terms of their ability to simulate the fundamental chemical processes that control $O_3$ formation and the sensitivity of $O_3$ to emissions reductions (Arnold et al., 1998). Process analysis is a method for explaining model simulations by adding algorithms to the AQM to store the integrated rates of species changes due to individual chemical reactions and other sink and source processes (Jeffries and Tonnesen, 1994; Tonnesen, 1995). By integrating these rates over time and outputting them at hourly intervals, process analysis provides diagnostic outputs that can be used to explain a model simulation in terms of the budgets of free radicals, production and loss of odd oxygen and $O_3$, and conversion of NOx to inert forms, as well as the effects of transport and other sink and source terms. Of particular importance to the CENRAP modeling, process analysis can also improve model diagnosis and performance evaluation efforts by identifying processes that are 'out of balance' (Tesche and Jeffries, 2002), by identifying situations for which the model formulation and/or implementation should not be expected to apply and by suggesting how ambient data can be used to evaluate model accuracy for key terms in the chemical processing of VOC and NOx (e.g., Imre et al., 1998).

Process Analysis (PA) is implemented in both CMAQ and CAMx and each model supports three complementary aspects of the method: (a) the integrated process rate (IPR), (b) integrated reaction rate (IRR) and (c) chemical process analysis (CPA). Several versions of process analysis (PA) have been implemented in air quality models (AQMs) including both trajectory models (Tonnesen, 1990, 1995) and grid models (Jang et al., 1995, Tonnesen and Dennis, 2000; Arnold et al., 1998; and Wang, 1997).

The fundamental approach in all versions of PA is similar: The AQM is modified to calculate the integral over time of the individual sink and source processes and each chemical reaction. These integrated sink/source process rates (IPR) and integrated reaction rates (IRR) can then be stored to a file and analyzed using a post-processor, or some processing can be performed internally in the model and a more limited set of process diagnostic information is output directly by the AQM. Chemical process analysis (CPA) is an improvement on the IRR method whereby some of the processing of IRR information is internalized within the AQM to output chemically meaningful parameters directly (e.g., budget terms for $O_3$, NOx and odd oxygen).

Process analysis measures for aerosol chemistry have not been analyzed as much as for ozone chemistry. Although the ozone chemistry process analysis is directly related to secondary sulfate and nitrate formation, there is additional process analysis information available in the aerosol modules that are not extracted in either CMAQ or CAMx. In particular, information on sulfate formation and oxidants from the aqueous-phase module and on the sulfate/nitrate equilibrium from the aerosol thermodynamics module would be a useful addition to the current process analysis output.

Because application of all three of these probing tools--source apportionment, DDM, and Process Analysis—are computational intensive and require a fair amount of analysis time to reap the benefits of using the methods, they do not lend themselves directly to annual simulations. However, each method has potential for use in addressing key episodic periods or geographical locations in the CENRAP domain where performance in the 2002 simulation may present a problem or where particular attention needs to be focused on emissions controls (a specific Class 1 area for example).

In such focused applications, one or more of these probing tools may indeed serve a purpose and will be considered where appropriate.

## 7.6 CENRAP 2002 Annual Model Evaluation Procedures

EPA guidance (EPA, 2001, pg. 227) suggests that the performance evaluation focus on two aspects:

- How well is the model able to replicate observed concentrations of components of $PM_{2.5}$, and total observed mass of $PM_{2.5}$? and
- How accurately does the model characterize the sensitivity of changes in component concentrations to changes in emissions?

Recognizing that the former is much easier to accomplish than the latter, EPA goes on to declare that testing of a model's reliability in estimating the actual effects of emissions changes is the more important. Over the past 20 years, a substantial body of information and analytical techniques has been developed to address the first aspect. Unfortunately, even today there are little rigorous methods available for quantifying the accuracy and precision of a model's predictions of ozone, PM or visibility changes as the result of emissions changes. In this section we explain how the CENRAP annual model testing will address the first aspect of the performance evaluation, i.e., how does the model compare against observed data. In section 7.10 we consider the second performance consideration.

### 7.6.1 Assessment of Ground-Level Gas-Phase and Aerosol Species

Given that visibility is expressed in terms of extinction and deciview built off of individual components of fine particulate matter, the model should be evaluated separately for each of the key fine particulate matter components that make up the extinction coefficient. Current EPA guidance suggests that the model should also be evaluated for ammonium as well as several key gas-phase species that are important for fine particulate modeling. For *particulate species* this includes $SO_4$ and/or S, $NH_4$, $NO_3$, mass associated with $SO_4$, mass associated with $NO_3$, elemental carbon (EC), organic carbon (OC), IP, mass of individual constituents of IP, and coarse matter (CM). The *gaseous species* include ozone ($O_3$), $HNO_3$, $NO_2$, PAN, $NH_3$, NOy, $SO_2$, CO, and $H_2O_2$.

As indicated above, for some of the CENRAP subregions there are very few measurements, if any, for most of the gaseous species. But ozone measurements are available and should be integrated into the model performance evaluation. Given the importance of the photochemistry and the radical cycle in forming secondary fine particulate matter, the ozone evaluation will provide some insight into the model's ability in this area. However, given the coarse grid resolution to be used (36 km and 12 km) perhaps greater weight should be given to the model evaluation results at the more rural ozone monitors (e.g., CASTNet) compared to urban ozone monitors. In addition to the list of PM and gaseous species from EPA's guidance document listed above, the CENRAP operational evaluation will also examine model's ability to estimate $PM_{2.5}$ mass, $PM_{10}$ mass, extinction ($b_{ext}$) and deciview (dV) using the equation above to obtain extinction from fine particulate concentrations.

At some of the IMPROVE sites there are also direct measurements of hourly extinction using transmissometer or nephelometer instruments. Thus, it would be scientifically interesting to evaluate the model estimated extinction with the hourly measured values at these sites. If such a comparison were to be made, day-specific hourly and site-specific RH and f(RH) values could be used to convert model estimated fine particulate matter to extinction rather than the monthly f(RH) EPA recommends for projecting visibility improvements (EPA, 2003a). Furthermore, there have been some discrepancies reported in the past between the extinction calculated using the reconstructed mass data and the hourly measurements. Thus, it would be interesting to calculate day-specific extinction using day-specific f(RH) values and the reconstructed mass data for comparison with the direct measurements and the model estimates. Whether and to what extend this more detailed investigation into light extinction is pursued would depend upon the availability of additional resources.

As part of the CMAQ/CAMx operational evaluation, model outputs will be compared statistically and graphically to observational data obtained from the IMPROVE, SEARCH, CASTNet, EPA-FRM, EPA-STN, and other monitoring networks. These monitoring data will be obtained from AIRS, VIEWS, and other appropriate organizations. These comparisons will likely include:

- Daily monthly, seasonal and annual averages for $SO_2$, $SO_4$, $NO_3$, EC, OC, $PM_{2.5}$, and $PM_{10}$, taking care to exclude periods of sampling interference in the observational data. We will look for systematic biases between the model results and IMPROVE observations, and if biases are found, identify possible sources of error in the model inputs.

- Hourly, high resolution PM species and gaseous species concentrations at sites where available (e.g., SEARCH, AIRS and EPA-Supersites).

- At sites with contrasting aerosol mass loadings, analysis of the temporal behavior of the major scattering and absorbing aerosol constituents along with the visibility trends, to establish correlations

The optional CMAQ/CAMx diagnostic model evaluations may entail several components, many of which can be identified presently. Of course, the actual diagnostic analyses to be performed and the scope of such analyses can only be determined once the initial operational model evaluation is underway. These potential diagnostics analyses will need to be carefully defined and rank-ordered in terms of their priority to ensure that they can be accommodated within available resources and schedule. Among the diagnostic model evaluation analyses that could be considered are:

- Evaluate seasonal trends in observations of organic and inorganic aerosol precursors and their effects on PM composition and visibility, and evaluate the ability of the model to capture these seasonal trends.

- Evaluate how well the model simulates various physicochemical processes by: (a) examining observed and modeled correlations between various species pairs, and (b) comparing model-predicted ratios of various species (individual or families) with observations to evaluate gas/particle partitioning (e.g., nitrate/total nitrate, $SO_4/SO_x$).

- Investigate the performance of the model at selected observational sites characterized by different chemical regimes that may be encountered either spatially or during different seasons to help identify any inadequacies in the model and to provide a better understanding of conditions under which model inferences may be weak.

- Compare hourly 24-hour average, and episode averages of the PM constituents for the 2002 36 km annual modeling period as well as the 12 km episodes across all sites.

- Create scatter plots of modeled vs. observed data and hourly and 24-hour averages by site and subregion to help identify any site-specific biases.

- Create time series plots of predicted and observed concentrations as appropriate.

- Evaluate for total sulfur ($SO_2 + SO_4$), nitrate ($HNO_3 + NO_3$) and ammonia ($NH_3 + NH_4$).

- Compare observed versus modeled mass fractions of PM constituents at various sites that are characterized by their proximity or remoteness relative to sources, or by specific meteorological conditions (e.g., frontal passage, stagnation, precipitation); these will enable identification of trends in the model of over- or under-prediction of specific PM constituents under these conditions.

- Calculate the measured and predicted relative abundance of key PM components and compare with EPA guideline recommendations and emergent alternative science recommendations (e.g., removing the soil component from the calculations, use of alternative relative importance equations [i.e., Boylan, 2004]).

- Pay particular close attention to the model performance at the Class I areas for $SO_4$, $NO_3$, EC, OC, IP and CM on the 20% best and 20% worst days to document whether certain of these days should be eliminated from the visibility projections due to inadequate model performance.

The suite of statistical metrics and graphical tools identified in the previous section for the core operational evaluation efforts would likely also be used to diagnose performance problems with the CMAQ/CAMx simulations should they exist and to highlight differences between model runs. Experience in ozone/PM modeling is the best basis upon which to identify obviously flawed simulation results. Efforts to improve the CMAQ/CAMx model's base case performance will be made, where necessary, warranted (i.e., to reduce the discrepancies between model estimates and observations), and consistent with the project resources and schedule; however, these model performance improvements efforts must be based on sound scientific principles. "Curve-fitting" exercises will be avoided.

## 7.7 Performance Goals and Benchmarks

Establishment of performance goals and benchmarks for regulatory modeling is a necessary but difficult activity. Here, performance goals refer to targets that we believe a good performing model should achieve, where as performance benchmarks are based on historical model performance measures for the best performing simulations. Performance goals are necessary in order to provide consistency in model applications and expectations across the country and to provide standardization in how much weight may be accorded modeling study results in the decision-making process. It is a problematic activity, though, because many areas present unique challenges (e.g., Houston, San Joaquin Valley, Los Angeles) and no one set of performance goals is likely to fit all needs. Equally concerning is the very real danger that modeling studies will be truncated when the 'statistics look right' before full assessment of the model's reliability is made. This has the potential from breeding built-in compensating errors (Reynolds et al., 1996) as modelers strive to get good statistics as opposed to searching for the explanations for poor performance and then rectifying them. A NARSTO review of more than two-dozen urban-scale ozone SIP applications found this tendency to be all too prevalent in the regulatory modeling of the 1990s. (Roth et al, 1997).

Nearly 15 years ago, research sponsored by the California Air Resources Board (Tesche et al., 1990) led to the agency's adoption of three performance goals for 1-hour ozone modeling in the state:

- Unpaired (in time and space) peak prediction accuracy ($\leq \pm$ 20%);
- Mean normalized bias in hourly averaged concentrations ($\leq \pm$ 15%); and
- Mean normalized gross error in hourly concentrations ($\leq$ 35%).

These performance goals for 1-hr ozone concentrations were adapted from previous surveys of several dozen urban-scale photochemical grid modeling studies (principally in California) focusing on ozone episodes of 1 to at most 3 days in duration. A surprising number of these studies did not include biogenic VOC emissions in the inventory under the then prevailing belief that biogenics were a negligibly small source category compared to automobile emissions. Most of the studies (Tesche, 1985, 1988; Tesche et al., 1985; 1990) comprising the data base from which the California ozone performance goals were derived entailed hourly ozone concentrations well above background levels (~40-50 ppb). As a result, it was common practice to use a "cutoff values" ranging between 40 ppb to 60 ppb to eliminate prediction-observations pairs that would cause these bias and error residual statistics to become extraordinarily large when measured concentrations were low. Accordingly, normalized statistics such as bias and error proved to be suitable in most applications since the observed concentrations were generally high. These three California ozone model performance goals were adopted by EPA (1991) as part of the nationwide photochemical modeling guidelines and have been heavily used since.

However, when these evaluation metrics and goals were later adapted to PM and PM species, difficulties arose because performance statistics that divide by low concentration observations become much less useful. Indeed, some PM species may approach zero (e.g., $NO_3$). In time, this has led to the introduction of the fractional and normalized mean bias and error metrics in addition to the mean normalized metrics and related performance expectations based on these alternative measures.

While the 1-hr metrics and goals still have value in interpreting ozone and some gas-phase species performance, it has been necessary to develop new performance metrics and goals for fine particulates. EPA's PM guidance document (EPA, 2001) guidance document identifies particulate matter components of interest to include: $SO_4$ and/or S, $NH_4$, $NO_3$, mass associated with $SO_4$, mass associated with $NO_3$, EC, OC, IP, and mass of individual constituents of inorganic primary particulate matter (i.e., IP). Gaseous pollutants of interest include ozone, $HNO_3$, $NO_2$, PAN, $NH_3$, $NO_y$, $SO_2$, CO, and $H_2O_2$. In addition, EPA guidance identifies several potentially useful statistical measures including: (a) accuracy of spatially averaged concentrations near a monitor, (b) fractional bias in means and standard deviations of predictions and observations, (c) normalized bias, (d) normalized gross error, (e) unpaired comparisons between predicted and observed peak concentrations. (Interested readers are referred to the EPA guidance document on the details of these metrics including mathematical formulae and implementation methods.)

As with ozone in the 1980s, actual experience with PM models has led to the development of the current performance expectations for these models. For example, $PM_{10}$ SIP model performance goals for mean normalized gross error of $\leq$ 30% for southern California (SCAQMD, 1997; 2003) and $\leq$ 50% for Phoenix (ENVIRON, 1998) have been used. As correctly pointed out by Seigneur and co-workers (2003), the current ability of regional PM models to predicting regional PM and visibility is an area of research with improvements needed for characterizing meteorology and emissions as well as PM models themselves. To this list we would add the need for improvements in model evaluation methodologies as well.

When EPA's draft guidance was developed nearly four (4) years ago, an interim set of fine particulate modeling performance goals were suggested for aggregated mean normalized gross error and mean normalized bias as follows:

| Pollutant | Gross Error | Normalized Bias |
|---|---|---|
| **$PM_{2.5}$** | **~30-50%** | **~"10%** |
| **Sulfate** | **~30-50%** | **~"20-30%** |
| **Nitrate** | **~20-70%** | **~"15-50%** |
| **EC** | **~15-60%** | **NA** |
| **OC** | **~40-50%** | **~"38%** |

Because regional-scale fine particulate and regional haze modeling is an evolving science, and considerable practical application and performance testing has transpired in the intervening years since these goals were postulated, we consider them general guidelines. Results of the WRAP, VISTAS, and MRPO model evaluation together with recommendations from science workshops (e.g., EPA's PM Model Performance Evaluation Workshop in February 2004) and recently published scientific studies (e.g., Boylan, 2004) will be used to provide support to these recommendations.

We regard the above goals as simply general interim guidance. Certainly, more information on likely performance expectations will be available once the CENRAP ozone and PM performance evaluation and sensitivity results are fully analyzed. In the CENRAP annual modeling we will generate the model performance statistics listed in Table 7-3 and make comparisons with EPA's interim goals as one means for:

- Establishing a benchmark on the annual 2002 CENRAP modeling episode, and

- Guiding the interpretation of the CMAQ and CAMx modeling results.

Equally as useful, we will endeavor to compare the results of the CENRAP multi-species evaluations of CMAQ and CAMx with the preliminary CENRAP modeling (Pun, Chen and Seigneur, 2004; Tonnesen and Morris, 2004), the results from the BRAVO study (Pitchford et al., 2004) as well as results from EPA and the other RPOs.

## 7.8 Diagnostic and Sensitivity Testing

Rarely does a modeling team find that the first simulation satisfactorily meets all (or even most) model performance expectations. Indeed, our experience has been that initial simulations that 'look very good', usually do so as the result of compensating errors. The norm is to engage in a logical, documented process of model performance improvement wherein a variety of diagnostic probing tools and sensitivity testing methods are used to identify, analyze, and then attempt to remove the causes of inadequate model performance. This is invariably the most technically challenging and time consuming phase of a modeling study. We anticipate that the annual CMAQ/CAMx model base case simulations will present some performance challenges that may necessitate focused diagnostic and sensitivity testing in order for them to be resolved. Hopefully, these diagnostic and/or sensitivity tests can be adequately carried out within the resources and schedule of the current work effort. If not, then it may be necessary to draw upon additional resources from the FY2 budget to conduct the necessary work. Where practical, diagnostic or sensitivity analyses, if needed, could be performed on selected episodes within the annual cycle, thereby avoiding the time-consuming task of running CMAQ/CAMx for the fully 2002 period. Below we identify the types of diagnostic and sensitivity testing methods that might be employed in diagnosing inadequate model performance and devising appropriate methods for improving the model response.

### 7.8.1 Traditional Sensitivity Testing

Model sensitivity experiments are useful in three distinct phases or 'levels' of an air quality modeling study and all will be used as appropriate in the CENRAP modeling with CMAQ and CAMx. These levels are:

- **Level I**. Model algorithm evaluation and configuration testing;
- **Level II**. Model performance testing, uncertainty analysis and compensatory error diagnosis, and
- **Level III.** Investigation of model output response (e.g., ozone, aerosol, deposition) to changes in precursors as part of emissions control scenario analyses.

Most of the Level I sensitivity tests with CMAQ and CAMx have already been completed by the model developers and the RPOs. However, given the open community nature of the CMAQ and

CAMx models, and the frequent science updates to the model and supporting databases, it is possible that some additional configuration sensitivity testing will be necessary.

Potential Level II sensitivity analyses might be helpful in accomplishing the following tasks:

- To reveal internal inconsistencies in the model;
- To provide a basis for compensatory error analysis;
- To reveal the parameters (or inputs) that dominate (or do not dominate) the model's operation;
- To reveal propagation of errors through the model; and
- To provide guidance for model refinement and data collection programs.

At this time, it is not possible to identify one or more Level II sensitivity runs that might be needed to establish a reliable annual 2002 CMAQ or CAMx base case. The merits of performing Level II sensitivity testing will depend upon whether performance problems are encountered in the operational evaluation. Also, the number of tests possible, should performance difficulties arise, will be limited by the additional resources and schedule. Thus, at this juncture, one cannot be overly prescriptive on the number and emphasis of sensitivity runs that may ultimately be desirable. However, from past experience with CMAQ, CAMx and other models, it is possible to identify examples of sensitivity runs could be useful in model performance improvement exercises with the annual 2002 CMAQ/CAMx simulation. These include:

- Modified biogenic emissions estimates;
- Modified on-road motor vehicle emissions;
- Modified air quality model vertical grid structure;
- Modified boundary conditions;
- Modified fire emissions;
- Modified EGU emissions;
- Modified ammonia emission estimates.
- Modified aerosol/$N_2O_5$/$HNO_3$ chemistry; and
- Modified $NH_3$ and $HNO_3$ deposition velocities.

If necessary, Process Analysis extraction outputs can be included in these Level II diagnostic sensitivity simulations in order to provide insight into why the model responds in a particular way to each input modification. Again, the number, complexity, and importance of these types of traditional sensitivity simulations can only be determined once the initial CMAQ/CAMx annual 2002 simulation(s) are executed.

Level III sensitivity analyses have two main purposes. First, they facilitate the emissions control scenario identification and evaluation processes. Today, four complimentary sensitivity "tools" can be used in regional photochemical models depending upon the platform being used. These methods include: (a) traditional or 'brute force' testing, (b) the direct decoupled method (DDM), (c) Ozone Source Apportionment Technology (OSAT) and PM Source Apportionment Technology (PSAT), and (d) Process Analysis (PA). Each method has its strong points and they will be employed where needed and as resources are available. The second purpose of Level III sensitivity analyses is to help quantify the estimated reliability of the air quality model in simulating the atmosphere's response to significant emissions changes. This important model evaluation need is addressed in further detail in section 7.9 below.

Based on experience in other regional studies, examples of Level III monthly or annual sensitivity runs for CENRAP annual modeling might include:

- Ozone, sulfate, nitrate, ammonium and other aerosol sensitivities to $SO_2$ emissions;
- Ozone, sulfate, nitrate, ammonium and other aerosol sensitivities to elevated point source $NO_x$ emissions;
- Ozone, sulfate, nitrate, ammonium and other aerosol sensitivities to ground level $NO_x$ emissions; and
- Sulfate, nitrate, ammonium and other aerosol sensitivities to ammonia.

Of course, traditional 'brute force' sensitivity experiments are just one way of quantifying these or other Level III sensitivities. Other methods that can be applies include DDM, OSAT, or PSAT simulations.

The need to perform sensitivity experimentation (Levels I, II, or III) will depend on the outcome of the CENRAP 2002 annual operational performance evaluations. If such a need arises, the ability to actually carry out selected sensitivity and/or diagnostic experiments will hinge on the availability of additional resources and sufficient time to carry out the analyses. Clearly, selection of the specific analysis method will depend upon the nature of the technical question(s) being addressed at the time.

### 7.8.2 Diagnostic Tests

A rich variety of diagnostic probing tools are available for investigating model performance issues and devising appropriate means for improving the model and/or its inputs. Previously, in section 7.4.4 we introduced the suite of 'probing tools' available for use in the CMAQ and CAMx modeling systems. Where the need exists (i.e., if performance problems are encountered) and assuming CENRAP elects to fund the use of the probing tool applications, these techniques could be employed as appropriate to assist in the model performance improvement efforts associated with the annual 2002 CMAQ/CAMx base case development. Here we describe an additional diagnostic method – indicator species and species ratios -- that is potentially useful not only in model performance improvement activities but also in judging the models reliability in estimating the impacts on air quality from future emissions. This method involves the use of so-called 'indicator species' and species ratios. If, during the conducting of the CENRAP annual simulations we determine that application of indicator species and species ratio techniques would be beneficial to the study (and if existing project resources allow), we will discuss with the CENRAP Project Manager the merits of including this additional probing tool as part of additional work efforts.

Beginning in the mid 1990s, considerable interest arose in the calculation of indicator species and species ratios as a means of diagnosing photochemical model performance and in assessing model credibility in estimating the effects of emissions changes. Major contributions to the development and refinement of this general diagnostic method over the past decade have been made many scientists including Milford et al., (1994), Sillman (1995, 1999), Sillman et al., (1997), Blanchard (2000), Blanchard and Fairley (2001), and Arnold et al., (2003). Indeed, a recent evaluation of CMAQ using indicator species ratios such as $O_3/NO_x$, $NO_z/NO_y$ (a measure of chemical aging), and $O_3/NO_z$ (a measure of the ozone production efficiency per $NO_x$ converted), showed not only good agreement with measurements (Arnold et al., 2003) but also convincingly demonstrated the utility of the method for diagnosing model performance in a variety of ways.

Traditionally, indicator species analyses have focused on ozone and its precursor and product species. However the method is equally applicable to PM species and species ratios given sufficient measurement data for comparisons. With some of the high-resolutions monitoring data available from the EPA Supersites (e.g., St. Louis in the CENRAP region), it is indeed feasible to compute relevant indicator species and ratios for PM and its component species. For example, Ansara and Pandis (1998) demonstrated how indicator species ratios could be applied to show how the modeled mass of PM might respond to sulfate, nitrate and ammonia emissions-related reductions.

## 7.9 Corroborative and Weight of Evidence Modeling Analyses

This section identifies additional modeling analyses that might be worth pursuing under additional funding to add strength to the core model evaluation efforts already planned as part of the CENRAP 2002 annual operational evaluation.

### 7.9.1 Corroborative Models

Noteworthy in EPA's new ozone, PM, and regional haze guidance documents is the encouragement of the use of alternative modeling methods to corroborate the performance findings and control strategy response of the primary air quality simulation model. This endorsement of the use of corroborative methodologies, stems from the common understanding that no single photochemical modeling system can be expected to provide exact predictions of the observed ozone and PM species concentrations in a region the size of CENRAP, especially over time scales spanning 1-hr to 1 year. Although the photochemical/PM models identified in EPA's PM/regional haze guidance document possess many up-to-date science and computational features, there still can be important differences in modeled gas-phase and aerosol predictions when alternative models are exercised with identical inputs.

Mindful of EPA's endorsement of corroborative modeling methods and the rigorous use of 'weight of evidence' investigations, we recommend that the most recent version of CMAQ and CAMx be carried through the study, including the evaluation of emissions control strategies. Among other things, this will permit us to more explicitly identify the expected range of model uncertainty and to corroborate the general effectiveness of the CMAQ and CAMx regional haze control strategies. Other corroborative modeling methods such as the CMAQ-AIM and CMAQ-MADRID should also be considered. However, as these models are derivatives of CMAQ they would not provide as robust independent corroboration as CAMx.

### 7.9.2 Weight of Evidence Analyses

EPA's guidance recommends three general types of 'weight of evidence' analyses in support of the attainment demonstration: (a) use of air quality model output, (b) examination of air quality and emissions trends, and (c) the use of corroborative modeling such as observation-based (OBM) or observation-driven (OBD) models. We will consider the use of one or more methods in conducting the CMAQ/CAMx modeling because it could significantly strengthen the credibility and reliability of the modeling available to the states for their subsequent use. The exact details of the 'weight of evidence' analyses must wait until the CENRAP modeling study evolves further. It is premature to prescribe which, if any of the WOE analyses would be performed since the model's level of performance with the 2002 episode is obviously not known at this time and the time and remaining project resources available to support WOE analyses is unknown as well. Nonetheless, we outline below our thoughts regarding what would likely be considered should the operational CMAQ/CAMx model evaluation need to be bolstered with WOE analyses.

**Use of Emissions and Air Quality Trends.** A limited scope emissions and trend analysis could be employed to support the 'weight of evidence' determinations. However, traditionally, these types of analyses are performed by the lead agency's own staff. With this expectation, we would coordinate our efforts with the CENRAP Modeling Workgroup to develop a trends analysis supporting the future year applications of CMAQ/CAMx.

**Use of Corroborative Observational Modeling**. While regulatory modeling studies for ozone attainment demonstrations have traditionally relied upon photochemical models to evaluate ozone control strategies, there has recently been growing emphasis on the use of data-driven models to corroborate the findings of air quality models. As noted, EPA's guidance now encourages the use of such observation-based or observation-driven models (OBMs/ODMs). We will consider the merits of using these techniques as supportive weight of evidence. While the OBD/OBM models cannot predict future year air quality levels, they do provide useful corroborative information on the extent to which specific subregions may be VOC-limited or NOx-limited, for example, or where controls on ammonia or SO2 emissions might be most influential in reducing $PM_{2.5}$. Information of this type, together with results of DDM and traditional 'brute-force' sensitivity simulations, can be extremely helpful in postulating emissions control scenarios since it helps focus on which pollutant(s) to control.

### 7.9.3 Comparison with Other RPOs

WRAP, VISTAS and MRPO are also modeling the 2002 annual period that provides an opportunity to compare the CENRAP modeling results with that from the other RPOs as an independent check of the modeling.

## 7.10 Assessing Model Reliability in Estimating the Effects of Emissions Changes

EPA identifies three methods (EPA, 2001, pg. 228) potentially useful in quantifying a model's reliability in predicting air quality response to changes in model inputs, e.g., emissions. These include:

- Examination of conditions for which substantial changes in (accurately estimated) emissions occur;

- Retrospective modeling, that is, modeling before and after historical significant changes in emissions to assess whether the observed air pollution changes are adequately simulated; and

- Use of predicted and observed ratios of 'chemical indicator species'.

We note that in some urban-scale analyses, the use of weekday/weekend information has been helpful in assessing the model's response to emissions changes. However, we suspect that this approach would not prove feasible with an annual episode over geographic domain as large as CENRAP.

The first two methods have actually been considered for over 15 years and were the subject of intensive investigations in the early 1990s in Southern California in studies sponsored by the South Coast Air Quality Management District (Tesche, 1991) and the American Petroleum Institute (Reynolds et al., 1996). To date, neither method has proven useful largely because of the great difficulty in developing historical emissions inventories of sufficient quality to make such an analysis credible and the difficulties in removing the influences of different meteorological conditions such that the modeling signal reflects only the model's response to emissions changes. It is difficult enough to construct reliable emissions inventories using today's modeling technology let alone construct retrospective inventories 5-10 years ago prior to the implementation of significant emissions control programs or major land use changes. The use of indicator species, however, offers some promise.

Recent analytical and numerical modeling studies have demonstrated how the use of ambient data and indicator species ratios can be used to corroborate the future year control strategy estimates of Eulerian air quality models. Blanchard et al., (1999), for example used data from environmental (i.e., smog) chambers and photochemical models to devise a method for evaluating the 1-hr ozone predictions of models due to changes in precursor NOx and VOC emissions. Reynolds et al., (2003) followed up this analysis, augmented with process analysis, to assess the reliability of SAQM photochemical model estimate of 8-hr ozone to precursor emissions cutbacks. With respect to secondary aerosol PM, the recent CMAQ evaluation by Arnold et al. (2003) clearly demonstrated how the use of indicator species analysis could be use to develop insight into the expected reliability and adequacy of a photochemical/PM model for simulating the effects of emissions control scenarios. These researchers used three indicator ratios (or diagnostic 'probes') to quantify the model's response to input changes:

- The ozone response surface probe [$O_3/NO_x$];

- The chemical aging probe [$NO_z/NO_y$ ]; and

- The ozone production efficiency probe [$O_3/NO_z$ ].

By closely examining the CMAQ's response to key input changes, properly focused in time and spatial location, Arnold et al., (2003) were able to conclude that the photochemical processing in CMAQ was substantially similar to that in the atmosphere

Thus, the extension of these techniques to address CMAQ and CAMx predictions for secondary aerosols will doubtless be quite challenging, but the use of indicator species (e.g., ammonia or $HNO_3$ limitation for nitrate particle formation) and species ratios appears to offer, at this time, the only real opportunity to quantify the expected reliability of the air quality model to correctly simulate the effects of emissions changes. In the CENRAP CMAQ and CAMx model evaluation, we will remain alert to opportunities to extend the indicator species ratio analyses to the problem of fine particulate and regional haze. This is one area where technical collaboration between the Emissions and Air Quality Modeling team and the CENRAP Modeling Workgroup can be especially fruitful in terms of identifying and testing emergent methods for challenging the model's ability to correctly simulate the effects of future year emissions changes. Finally, we note that this is truly a current research area and as such falls outside the scope of the current CENRAP annual modeling effort. However, given its importance, we will remain alert to opportunities to utilize newly available methods should this prove feasible within the CENRAP resources and schedule.

**Table 7-1**. Ground-level ambient data monitoring networks and stations available in the CENRAP states for CY-2002.

| Monitoring Network | Chemical Species Measured | Sampling Frequency; Duration | Approximate Number of Monitors |
|---|---|---|---|
| **IMPROVE** | Speciated $PM_{2.5}$ and $PM_{10}$ | 1 in 3 days; 24 hr | 11 |
| **CASTNET** | Speciated $PM_{2.5}$, Ozone | Hourly, Weekly; 1 hr, Week | 3 |
| **SEARCH** | 24.hr PM25 (FRM Mass, OC, BC, SO4, NO3, NH4, Elem.); 24-hr PM coarse (SO4, NO3, NH4, elements); Hourly PM2.5 (Mass, SO4, NO3, NH4, EC, TC); and Hourly gases (O3, NO, NO2, NOy, HNO3, SO2, CO) | Daily, Hourly; | 0 |
| **NADP** | WSO4, WNO3, WNH4 | Weekly | 23 |
| **EPA-FRM** | Only total fine mass ($PM_{2.5}$) | 1 in 3 days; 24 hr | **(?)** |
| **EPA-STN** | Speciated $PM_{2.5}$ | Varies; Varies | 12 |
| **AIRS/AQS** | CO, NO, $NO_2$, $NO_x$, $O_3$ | Hourly; Hourly | 25 |
| **EPA-SPEC** | Various as part of St. Louis Super Site | Various | 1+ |

**Table 7-2**. Aircraft sampling programs performed during CY-2002.

| Aircraft Program | Meteorological Parameters & Chemical Species Measured | Sampling Program & Flight Duration | Approximate Number of Flights; Days; Aircraft |
|---|---|---|---|
| **University of Maryland (UMD); Univ. Research Foundation (URF)** | ***Meteorology:*** WS, WD, Temp, RH, ***Air Quality***: $O_3$, NO, $NO_2$, CO, $SO_{2,}$ aerosol absorption, aerosol scattering. | 23 May to 3 Oct; Typically 3 hrs | 54 flights, 54 days, 1 aircraft |
| **Midwest RPO & Wisconsin DNR** | ***Meteorology***: WS, WD, Temp, RH, dew point, pressure ***Air Quality***: $O_3$, NO, $NO_2$, $NO_x$, $NO_y$, speciated VOCs, carbonyls, $HNO_3$, $NH_3$, Hg, $SO_4$, OC, EC, $PM_{2.5}$, light scattering (Neph), visibility pictures. | 1 June to 22 Nov; Typically 3-5 hrs | 133 flights; 29 days; 2 aircraft (WDNR and Jacko Aircraft) |

**Table 7-3.** Core statistical measures to be used in the CENRAP 2002 annual air quality model evaluation with ground-level data (see ENVIRON, 2003b,d for details).

| Statistical Measure | Shorthand Notation | Mathematical Expression | Notes |
|---|---|---|---|
| Accuracy of paired peak (***$A_p$***) | **Paired_Peak** | $\frac{P - O_{peak}}{O_{peak}}$ | $P_{peak}$ = paired (in both time and space) peak prediction |
| Coefficient of determination (***$r^2$***) | **Coef_Determ** | $\frac{\left[\sum_{i=1}^{N}(P_i - \overline{P})(O_i - \overline{O})\right]^2}{\sum_{i=1}^{N}(P_i - \overline{P})^2 \sum_{i=1}^{N}(O_i - \overline{O})^2}$ | $P_i$ = prediction at time and location *i*; $O_i$ = observation at time and location *i*; $\overline{P}$ = arithmetic average of $P_i$, i=1,2,…, N; $\overline{O}$ = arithmetic average of $O_i$, i=1,2,…,N |
| Normalized Mean Error (***NME***) | **Norm_Mean_Err** | $\frac{\sum_{i=1}^{N}\lvert P_i - O_i\rvert}{\sum_{i=1}^{N} O_i}$ | Reported as % |
| Root Mean Square Error (***RMSE***) | **Rt_Mean_Sqr_Err** | $\left[\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2\right]^{1/2}$ | Reported as % |
| Fractional Gross Error (***$F_E$***) | **Frac_Gross_Err** | $\frac{2}{N}\sum_{i=1}^{N}\left\lvert\frac{P_i - O_i}{P_i + O_i}\right\rvert$ | Reported as % |
| Mean Absolute Gross Error (***MAGE***) | **Mean_Abs_G_Err** | $\frac{1}{N}\sum_{i=1}^{N}\lvert P_i - O_i\rvert$ | |
| Mean Normalized Gross Error (***MNGE***) | **Mean_Norm_G_Err** | $\frac{1}{N}\sum_{i=1}^{N}\frac{\lvert P_i - O_i\rvert}{O_i}$ | Reported as % |

| Statistical Measure | Shorthand Notation | Mathematical Expression | Notes |
|---|---|---|---|
| Mean Bias (***MB***) | **Mean_Bias** | $\frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right)$ | Reported as concentration (e.g., µg/m$^3$) |
| Mean Normalized Bias (***MNB***) | **Mean_Norm_Bias** | $\frac{1}{N}\sum_{i=1}^{N}\frac{\left(P_i - O_i\right)}{O_i}$ | Reported as % |
| Mean Fractionalized Bias (Fractional Bias, ***MFB***) | **Mean_Fract_Bias** | $\frac{2}{N}\sum_{i=1}^{N}\left(\frac{P_i - O_i}{P_i + O_i}\right)$ | Reported as % |
| Normalized Mean Bias **(*NMB*)** | **Norm_Mean_Bias** | $\frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N}O_i}$ | Reported as % |
| Bias Factor (***BF***) | **Bias Factor** | $\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_i}{O_i}\right)$ | Reported as BF:1 or 1: BF or in fractional notation (BF/1 or 1/BF). |

## 8.0 VISIBILITY ASSESSMENT AND COMPARISON

This chapter provides a summary of how the CENRAP 2002 annual modeling results will be used to satisfy the requirements of the Regional Haze Rule (RHR) as part of the Section 308 visibility State Implementation Plans (SIPs) and Tribal Implementation Plans (TIPs) due in 2007/2008. Note that the CENRAP 2002 modeling results may also be used in some capacity as the regional component in some State's fine particulate ($PM_{2.5}$) and 8-hour ozone SIPs that are also due in 2007/2008. However, how the modeling results will be used to address $PM_{2.5}$ and 8-hour ozone issues is not discussed in this chapter and will be addressed by the pertinent States at a later time. The purpose of this section is to start planning how the modeling results will be used to project visibility changes for the Section 308 SIPs/TIPs to optimize the presentation of results and identify any potential pitfalls early on.

### 8.1 Regional Haze Rule Requirements

The Regional Haze Rule (RHR) regulations were published by EPA in 1999 (Federal Register 35769, July 1, 1999) and are designed to address the requirements of Section 169A and 169B of the Clean Air Act (CAA). The CAA and RHR call for the protection of visibility at the 156 "mandatory Federal Class I areas."[1] Section 169A of the Clean Air Act Amendments (CAAA) established a national visibility goal to remedy existing impairment due to air pollution at the Class I areas. This is accomplished by defining a visibility goal of "natural conditions" to be achieved at each Class I area by 2064. The RHR requires states with Class I areas to develop SIPs that include reasonable progress goals for improving visibility in each Class I area and emission reduction measures to meet those goals. For the initial control strategy SIPs due in 2007/2008, states are required to adopt progress goals for improving visibility from baseline conditions. The 2000-2004 five-year period is used to define baseline conditions and the first future progress period is 2018, that has been interpreted as either the 2013-2017 or 2014-2018 five year periods. A state is required to set progress goals for each Class I area in the state for two visibility metrics:

- Provide for an improvement in visibility for the most impaired (i.e., 20% worst) visibility days; and

- Ensure no degradation in visibility for the least impaired (i.e., 20% best) visibility days.

The reasonable progress goals must provide for a rate of improvement sufficient to attain "natural conditions" by 2064, or justify any alternative rate. States are to define controls to meet progress goals every 10 years, starting in 2018, that defines progress periods ending in 2018, 2028, 2038, 2048, 2058 and finally 2064. States will determine whether they are meeting their goals by comparing visibility conditions from one five-year period to another (e.g., 2000-2004 to 2013-2017). As stated in 40 CFR 51.308 (d) (1), baseline visibility conditions, progress goals, and

[1] Areas designated as mandatory Class I areas are those National Parks exceeding 6,000 acres, wilderness areas and national memorial parks exceeding 5,000 acres, and all international parks that were in existence on August 7, 1977. Since then some new Class I areas have been defined and some existing Class I areas have been expanded. However, these areas are not included as part of the 156 "Federal mandated Class I areas" for which the RHR applies. In this document the term "Class I area" refers to the 156 "Federal mandated" Class I Areas.

changes in visibility must be expressed in terms of deciview (dv) units. The deciview unit of visibility impairment is derived from light extinction ($b_{ext}$) as follows:

$$dv = 10 \ln (b_{ext}/10)$$

Where light extinction ($b_{ext}$) is expressed in terms of inverse megameters ($Mm^{-1} = 10^{-6}\ m^{-1}$). Section 8.2.1.1 below discusses how $b_{ext}$ is calculated using the observed fine particulate concentrations from the IMPROVE monitors and Section 8.2.1.2 below defines how the modeling results will be mapped to the IMPROVE measurements for the purpose of projecting visibility improvements.

## 8.2 EPA Guidance for Complying with Regional Haze Rule

EPA has published three guidance documents that relate to how modeling results should be used to demonstrate compliance with the RHR:

"Guidance for Demonstrating Attainment of Air Quality Goals for PM 2.5 and Regional Haze, Draft 2.1 (EPA, 2001).

"Guidance for Tracking Progress Under the Regional Haze Rule" (EPA, 2003a).

"Guidance for Estimating Natural Visibility Conditions Under the Regional Haze Rule" (EPA, 2003b).

The first EPA modeling guidance document listed above (EPA, 2001) discusses the use of modeling results to demonstrate progress toward the RHR goals. This is a draft document that is scheduled to be updated by the end of 2004. When the draft modeling guidance is updated that may necessitate an update to this Modeling Protocol to keep it consistent with current EPA guidance. The second document (EPA, 2003a) focuses on monitored visibility and how to define the visibility baseline and how to track visibility goals. The third EPA guidance document discusses procedures for defining "natural conditions" for a Class I area that is the visibility goal in 2064. In the discussion below we are assuming that other CENRAP participants would define the "natural conditions" for each CENRAP Class I area. In this Chapter we address the use of the modeled and monitored fine particulate concentrations at the Class I areas for projecting visibility changes from the current to future years.

Below we describe EPA's approach for projecting visibility improvements. Some concerns have been raised regarding some of the assumptions in EPA's visibility projection approach, such as: (1) EPA assumes that sulfate is fully neutralized by ammonium which may not always be the case; (2) EPA's natural background fails to account for sodium chloride which may be an important component at some sites; and (3) the adjustment factor assumed to convert organic carbon (OC) to organic matter (OM) (1.4) may be too low. There may be some updates to EPA's visibility projection approach, but the basic approach will likely remain unchanged.

### 8.2.1 Procedures for Projecting Visibility Changes

EPA guidance recommends using the model in a relative sense to project future-year visibility conditions (EPA, 2001; 2003a). This is done through the use of Relative Reduction Factors (RRFs) that are defined as the ratio of the future-year to the current-year modeling results. The RRFs are applied to the baseline visibility conditions to project future-year visibility. The major features of EPA's recommended visibility projections are as follows (EPA, 2003a,b):

- Monitored data should be used to define current air quality.
- Monitored concentrations of $PM_{10}$ are divided into six major components, the first five of which are assumed to be $PM_{2.5}$ and the sixth is $PM_{2.5\text{-}10}$.
    - SO4 (sulfate);
    - NO3 (particulate nitrate);
    - OC (organic carbon);
    - EC (elemental carbon);
    - OF (other fine particulate); and
    - CM (coarse matter).
- Models are used in a relative sense to develop relative reduction factors (RRFs) between future and current predicted concentrations of each component.
- Component-specific relative reduction factors are multiplied by current monitored values to estimate future component concentrations.
- Estimates of future component concentrations are consolidated to provide an estimate of future air quality, which can be related to a goal for regional haze.
- Future estimated air quality is compared with the goal to see if the simulated control strategy results in the goal being met.
- It is acceptable to assume that all measured sulfate is in the form of ammonium sulfate [$(NH_4)_2SO_4$] and all particulate nitrate is in the form of ammonium nitrate [$NH_4NO_3$].

In order to facilitate tracking visibility progress, three important visibility parameters are required for each Class I area:

<u>Baseline Conditions</u>: Baseline Conditions represent visibility for the 20% best and 20% worst days for the initial five-year baseline period of the regional haze program. Baseline Conditions are calculated using monitor data collected during the 2000-2004 five-year period.

<u>Current Conditions</u>: Current Conditions for the best and worst visibility days are calculated from the latest five-year average based on the latest five-years of monitored data (which is currently 1999-2003, but expected to be 2000-2004 by the time the CENRAP future-year visibility projections will be made in the latter half of 2005).

Natural Conditions: Estimates of natural visibility conditions for the 20% best and 20% worst days at a Class I area that is the goal of the RHR is 2064 that has been defined as visibility conditions that would be experienced in the absence of human-caused impairment" (EPA, 2003b).

#### 8.2.1.1 Calculation of Baseline Conditions

Baseline and Current Conditions for Class I areas are calculated using fine and coarse particulate concentration measurements at Interagency Monitoring of Protected Visual Environments (IMPROVE) monitors. Currently, each Class I area in the CENRAP domain has an associated IMPROVE PM monitor. The IMPROVE monitors do not directly measure visibility, but instead measure speciated fine particulate ($PM_{2.5}$) and total $PM_{2.5}$ and $PM_{10}$ mass concentrations from which visibility is obtained through the IMPROVE reconstructed mass extinction equation.

Visibility conditions are estimated starting with the IMPROVE reconstructed 24-hour average particulate matter (PM) mass measurements for six PM species:

- Sulfate [$SO_4$];
- Particulate Nitrate [$NO_3$];
- Organic Matter [OM];
- Elemental Carbon [EC];
- Other Fine Particulate [Soil]; and
- Coarse Matter [CM].

The IMPROVE monitors do not directly monitor some of these species so assumptions are made as to how the IMPROVE measurements can be adjusted and combined to obtain these six components of light extinction. For example, sulfate and particulate nitrate are assumed to be completely neutralized by ammonium and are assumed to occur solely in the fine particulate mode (that is any coarse mode sulfate and nitrate in the real atmosphere may be present in the CM IMPROVE measurement). Concentrations for the above six components of light extinction in the IMPROVE reconstructed mass (RCM) extinction equation are obtained from the IMPROVE measured species as follows:

**Table 8-1.** Definition of IMPROVE Reconstructed Mass (RCM) species from measured IMPROVE species.

| IMPROVE RCM | IMPROVE Measured Species |
|---|---|
| $SO_4$ | 1.375 x ( 3 x S) |
| $NO_3$ | 1.29 x NO3_ |
| OM | 1.4*OC1 + 1.4*OC2 + 1.4*OC3 + 1.4*OC4 + 1.4*OP |
| EC | EC1 + EC2 + EC3 – OP |
| Soil | 2.2*AL + 2.49*SI + 1.63*CA + 2.42*FE + 1.94*TI |
| CM | MT – MF |

Where:

- S is elemental sulfur as determined from proton induced x-ray emissions (PIXE) analysis of the IMPROVE Module A that is multiplied by 3 to account the presence of oxygen, if S is missing then the sulfate (SO4) measured by ion chromatography analysis of the Module B is used to replace (3 x S). It is assumed to be completely neutralized by ammonium (1.375 x SO4).
- NO3_ is the particulate nitrate measured by ion chromatography analysis of the Module B. It is assumed to be completely neutralized by ammonium (1.29 x).
- The IMPROVE Organic Carbon (OC) measurements are multiplied by 1.4 to obtain Organic Matter (OM) to adjust the OC mass for other elements assumed to be associated with OC.
- Elemental Carbon (EC), which is also referred to as Light Absorbing Carbon (LAC) is determined by TOR analysis and is the sum of EC fractions minus the pyrolized fraction.
- Soil is determined as a sum of the mass of those elements (measured by PIXE) predominately associated with soil, whose mass is adjusted for oxygen in the common compounds.
- MT and MF are total $PM_{10}$ and $PM_{2.5}$ mass, respectively.

Associated with each PM species is an extinction coefficient that converts concentrations (in :g/m$^3$) to light extinction (in inverse mega meters, $Mm^{-1}$). Sulfate and nitrate are hygroscopic so relative humidity adjustment factors [f(RH)] are used to modify the extinction coefficients that increase the particle's extinction efficiency with increasing RH to account for the particles taking on water and having higher light scattering properties. Note that some Organic Matter (OM) compounds may also have hygroscopic properties, but the IMPROVE reconstructed mass extinction equations assume OM is non-hygroscopic.

$$B_{Sulfate} = 3 \times f(RH) \times [SO_4]$$
$$b_{Nitrate} = 3 \times f(RH) \times [NO_3]$$
$$b_{EC} = 10 \times [EC]$$
$$b_{OM} = 4 \times [OM]$$
$$b_{Soil} = 1 \times [Soil]$$
$$b_{CM} = 0.6 \times [CM]$$

Monthly average f(RH) factors are used as recommended in EPA's guidance (EPA, 2003a). These values have been recently updated (SAIC, 2003) and are available at: ***ftp://ftp.saic.com/raleigh/RegionalHaze_2002FRHcurve/fRH_analysis/***.

The total light extinction ($b_{ext}$) is assumed to be the sum of the light extinction due to the six PM species listed above plus Rayleigh (blue sky) background ($b_{Ray}$) that is assumed to be 10 $Mm^{-1}$.

$$b_{ext} = b_{Ray} + b_{Sulfate} + b_{Nitrate} + b_{EC} + b_{OM} + b_{Soil} + b_{CM}$$

The total light extinction ($b_{ext}$) in $Mm^{-1}$ is related to visual range (VR) in km using the following relationship:

$$VR = 3912 / b_{ext}$$

The Regional Haze Rule requires that visibility be expressed in terms of deciview (dv) that uses natural logarithms of the extinction as follows:

$$dv = 10 \ln(b_{ext}/10)$$

The equations above using data from the associated IMPROVE monitor are used to estimate the daily average visibility at each Class I area. For each year from the 2000-2004 baseline these daily average visibility values, in terms of deciview (dv), are then ranked from highest to lowest. The 20% worst days visibility for each given year is given as the average visibility across the 20% highest visibility days in deciview. Similarly, the 20% best days are given as the average visibility across the 20% lowest visibility days in terms of deciview. The Baseline Conditions is the average of the 20% best and 20% worst days across the five-years of 2000-2004. The Current Conditions is the average of the 20% best and 20% worst days based on the latest five-years of available data, which would be either 1999-2003 or 2000-2004 for the CENRAP annual modeling analysis depending on when the 2004 data become available.

#### 8.2.1.2 Mapping of Modeling Results to the IMPROVE Measurements

As noted above, to project future-year visibility at Class I areas the modeling results are used in a relative sense to scale current observed visibility for the 20% best and 20% worst visibility days using RRFs that are the ratio of modeling results for the future-year to current-year. This scaling is done separately for each of the six components of light extinction in the IMPROVE reconstructed mass extinction equations. The modeled species do not necessarily exactly match up with the IMPROVE reconstructed mass species, thus assumptions must be made to map the modeled species to the IMPROVE reconstructed mass species for the purpose of projecting visibility improvements. For example, the models explicitly simulates ammonium and sulfate may or may not be fully neutralized in the model, whereas the IMPROVE reconstructed mass equations assume sulfate is fully neutralized by ammonium. For the CMAQ Version 4.4 model and the CAMx V4.10s using the M5 coarse/fine mode configuration that are being used in the CENRAP annual modeling, the mapping of modeled species to IMPROVE reconstructed mass (RCM) species listed in Table 8-2.

**Table 8-2a.** Proposed mapping of CMAQ V4.4 modeled species concentrations to IMPROVE Reconstructed Mass (RCM) species.

| IMPROVE RCM | CMAQ V4.3 Species |
|---|---|
| $SO_4$ | 1.375 x (ASO4J + ASO4I) |
| $NO_3$ | 1.29 x (ANO3J + ANO3I) |
| EC | AECJ + AECI |
| OM | AORGAJ + AORGAI + AORGPAJ + AORGPAI + AORGBJ + AORGBI |
| Soil | A25J + A25I |
| CM | ACORS + ASEAS + ASOIL |

**Table 8-2b.** Proposed mapping of CAMx V4.10s using M4 (coarse/fine) modeled species concentrations to IMPROVE Reconstructed Mass (RCM) species.

| IMPROVE RCM | CAMx V4.10s M4 Species |
|---|---|
| $SO_4$ | 1.375 x PSO4 |
| $NO_3$ | 1.29 x PNO3 |
| EC | PEC |
| OM | POA + SOA1 + SOA2 + SOA3 + SOA4 |
| Soil | FCRS |
| CM | CCRS + CPRM |

If a different model is used (e.g., REMSAD, CMAQ-MADRID, CMAQ-AIM or CAMx using multi-sectional algorithms) or updates to CMAQ or CAMx change its species definitions, then the species mappings would have to be defined specific for that model. Note that in the above species mapping it is assumed that all of the CMAQ and CAMx estimated sulfate and particulate nitrate are in the fine mode. If a fully sectional model was used (e.g., CMAQ-AIM or CAMx full-science), then a decision would have to be made whether to map the coarse mode sulfate and nitrate to the IMPROVE fine $SO_4$ and $NO_3$ species (as implicitly assumed for CMAQ and CAMx M4 using mappings in Table 8-2) or to include it in the IMPROVE CM species. In fact this is also an issue for the CMAQ V4.3 modal approach in which the lognormal size distribution for sulfate and nitrate in the accumulation mode (i.e., ASO4J and ANO3J) may have some mass greater than 2.5 :m, however the current usual convention for CMAQ is to assume all the sulfate and nitrate are in the fine mode.

### 8.2.1.3 Using Modeling Results to Project Changes in Visibility

Modeling results are used in a relative fashion to project future-year visibility using relative reduction factors (RRFs). RRFs are expressed of the ratio of the modeling results for the future-year to the results of the base year and are Class I area and PM species specific. RRFs are applied to the individual 24-hour observed PM species measurements for the 20% Worst and 20% Best days that make up the Baseline Conditions to project future-year PM levels from which visibility can be assessed using the IMPROVE RCM extinction equations listed above. EPA has identified the following six steps to project future-year visibility for the 20% best and 20% worst days (EPA, 2001; 2003a):

1. For each Class I area visibility is ranked using IMPROVE reconstructed PM mass (RCM) extinction equation for each year that comprises the five-year Baseline Conditions (i.e., 2000-2004).

2. Calculate the arithmetic average of the 20% best and 20% worst visibility days, in deciview (dv), for each year of the five-year baseline period and then calculate the five-year average of the 20% best and 20% worst days. Document which days during the five-year baseline period comprise the 20% best and 20% worst observed visibility days.

3. Use an air quality model to simulate the baseline period emissions and future-year emissions. Extract 24-hour average PM species concentrations "near" each Class I area. Calculate the average PM species concentration estimates in the current and future-year

simulation across the 20% best and 20% worst observed visibility days for each year in the five-year baseline period. Average across the five-years the average PM species concentrations for the 20% best and 20% worst days and for each year in the five-year baseline period. Calculate Class I area and best/worst 20% days RRFs for each of the six PM components as the ratio of the five-year average estimated PM species concentration in the future-year to current year.

4. Multiply the 20% best/worst days and PM species dependent RRFs by the observed 24-hour PM species concentrations for each day from the 20% best/worst days to obtain future-year daily average PM species concentrations for each day from the 20% best and 20% worst visibility days from the five-year baseline period.

5. Using the future-year estimated 24-hour PM species concentrations for the best/worst 20% observed days, calculate extinction using the IMPROVE reconstructed mass extinction equation and daily deciview, perform annual averaging of the visibility (dv) estimates for the 20% best/worst days for each year and then obtain the five-year average visibility (dv) for the future-year.

When selecting model estimated PM species concentrations "near" the monitor, EPA (2001) recommends taking a spatial average of PM concentrations across a grid cell resolution dependent NX by NY array of cells centered on the grid containing the monitor. For the CENRAP annual modeling using a 36 km grid, just the model estimates for the grid cell containing the monitor will be used (i.e., NX=NY=1). For the 12 km modeling, EPA recommends that NX=NY=3, that is the model estimated PM species concentrations are averaged across 9 grid cells centered on the IMPROVE monitor (EPA, 2001).

#### 8.2.1.4. Exclusion of Days from 20% Best/Worst Conditions

As discussed in Chapter 7, the model performance evaluation will pay particular close attention to the model's PM performance at each Class I area for the 20% Best and 20% Worst days. If the model performance is extremely poor for one or more of these days at a Class I area, it may be desirable to exclude that day when calculating the RRFs for visibility projections. For example, if a day is one of the Worst 20% days at a Class I area and the model estimates clean background concentrations then it would be inappropriate to use the modeling results, even in a relative fashion through the RRFs, to project future-year visibility. The Modeling Team would identify such suspect days and work with the CENRAP Modeling Workgroup on whether they should be excluded from the visibility projections on a case-by-case basis.

## 8.3 CENRAP Future-Year Modeling

The CENRAP future-year modeling will use the 2002 MM5 meteorological conditions. That is, we will assume that the meteorological conditions for the future-year are the same as for 2002. This will allow for the comparison of the changes in visibility at Class I areas from the current (2002) to future-year due to changes in emissions. This means that the effects of climate change, land use variations and climatic variations will not be accounted for in the future-year meteorological inputs. Several other decisions concerning the future-year to be modeled, model(s) to be used and modifications to the model inputs to reflect future years do need to be made such as:

### 8.3.1 Future-Year to be Modeled

Visibility projects are needed for the 20% best and 20% worst days from the five-year baseline period of 2000-2004 to 2018, which has been interpreted as the five-year future-year periods of 2013-2017 or 2014-2018. The EPA guidance documents are inconsistent on what future-year should be used for the assessment. The 2001 modeling guidance says that the mid-year from the 2013-2017 five-year projection period should be used, which would be 2015 (EPA, 2001, pg. 221). The 2003 guidance for tracking progress doesn't discuss the future-year that should be modeled, but does mention that goals should be compared across 5-year planning periods "(e.g., 2000-2004 to 2013-2017) and progress should be measured as improvement from 2004 to "2018" (EPA, 2003, pg. 1-6). CENRAP and other RPOs will be looking to EPA for providing more definitive guidance as to which future-year should be used for modeling visibility progress and this Modeling Protocol will be updated when such information is available, which should be when EPA released the revised $PM_{2.5}$ and regional haze guidance document by the end of 2004. Thus, in the discussion below we just refer to the "future-year" with the actual year to be determined during 2004. Current thinking is that the 2018 future-year will be used in CENRAP.

### 8.3.2 Future-Year Emissions

The future-year emissions or growth and control factors will be provided to the CENRAP annual Emissions and Air Quality Modeling Team (ENVIRON/UCR) by the CENRAP emissions projection contractor (TBD). They would be processed into the gridded speciated hourly three-dimensional emissions inputs for photochemical grid modeling using the SMOKE emissions model and the procedures discussed in Section 5.2. The same biogenic emissions as used in the 2002 Base Case modeling will be used for the future-year modeling. This assumes that the same land use and biomass distribution as used in the 2002 Base Case emissions would exist in the future-year emission scenarios. The effects of urban sprawl, increased agricultural, deforestation, locations of new roads, etc. between the current (2002) and future-year would not be accounted for. Typical Year future-year fire emissions that are consistent with the typical year fire emissions used in the 2002 Typical Year emissions scenario would be used in the future-year. Future-year Typical Year emissions for Electrical Generating Units (EGUs) would also be used for the future-year scenarios.

### 8.3.3 Future-Year Initial and Boundary Conditions

The same Initial Conditions (ICs) as used in the 2002 Base Case would be used in the future-year modeling. Because a ~15 day spin up period is being used, ICs will have minimal if any influence on the model estimated concentrations.

Boundary condition (BC) concentrations that are assumed along the lateral edges of the continental US national RPO 36 km modeling domain for current (2002) year are based on a 2002 simulation of the GEOS-CHEM global climate model. Ideally, the future-year BCs should be based on a GEOS-CHEM simulation using the 2002 meteorology and the future-year global emission inputs. However, currently such a simulation is not available. In the absence of a future-year emissions scenario GEOS-CHEM simulation, we recommend that the same CMAQ/CAMx BCs be used in the future-year as used in the 2002 base year (i.e., 2002 GEOS-CHEM results) as any adjustments to them would be arbitrary and speculative.

### 8.3.4 Other Future-Year Modeling Inputs

All other future-year CMAQ and CAMx modeling inputs would be identical to the 2002 Base Case simulation including meteorology, photolysis rates, geophysical, and other inputs.

## 8.4 Presentation of Results

There are various ways that the future-year modeling results can be presented to display the visibility projections and convey the key findings. Although the projection of visibility improvements at the Class I areas is a key element of the Regional Haze Rule (RHR), analyzing the results solely at the Class I areas is a very narrow view of them, thus we intend to present the future-year modeling results in other forms to convey a more complete picture.

### 8.4.1 Projection of Visibility Improvements at Class I Areas

The results of the visibility projections at each Class I area in and around the CENRAP domain would be displayed graphically as well as in tabular summaries. Spread sheets of the results would be generated ordered by state and provided to CENRAP. The visibility projections at each Class I area would be compared with the 2018 visibility progress goal calculated following EPA's "Guidance for Tracking Progress Under the Regional Haze Rule (EPA, 2003a).

Figure 8-1 displays a sample plot of a visibility progress goal prepared for the Grand Canyon National Park as part of the WRAP Section 309 SIP work effort. The modeled projected visibility could be plotted on the same figure as a graphical way of presenting the results. Shown as a straight line near the bottom of Figure 8-1 is the EPA default "natural conditions" for the 20% worst visibility days as the green line with green triangles that represents the visibility goal at Grand Canyon of 6.97 dv in 2064. The black diamonds on the left side of the plot are the "current conditions" that are based on IMPROVE observations for the 20% worst days for five-year periods ending in 1993 (i.e., 1989-1993) to 2001 (i.e., 1997-2001). The "current conditions" for the latest five years of data, which is 1997-2001 in this graph (the Section 309 SIP analysis was performed in 2003 before the 2002 IMPROVE data were available), is assumed as the "baseline conditions" (i.e., 2000-2004) that is the starting point in 2004 (12.00 dv) for the glide path or linear uniform rate of progress to natural conditions in 2064 (purple line with squares). In this example, the visibility progress goal in 2018 would be 10.83 dv so that a 1.17 dv reduction in visibility is needed by 2018 to meet the visibility progress target in 2018 at the Grand Canyon National Park (GMNP). Note that we would expect higher observed visibility levels for the 20% worst days at many of the CENRAP Class I areas due to higher sulfate concentrations in the eastern US, but

Figure 8-1 illustrates the types of graphs that can be utilized to display visibility progress goals combined with visibility progress projections.

Tabular summaries could list for each Class I area, in terms of deciview, the observed Baseline Conditions, the 2018 Progress Goal, the difference (i.e., visibility progress target) and then the modeled difference for each future-year emissions scenario. Such a table would clearly state which future-year emission scenarios met the RHR 2018 Progress Goals at which Class I areas.

Another useful display of the projected visibility improvements are stacked bar charts of extinction showing the contributions of each of the six major PM components to light extinction for the IMPROVE measured current period to the future-year period and the 20% best and 20% worst days. These side–by-side current to future-year extinction comparisons can graphically display which extinction components are getting better and which are getting worse. They can also be used to identify why the emission controls may not have had as large an effect on visibility as expected. For example, an $SO_2$ emissions control strategy may have had the desired effect in reducing sulfate concentrations but the visibility estimates remain relatively unchanged as the lower sulfate concentrations frees up ammonia that then forms particulate nitrate that compensates for the lower sulfate in the visibility equations (i.e., nitrate replacement). Such a chemical shift in the extinction budget would be readily apparent in the stacked bar charts. The most useful forms to convey the modeling results will evolve over the project.

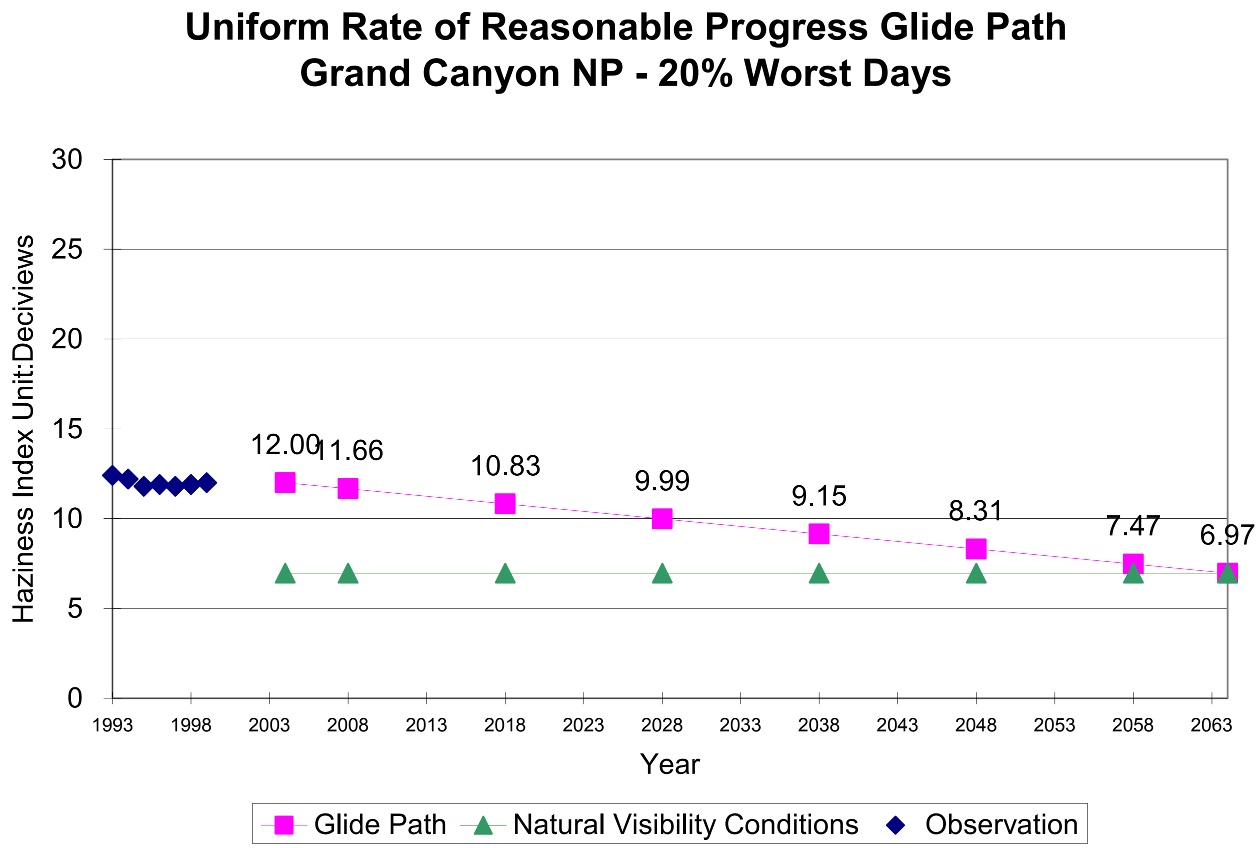


**Figure 8-1.** Example calculation of visibility progress goals using IMPROVE data collected at the Grand Canyon National Park for the 20% Worst days showing EPA default "Natural Conditions" (green triangles) for the 20% Worst days that is the 2064 Goal (6.97 Dv), "Current Conditions" Observed Visibility (Dv) as black triangles whose latest value (12.00 Dv) is assumed as the starting point for the 2000-2004 baseline (first purple square), and the glide path from 2004 baseline to "Natural Conditions" in 2064 assuming linear uniform rate of progress.

### 8.4.2 Spatial Maps of Results

Spatial maps of changes in individual species, total PM mass, extinction and deciview from the 2002 Typical Year to future-year simulations would be used to provide a big picture view of the changes in regional air quality and visibility from 2002 to the future–year. The display of hourly and 24-hour animations as well as longer-term averages (e.g., monthly, seasonal and annual) may provide more insight into the modeling results.

### 8.4.3 Time Series Plots

Time series plots of concentrations extinction and deciview at key receptor locations, such as Class I areas, may also provide additional insight into the modeling results. Hourly time series for selected periods and locations may provide insight into short-term changes, whereas changes in

24-hour concentrations for the year provides information for all days, not just the ~40 days that constitute the  20% best/worst days of the year.

### 8.4.4 Box Plots and Q-Q Plots

Box and whisker plots of the projected extinction and deciview at Class I areas as well as absolute modeled extinction and deciview at all sites can be used to see changes in visibility across all concentrations, not just the 20% best/worst days.  Quantile-Quantile plots of the changes in frequency distribution projected using RRFs and absolute results from the model would also provide an indication of the changes in visibility from the current to future-year across all days.

# 9.0 DATA MANAGEMENT

Data management and data security procedures are critical components of the CENRAP regional fine particulate and haze modeling. Very large data files are used in each component of the modeling process, including processing of the meteorology data, emissions processing, and visibility modeling with CMAQ and CAMx. An annual simulation on the CENRAP 36-km domain requires approximately 2 Terabytes (Tb) of disk storage for CMAQ, with CAMx requiring less than half that storage need due to use of more efficient file formats (e.g., doesn't use 3D emission inputs). This chapter describes data management practices.

For all critical files we will maintain backup copies either on tapes or redundant disk systems. In addition, because model simulations will be performed separately by ENVIRON and UCR, each institution will maintain its own copy and backup of critical input and output files. Because there are differences in system configurations at each of the modeling centers (ENVIRON and UCR), the data backup and archiving are discussed separately for each center, below.

CMAQ and CAMx generate large output files of which most information is rarely used (for example, model output for layers other than the surface layer). A common practice is to extract layer 1 data from the 3D concentrations files for analysis and then deleted the 3-dimensional concentration output files. Whether this will be done for the CENRAP runs remains to be seen. In any event, we will permanently saved the script, executables and input data so that 3-D concentration files can be regenerated if needed.

To promote efficient, reliable communication among project participants, the modeling team will create several different listservs to aid in dissemination of information and as a primary means for distributing emissions and air quality modeling information to the CENRAP Project Manager, Administrative Project Manager and Modeling Workgroup.

## 9.1 Project Website

For the preliminary CENRAP modeling analysis performed by UCR and ENVIRON, a project webpage at UCR was used to publish and disseminate all project results. The CENRAP preliminary modeling webpage address is: http://pah.cert.ucr.edu/aqm/cenrap/index.shtml. We will establish a new webpage for the CENRAP 2002 annual modeling. The project webpage will be the primary mechanism for communicating results of model simulations and analysis to funding agencies and other interested parties.

## 9.2 Data Transfer

Data transfer among the modeling centers and between other CENRAP participants or contractors will be accomplished using a combination of email, ftp downloads and portable disk drives depending on the size of the data transfer. For data files smaller than a few MB email typically works well and is most efficient. For data files of less than about 500 MB file transfer protocol (ftp) is typically the fastest and most efficient method. ENVIRON and UCR each maintain webpages and ftp pages that can be used for exchanging data. In addition, each modeling center has several portable disk drives with both USB2 and firewire interfaces that can be FedEx among project participants to exchange large data sets. Portable disk drives range in size from 80 to 300 GB and are adequate for all large files data transfers. The approach described here has been used throughout the WRAP and VISTAS projects and has proven to be economical and efficient.

## 9.3 Data Backup and Archiving

Data backup and archiving will be performed at each of the modeling centers. Copies of critical project data will be maintained at each modeling center to provide redundant backup of key project data. In addition, each modeling center will perform backups of key project data to tape or redundant disk storage systems. Data storage and back up resources at each modeling center are described next.

### 9.3.1 ENVIRON

The ENVIRON modeling center has over 15,000 Gigabytes (>15 Terabytes, Tb) of disk storage available to the UNIX/Linux workstations. All of the workstations are networked together and are accessible from each employee's desktop PC. All workstations have CD-ROM drives and can access DLT, 4mm DAT and 8mm Exabyte tape drives for data backup and data transfer. ENVIRON can also create CDs (CD-R and CD-RW) and DVDs (DVD+ and DVD-) for data backup and distribution. For CENRAP 2002 annual modeling, most CMAQ and CAMx simulations would be performed on one of the 9 node Beowulf Linux Cluster that includes one master node and 8 processing nodes. Each node consists of two AMD Athlon 2600+ (or faster) processors. The master node has 2 Gb of memory and is connected to a 2.8 Tb RAID disk system. Each secondary processing node includes 1 Gb of memory. The ENVIRON Novato computing center also includes approximately 10 dual processor Linux workstations with processing speeds of 1700+ to 3000+. Three older Unix workstations are also available, SUN, DEC and SGI. The SMOKE modeling would likely be performed using a dual processor Linux box. The Linux computer systems are located in their own room with their own dedicated air conditioning (AC) system. The room includes a temperature sensitive power shut off device that will shut off the power to all computers in case the AC breaks down so that catastrophic failure due to too high temperatures does not occur. Backups are made on IDE disk drives that are removed from the computer and stored on a shelf to protect against power surges destroying the backup data.

### 9.3.2 UC Riverside

Data storage systems at UCR include more than 22 TB of disk space configured as RAID5 disk systems. All computers and disk systems are connected using high speed Gigabit Ethernet for efficient simulation and analysis of large datasets. To provide maximum data security the systems are located behind the UCR firewall and an additional firewall internally within the laboratory. A separate system is used for the project websites and ftp site to allow project data to be accessed thought the UCR T1 internet connection. The data backup/archiving system include 8mm tape drives and DLT and Super DLT auto loading cartridge system capable of performing unattended archive/backups of over 1 Tb (uncompressed). Key disk systems have hot-swappable hard drives with stand-by spare drives and redundant power supplies. Copies of critical project data are backed up to tape and to a redundant RAID5 disk system. The compute clusters and disk systems are located in a locked, secure room with a dedicated climate control system and with backup air conditioning. The laboratory has a full time systems administrator to perform system backups, maintenance and updates and Dr. Tonnesen's group includes a second full time systems administrator who performs weekly backups of critical data to tape.

# 10.0 DOCUMENTATION

This section describes the documentation that be provided during CENRAP annual modeling and the potential for modifications to this Modeling Protocol and QAPP that might become necessary as the study unfolds.

## 10.1 Planned Documentation

Documentation associated with the emissions and air quality modeling performed during the CENRAP annual modeling will include all relevant input data bases and scripts associated with the pre- and post-processing associated with model input development, model application, sensitivity and diagnostic analyses, and performance evaluations. PowerPoint presentations, technical memorandums, interim and final reports that describe the methodologies and results of the model performance evaluation, model intercomparison, and visibility assessment will be provided. Table 10-1 below lists the current schedule of deliverables under the CENRAP 2002 annual modeling and analysis study.

**Table 10-1.** Current list of deliverables and schedule under the CENRAP 2002 annual modeling study.

| Deliverable | Deliverable Due Date |
|---|---|
| Task 1. Develop QAPP. | |
| Draft due to CENRAP | November 15, 2004 |
| Final due to CENRAP | Within 1 week of comments |
| Task 2. Develop Modeling Protocol/Work plan. | |
| Draft due to CENRAP | November 15, 2004 |
| Final due to CENRAP | Within 1 week of comments |
| Task 3. Develop model-ready meteorology inputs. | |
| Presentation (in electronic format) due to CENRAP | December 20, 2004 |
| Model-ready 2002 meteorology inputs completed | December 31, 2004 |
| Task 4. Develop base case model-ready emission inputs. | |
| Monthly progress reports due to CENRAP | 2nd week of following month |
| Technical memorandum due to CENRAP | February 21, 2005 |
| Presentation (in electronic format) due to CENRAP | February 21, 2005 |
| Model-ready emissions completed | January 31, 2005 |
| Task 5. Perform 2002 base case modeling evaluation. | |
| Monthly progress reports due to CENRAP | 2nd week of following month |
| Recommendation on model configuration due to CENRAP | March 15, 2005 |
| Draft report due to CENRAP | April 30, 2005 |
| Base case model simulations completed | April 30, 2005 |
| Task 6. Develop future case model-ready emission inputs. | |
| Monthly progress report due to CENRAP | 2nd week of following month |
| Technical memorandum due to CENRAP | April 30, 2005 |
| 2018 Model-ready emissions completed | April 30, 2005 |
| Task7. Perform 2018 modeling evaluation, sensitivities, and control strategies. | |
| Monthly progress report due to CENRAP | 2nd week of following month |
| Draft sensitivity and source apportionment report due to | August 31, 2005 |

| Deliverable | Deliverable Due Date |
|---|---|
| CENRAP | |
| Final sensitivity and source apportionment report due to CENRAP | Within 1 week of comments |
| 2018 model simulations completed | April 30, 2006 |
| Task 8. Reports and recommendations | |
| Draft report due to CENRAP | December 31, 2005 |
| Final report due to CENRAP | April 30, 2006 |

In addition, copies of all scripts, reports and documentation provided to the CENRAP Modeling Workgroup will be maintained on the project website. A final report summarizing all aspects of the project will also be provided.

Reporting on each task in the 2002 CENRAP modeling will consist of documentation of the data sources, methods, results, and findings. Individual task deliverables shall reflect any changes and revisions that occur over duration of study. At the completion of the CENRAP annual modeling, a draft final report will be prepared that details, documents and summarizes the results. This documentation will conform to the recommendations set forth in EPA's "Draft Guidance for Demonstrating Attainment of Air Quality Goals for $PM_{2.5}$ and Regional Haze (Jan '01)" and any subsequent versions. The final report will contain: (a) an executive summary abstract that provides a brief overview and summary of the modeling effort, emissions and air quality models used, model configuration, model performance evaluation overview and results, and rationale for the selected configuration, (b) technical detail covering all relevant aspects of the CENRAP 2002 emissions and PM grid modeling, and (c) a discussion on data accessibility and availability for review by the public. The report will be provided in electronic form, (e.g., Word 2000 and pdf formats) and shall be submitted to the CENRAP Project Manager and also posted on the project web site.

## 10.2 Procedures for Updating Modeling Protocol and QA Plan

One of the underlying realizations stemming from the preliminary CENRAP modeling activities was the awareness that the science of "One-Atmosphere" PM/regional haze modeling is advancing very rapidly. Part of this stems from the parallel activities being carried out by the RPOs and EPA; some if it is due to other ongoing 8-hr ozone and PM modeling studies being performed by various states. In addition, EPA is in the process of revising its $PM_{2.5}$ and regional haze guidance documents. Collectively, it is quite likely that there will be new opportunities to strengthen the modeling algorithms, input data sets, and evaluation procedures throughout the duration of the CENRAP annual modeling study. Moreover, when EPA updates the regional haze modeling guidance that may also trigger a need to update the CENRAP modeling approach and Modeling Protocol. Given the ongoing model refinement activities and the need for strong coordination with other CENRAP contractors, it may be necessary to modify certain aspects of this Modeling Protocol. In this event, modification will be made in consultation with the CENRAP Project Manager and the revised protocol will be submitted to the CENRAP Modeling Workgroup for approval.

# REFERENCES


Adelman, Z. 2004. Quality Assurance Protocol: WRAP RMC Emissions Modeling with SMOKE, Prepared for WRAP Modeling Forum, January.

Anthes, R.A. 1977. "A Cumulus Parameterization Scheme Utilizing a One-Dimensional Cloud Model", *Mon. Wea. Rev.,* Vol. 105, pp. 270-286.

Anthes, R.A. and T.T. Warner. 1978. "The Development of Mesoscale Models Suitable for Air Pollution and Other Mesometeorological Studies", *Mon. Wea. Rev.*, Vol. 106, pp, 1045-1078.

Arnold, J.R.; R.L. Dennis, and G.S. Tonnesen. 1998. Advanced techniques for evaluating Eulerian air quality models: background and methodology. In: Preprints of the 10th Joint Conference on the Applications of Air Pollution Meteorology with the Air & Waste Management Association, January 11-16, 1998, Phoenix, Arizona. American Meteorological Society, Boston, Massachusetts, paper no. 1.1, pp. 1-5.

Arnold, J.R., R.L. Dennis, and G.S. Tonnesen. 2003. "Diagnostic Evaluation of Numerical Air Quality Models with Specialized Ambient Observations: Testing the Community Multiscale Air Quality Modeling System (CMAQ) at Selected SOS 95 Ground Sites", *Atmos. Environ.*, Vol. 37, pp. 1185-1198.

ARS. 2002a. *Analysis of IMPROVE and SEARCH Speciated $PM_{2.5}$ Data.* Air Resource Specialists, Inc. Report to VISTAS, August 28, 2002. http://www.vistas-sesarm.org.

ARS. 2002b. *Analysis of Alternative Methods.* Air Resource Specialists, Inc. Report to VISTAS, August 28, 2002. http://www.vistas-sesarm.org.

ARS. 2002c. *Review of Selected Factors Responsible for High and Low Extinction.* Air Resource Specialists, Inc. Report to VISTAS, October 14, 2002. http://www.vistas-sesarm.org.

Baker, K. 2004. Presentations on latest MRPO modeling results at November 11, 2004 MRPO meeting in Chicago, Illinois. Available at: http://www.ladco.org/tech/photo/photochemical.html.

Barchet, W.R., and R.L. Dennis. 1990. "ANAPAP Model Evaluation, Volume 1: Protocol", prepared for the U.S. Environmental Protection Agency, prepared by Battelle Pacific Northwest Laboratories, Richland, WA.

Boylan, J. W., 2004. "Calculating Statistics: Concentration Related Performance Goals", paper presented at the EPA PM Model Performance Workshop, Chapel Hill, NC. 11 February.

Byun, D.W., and J.K.S. Ching. 1999. "Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System", EPA/600/R-99/030.

Carter, W.P.L. 1999. Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment, Draft report to the California Air Resources Board, Contracts 92-329 and 95-308, 9/13/99.

CENRAP. 2003. CENRAP Long Range Plan. Central States Regional Air Planning Association, Oklahoma City, Oklahoma. October 29.

Chang, J.S., R.A. Brost, I.S.A. Isakson, S. Madronic, P. Middleton, W.R. Stockwell and C.J. Walcek. 1987. A three-dimensional Eulerian acid deposition model: Physical concepts and formulation. *J. Geophys. Res.,* 92, 14681-14700.

Chang, J. S., et al. 1997. "The SARMAP Air Quality Model", prepared for the California Air Resources Board, prepared by the State University of New York, Albany, New York.

Coe D. and S. Reid. 2003. "Research and Development of Ammonia Emission Inventories for the Central States regional Air Planning Association". Final Report. Sonoma Technology, Inc. (STI). October 30.

Coella, P., and P. L., Woodward. 1984. "The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations", *J. Comp. Physics*, Vol. 54, pp. 174-201.

Dennis, R. L., et al. 1990. "Evaluation of Regional Acid Deposition Models", State-of-Science/Technology Report No. 5, National Acid Precipitation Assessment Program, Washington, D.C.

Dentener, F. J., and P. J. Crutzen. 1993. Reaction of $N_2O_5$ on tropospheric aerosols: Impact on global distributions of $NO_x$, $O_3$, and OH, *J. Geophys. Res., 98*, 7149-7163.

Donahue, N.M., M.K. Dubey, R. Mohrschladt, K.L. Demerjian, and J.G. Anderson. 1997. High-pressure flow study of the reactions $OH+NO_x \rightarrow HONO_x$: Errors in the falloff region. *J. Geophys. Res.* **102**, 6159-6168.

Dudhia, J. 1993. "A Non-hydrostatic Version of the Penn State/NCAR Mesoscale Model: Validation Tests and Simulation of an Atlantic Cyclone and Cold Front", *Mon. Wea. Rev.*, Vol. 121. pp. 1493-1513.

Edgerton, E. S., S*ources of PM2.5 Carbon in the Southeast US*, presentation at the MARAMA Science Meeting, Baltimore, MD, January 22, 2003. http://www.marama.org.

ENVIRON, Alpine Geophysics, UC Riverside and UC Davis. 2003a. "VISTAS Emissions and Air Quality Modeling – Phase I Task 4a/b Report: Review of Model Sensitivity Simulations and Recommendations of Initial CMAQ Model Configurations and Sensitivity Tests." Revised Draft July 25.

ENVIRON, Alpine Geophysics , UC Riverside and UC Davis. 2003b. "VISTAS Emissions and Air Quality Modeling – Phase I Task 3 Report: Review and Assessment of Available Ambient Air Quality Data to Support Modeling and Modeling Performance Evaluation for the Three VISTAS Phase I Episodes." Revised Draft July 22.

ENVIRON. 2002. "User's Guide Comprehensive Air Quality Model With Extensions (CAMx) Version 3.10." ENVIRON International Corporation, Novato, California (available at www.camx.com) April.

ENVIRON. 2003a. "Development of an Advanced Photochemical Model for Particulate Matter: PMCAMx." ENVIRON International Corporation, Novato, CA. Prepared for Coordinating Research Council, Inc. Project A-30 (available at www.crcao.com).

ENVIRON, 2003b. "VISTAS Emissions and Air Quality Modeling – Phase I Task 2 Report: Recommended Model Configurations and Evaluation Methodology for Phase I Modeling." Prepared by ENVIRON International Corporation, Alpine Geophysics, LLC and University of California at Riverside. Novato, California. (available at: http://pah.cert.ucr.edu/vistas/docs.shtml). August 4.

ENVIRON, 2003c. "VISTAS Emissions and Air Quality Modeling – Phase I Task 3 Report: Review and Assessment of Available Ambient Air Quality Modeling and Performance Evaluation for the Three VISTAS Phase I Episodes." Prepared by ENVIRON International Corporation, Alpine Geophysics, LLC and University of California at Riverside. Novato, California. (available at: http://pah.cert.ucr.edu/vistas/docs.shtml). July 22.

ENVIRON, 2003d. "VISTAS Emissions and Air Quality Modeling – Phase I Task 4a/b Report: Review of Model Sensitivity Simulations and Recommendations of Initial CMAQ Model Configurations and Sensitivity Tests." Prepared by ENVIRON International Corporation, Alpine Geophysics, LLC and University of California at Riverside. Novato, California. (available at: http://pah.cert.ucr.edu/vistas/docs.shtml). July 25.

EPA, 1991. "Guidance for Regulatory Application of the Urban Airshed Model (UAM), "Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C.

EPA, 1999. "Draft Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hr Ozone NAAQS". Draft (May 1999), U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, N.C

EPA. 2001. "Guidance for Demonstrating Attainment of Air Quality Goals for $PM_{2.5}$ and Regional Haze", Draft Report, U.S. Environmental Protection Agency, Research Triangle Park, NC.

EPA. 2003a. "Guidance for Tracking Progress Under the Regional Haze Rule", U.S. Environmental Protection Agency, Research Triangle Park, NC.

EPA. 2003b. "Guidance for Estimating Natural Visibility Conditions Under the Regional Haze Rule", U.S. Environmental Protection Agency, Research Triangle Park, NC.

Fahey, K.M., and S.N. Pandis. 2001. Optimizing model performance: variable size reduction in cloud chemistry modeling. *Atmos. Environ*. Vol. 35, pp. 4471-4478.

FLAG. 2000. Federal Land Manager's Air Quality Related Values Workgroup (FLAG) Phase I Report. (http://www2.nature.nps.gov/ard/flagfiee/AL.doc).

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

Gipson, G, and J. Young. 1999. Process Analysis, in *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*, Chapter 16, U.S. EPA, Office of Research and Development, Washington, DC 20460.

Green M., J. Watson, H. Kuhns, R. Morris, A. Pollack, P. Fields and M. Wolf. 2002. "A Scoping Study for Haze in the CENRAP Region". Final Report. Desert Research Institute (DRI), EBVIRON and ERG. December 18.

Grell, G. A., J. Dudhia, and D. R. Stauffer. 1994. "A Description of the Fifth Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note, NCAR TN-398-STR, 138 pp.

Hanna, S.R., A.G. Russell, J. Wilkinson, and J. Vukovich. 2002. Review of BEIS3 Formulation and Consequences Relative to Air Quality Standards: Estimation of Uncertainties in BEIS3 Emissions Outputs. EPRI Technical Report 1005159, EPRI, 3412 Hillview Ave., Palo Alto, CA 94304.

Hanna, S.R., A.G. Russell, J. Wilkinson, and J. Vukovich. 2003. Review of BEIS3 Formulation and Consequences Relative to Air Quality Standards: Estimation of Effects of Uncertainties in BEIS3 Emissions on Uncertainties in Ozone Predictions by Chemical Transport Models, EPRI, Palo Alto, CA: 2002 (report is currently under review by EPRI).

Hanna, S. R. 1994. "Mesoscale Meteorological Model Evaluation Techniques with Emphasis on Needs of Air Quality Models", in Mesoscale Modeling of the Atmosphere, R. A. Pielke and R. P. Pierce (Eds.), American Meteorological Society, Boston, MA.

Hermann, H. et al., 2000. manuscript submitted to *J. Geophys. Res.*

IPCC. 2001. "Climate Change 2001: The Scientific Basis -- Contribution of Working Group I to the Third Assessment Report of IPCC." Intergovernmental Panel on Climate Change. ISBN: 0521014956.

Jacob, D. J. 1999. Heterogeneous chemistry and tropospheric ozone. *Atmos. Environ.*, special issue, 1999 NARSTO Assessment.

Johnson, M. 2004. "Annual 2002 MM5 v363 Simulation Evaluation." Iowa Department of natural Resources, Air Quality Bureau. November. (Available at www.ladco.org/tech/photo/present/mm5v363_eval.pdf).

Khasibatla, P. et al. 1997. Impact of inert organic nitrate formation on ground-level ozone in a regional air quality model using the Carbon Bond Mechanism 4. *Geophys. Res. Letters*, **24**, 3205-3208.

Liu, G., C. Hogrefe, and S.T. Rao. 2003. "Evaluating the Performance of Regional-Scale Meteorological Models: Effects of Clouds Simulation on Temperature Prediction", *Atmos. Environ.*, Vol. 36, pp. 1691-1705.

Malm, W.C.; Pitchford, M.L; Scruggs, M.; Sisler, J.F.; Ames, R.; Cepeland, S.; Gebhart, K.; Day, D.E. 2000. *Spatial and Seasonal Patterns and Temporal Variability of Haze and Its Constituents in the United States. Report III.* Cooperative Institute for Research in the Atmosphere, May.

Matthias-Maser, S. and R. Jaenicke. 1995. "The size distribution of primary biological aerosol particles with radii >0.2 mm in an urban-rural influenced region." *Atmos. Res*., 39, 279-286.

McNally, D.E., and T.W. Tesche. 1994. "MAPS2.3 User's Guide", Alpine Geophysics, LLC, Golden, CO.

McNally, D. E. 1997. "Development of Methodology for Mapping MM5 Fields onto Arbitrary Eulerian Photochemical Air Quality Simulation Models (PAQSM)", Alpine Geophysics, LLC, Arvada, CO.

Morris, R.E., B. Koo, S. Lau, T.W. Tesche, D. McNally, C. Loomis, G. Stella, G. Tonnesen and Z. Wang. 2004. "VISTAS Emissions and Air Quality Modeling – Phase I Task 4cd Report: Model Performance Evaluation and Model Sensitivity Tests for Three Phase I Episodes", prepared for the VISTAS Technical Analysis Committee, prepared by ENVIRON International Corporation, Alpine Geophysics, LLC, and the University of California, Riverside (CE-CERT).

Morris, R.E., D. McNally, T. Tesche, G. Tonnesen, J. Boylan and P. Brewer. 2004. "Regional Haze Modeling over the VISTAS States: Preliminary Verification of Models-3/CMAQ for the 2002 Annual Period." Presented at AWMA Visibility Conference, Asheville, North Carolina. October.

NARSTO (1999) NARSTO Quality Planning Handbook, Appendix A: Quality Integrated Work Plan Template for Monitoring and Measurement Research and Development Projects. www.cgenv.com/Narsto.

Pielke, R. A., and M. Uliasz. 1998. "Use of Meteorological Models as Input to Regional and Mesoscale Air Quality Models C Limitations and Strengths", *Atmos. Environ*. Vol. 32, No. 8, pp. 1455-1466.

Pilinis, C. and J. H. Seinfeld. 1987. "Continued Development of a General Equilibrium Model for Inorganic Multicomponent Atmospheric Aerosols", *Atmos. Environ*. Vol.21, pp.2453.

Pilinis, C., K.P. Capaldo, A. Nenes, and S.N. Pandis. 2000. "MADM - A new multicomponent aerosol dynamics model." *Aerosol Sci. Tech*., Vol. 32(5), pp. 482-502. 2000.

Pitchford M.L., I. Tombach, M. Barna, K. Gebhart, M. Green, E. Knipping, N. Kumar, W. Malm, B. Pun, N. Schichtel and C. Seigneur. 2004. "Big Bend Aerosol and Visibility Study". Final Report. September.

Pleim, J.E. and J.S. Chang. 1992. A non-local closure model for vertical mixing in the convective boundary layer. *Atmos. Env*., **26A**, pp. 965-981.

Pun B., S-Y. Chen and C. Seigneur. 2004. "CENRAP Model Simulations and Performance Evaluations". Atmospheric and Environmental Research, Inc. (AER). August.

Reid S., S. Brown, D. Sullivan, H. Arkinson, T. Funk, and P. Stiefer. 2004a. "Research and Development of Planned Burning Emission Inventories for the Central States Regional Air Planning Association". Final Report. Sonoma Technology, Inc. (STI). July 30.

Reid S., D. Sullivan, B. Penfold, T. Funk, T. Tamura, P. Stiefer, S. Raffuse and H. Arkinson. 2004. "Emission Inventory Development for Mobile Sources and Agricultural Dust Sources form the Central States". Final Report. Sonoma Technology, Inc. (STI). September 22.

Russell, A.G., and R.L. Dennis. 2000. "NARSTO Critical Review of Photochemical Models and Modeling", *Atmos. Environ*., Vol. 34, No. 12-14, pp. 2283-2324.

Schere, K.L., and R.A. Wayland. 1989. "EPA Regional Oxidant Model (ROM2.0): Evaluation on 1980 NEROS Data Bases", EPA/600/3-80/057, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

Seaman, N.L. 1995. "Status of Meteorological Pre-Processors for Air Quality Modeling", *International Conf. On Particulate Matter*, Air and Waste Mgt. Assn., Pittsburgh, PA.

Seaman, N.L. 1995. "Status of Meteorological Pre-Processors for Air Quality Modeling", *International Conf. On Particulate Matter*, Air and Waste Mgt. Assn., Pittsburgh, PA.

Seaman, N.L. 1996. "Study of Meteorological Variables Needed in Air-Quality Modeling', Annual Progress Report prepared for the Coordinating Research Council (CRC), Interim Report, Project A-11, prepared by the Department of Meteorology, Penn State University, State College, PA.

Seaman, N.L. 2000. "Meteorological Modeling for Air Quality Assessments", *Atmos. Environ*., Vol. 34, No. 12-14, 2231-2260.

Seaman, N. L., and D. R. Stauffer. 1996. "SARMAP Meteorological Model Final Report", prepared for the San Joaquin Valleywide Air Pollution Study Agency, prepared by the Department of Meteorology, Pennsylvania State University, University Park, PA.

Seaman, N.L., and S.A. Michelson. 1998. "Mesoscale Meteorological Structure of a High-Ozone Episode During the 1995 NARSTO-Northeast Study", *J. Appl. Meteo.*, (submitted).

Seaman, N.L., D.R. Stauffer, and T.W. Tesche. 1992. "The SARMAP Meteorological Model: A Four-Dimensional Data Assimilation Technique Used to Simulate Mesobeta-Scale Meteorology During a High-Ozone Episode in California", International Specialty Conference on Tropospheric Ozone Nonattainment and Design Value Issues, U.S. EPA/AWMA, 27-30 October, Boston, MA.

Seaman, N.L., D.R. Stauffer, and L.M. Lario. 1995. "A MultiScale Four-Dimensional Data Assimilation System Applied to the San Joaquin Valley During SARMAP. Part I: Modeling Design and Basic Performance Characteristics", *J. Appl. Meteo.*, Vol. 34, pp. 1739-1761.

Seaman, N.L., D.R. Stauffer, and D.E. McNally. 1996a. "Application of the MM5-FDDA Meteorological Model to the Southern California SCAQS-1997 Domain: Preliminary Test Using the SCAQS August 1987 Case", Ninth Joint Conference on Applications of Air Pollution Meteorology, American Meteorological Society, 28 January-2 February, Atlanta, GA.

Seinfeld, J.H. and S.N. Pandis. 1998. Atmospheric Chemistry and Physics: From Air Pollution to Global Change. Wiley, New York.

Seigneur, C., et al. 2000. "Guidance for the Performance Evaluation of Three-Dimensional Air Quality Modeling Systems for Particulate Matter and Visibility", *J. Air & Wast Manage. Assoc.* Vol. 50, pp. 588-599.

Simoneit, B.R.T., J.N. Cardoso and N. Robinson. 1990. "An assessment of the origin and composition of higher molecular weight organic matter in aerosols over Amazonia." *Chemosphere*, 21, 1285-1301.

Solomon, P.S., T. Klamser-Williams, P. Egeghy, D. Crumpler and J. Rice. 2004. "STN/IMPROVE Comparison Study Preliminary Results". Presented at PM Model Performance Workshop. Chapel Hill, NC. February 10.

Stockwell, W.R., P. Middleton, J.S. Chang, and X. Tang. 1990. The second generation Regional Acid Deposition Model chemical mechanism for regional air quality modeling. *J. Geophys. Res.,* **95**, 16,343-16,367.

Stockwell, W.R., P. Middleton, J.S. Chang, and X. Tang, X. 1997. The second generation Regional Atmospheric Chemistry Mechanism regional air quality modeling. *J. Geophys. Res.,* **102**.

Strader, R., F.W. Lurmann and S.N. Pandis. 1999. "Evaluation of secondary organic aerosol formation in winter." *Atmos. Environ*. Vol. 33, pp. 4849-4864.

Strait, R., S. Roe and J. Vukovich. 2004. "2002 Base Case Modeling Inventory – Work Plan." E.H. Pechan and Associates, Inc. Durham, North Carolina. July 30.

Tang, Y. 2002, "A Case Study of Nesting Simulation for the Southern Oxidants Study 1999 at Nashville", *Atmos. Environ.*, Vol. 36, pp. 1691-1705.

Tesche, T. W., et al., 1990. "Improved Treatment of Procedures for Evaluating Photochemical Models", prepared for the California Air Resources Board, prepared by Alpine Geophysics, Crested Butte, CO. Contract No. A832-103.

Tesche, T.W., and D.E. McNally. 1993a. "Operational Evaluation of the CAL-RAMS Meteorological Model for LMOS Episode 1: 26-28 June 1991", prepared for the Lake Michigan Air Directors Consortium, prepared by Alpine Geophysics, LLC, Crested Butte, CO.

Tesche, T.W., and D.E. McNally. 1993b. "Operational Evaluation of the CAL-RAMS Meteorological Model for LMOS Episode 2: 17-19 July 1991", prepared for the Lake Michigan Air Directors Consortium, prepared by Alpine Geophysics, LLC, Crested Butte, CO.

Tesche, T.W., and D.E. McNally. 1993c. "Operational Evaluation of the CAL-RAMS Meteorological Model for LMOS Episode 3: 25-26 August 1991", prepared for the Lake Michigan Air Directors Consortium, prepared by Alpine Geophysics, LLC, Crested Butte, CO.

Tesche, T.W., and D.E. McNally. 1993d. "Operational Evaluation of the CAL-RAMS Meteorological Model for LMOS Episode 4: 20-21 June 1991", prepared for the Lake Michigan Air Directors Consortium, prepared by Alpine Geophysics, LLC, Crested Butte, CO.

Tesche, T. 1994. Evaluation Procedures for Regional Emissions, Meteorological, and Photochemical Models. Presented at the 86th Annual Meeting for the Air and Waste Management Association, 14-18 June, Denver, CO.

Tesche, T.W. and D.E. McNally. 1996a. "Superregional Ozone Modeling and Analysis Study C Phase I: Work Element 3: Assessment of the OTAG Data Sets -- Task 2 Technical Memorandum: Review of the OTAG Meteorological Inputs and Outputs", prepared for the Midwest Ozone Group, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W. and D.E. McNally. 1996b. "Superregional Ozone Modeling and Analysis Study C Phase II: Work Element 5 Technical Report: Comparative Evaluation of the MM5 and RAMS Models for the July 1991 OTAG Episode", prepared for the Midwest Ozone Group, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W., and D.E. McNally. 1997a. "Superregional Ozone Modeling and Analysis Study C Final Report: Assessment of the Reliability of the OTAG Modeling System", prepared for the Midwest Ozone Group, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W., and D.E. McNally. 1997b. "Methodology for Climatological Analysis of Multi-Year Aerometric Data to Support the Development of the Breton Conceptual Model and Field Program Design", prepared for Walk Heydel Environmental, prepared by Alpine

Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W., and D.E. McNally. 1997c. "The Use of the San Joaquin Valley Meteorological Model in Preparation of a Field Program in the South Coast Air Basin and Surrounding Regions of Southern California: Volume I: Final MM5 Evaluation for the 3-6 August 1990 SARMAP Episode", prepared for the California Air Resources Board, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W. and D.E. McNally. 1999. "Comparative Evaluation of the MM5 and RAMS3c Prognostic Meteorological Models Over the Midwestern U.S. for Two 1999 LMOS Intensive Measurement Episodes", prepared for the Coordinating Research Council, Project A-25 draft Final Report, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W., and D.E. McNally. 2001. "Evaluation of CAMx and Models-3/CMAQ Over the Lower Lake Michigan Region with Inputs from the RAMS3c and MM5 Models", prepared for the Coordinating Research Council, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche, T.W., and H.E. Jeffries. 2002. "Findings From BCCA-AG Science Studies", prepared for the Business Coalition for Clean Air- Appeals Group, (2002), prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Timin, B. 2002. "$PM^{-}_{2.5}$ and Regional Haze Modeling Guidance". Prepared by the U.S. EPA/OAQPS. April 24.

Timin, B. 2004. "$PM_{2.5}$ Model Performance Evaluation: Purpose and Goals". Presented at PM Model Evaluation Workshop, Chapel Hill, NC. February.

Tonnesen, G.S. 1999. Effects of uncertainty in the reaction of the hydroxyl radical with nitrogen dioxide on model simulated ozone control strategies. *Atmos. Environ.* **33***:*1587-1598.

Tonnesen, G.S. and R.L. Dennis. 2000. Analysis of radical propagation efficiency to assess ozone sensitivity to hydrocarbons and $NO_x$. Part 1: Local indicators of odd oxygen production sensitivity. *J. Geophys. Res.*, **105**, 9213–9225.

Tonnesen, G., R.E. Morris, J. Vimont, M. Uhl, K. Briggs and T. Moore. 2003. "The WRAP Regional Modeling Center – Application and Evaluation of Regional Visibility Models." Presented at AWMA Annual Meeting and Exhibition. June.

Tonnesen G. and R. Morris. 2004. "Evaluation of Existing Regional Haze Modeling Results for the CENRAP Region". UCR and ENVIRON. July.

Watson, J.G. and J.C. Chow. 1999. Reconciling Urban Fugitive Dust Emissions Inventory and Ambient Source Contribution Estimates: Summary of Current Knowledge and Needed Research. Document No. 61110.4D2, Desert Research Institute, Reno NV, September 3.

Western Regional Air Partnership (2003) Western Regional Air Partnership 2004 Work Plan, October.

Western Regional Air Partnership (2003a) Strategic Plan 2003-2008 of the Western Regional Air Partnership, September.

Xiu, A. and J.E. Pleim (2000) Development of a land surface model. Part I: application in a mesoscale meteorology model. *J. App. Met.*, **40**, pp. 192-209.

Zaveri, R.A and L.K. Peters. 1999. A new lumped structure photochemical mechanism for large scale applications, *J. Geophys. Res*., 30,387-30,415

Zheng,, M.; Cass, G.R.; Schauer, J.J.; Edgerton, E.S. *Environ. Sci. Tech.* 2002, 36, 2361-2371.