Regional Air Quality Model Evaluation: Framework and Tools

Alice Gilliland, Wyat Appel, Kristen Foley, Robert Gilliam, Robert Pinder, S.T. Rao, Kenneth Schere, and Jenise Swall

2009 AMAD Peer Review: Session 2
Model Evaluation: Establishing Model's Credibility

Environmental Issue

Atmospheric chemical transport models, such as the USEPA Community Multiscale Air Quality (CMAQ) model, are used for air quality management at the local, state, and federal levels. CMAQ is used by the Regional Planning Organizations (RPOs), by the state agencies for designing their emission control policies as part of State Implementation Plans, and by the USEPA Office of Air Quality Planning and Standards (OAQPS) for federal rulemaking such as the recently vacated Clean Air Interstate Rule (CAIR).

As the developers of the CMAQ model, the Division also has the primary responsibility to conduct comprehensive evaluation studies to rigorously assess model performance in simulating the spatio-temporal features embedded in the air quality observations.

Research Objectives

❖ The evaluation program has strived for comprehensive assessment, considering the quality of the model inputs (e.g., emissions, meteorology) as well as various processes simulated within the CMAQ model.

❖ A fundamental goal of the CMAQ evaluation program has been to conduct evaluation studies that not only characterize model performance but also identify what model improvements (inputs or processes) are needed.

❖ Furthermore, evaluating an air quality model's response to emission changes is central to how the model is used for air quality management.

❖ An additional goal to expanding model evaluation is to develop uncertainty estimates to assist the air quality management process. As the CMAQ evaluation program was developed over the past five years, the overarching objective was to incorporate all of these aspects into the program. To describe and coordinate the team's research efforts, a **Model Evaluation Framework** was developed, as discussed in the next section.

❖ Another objective of the CMAQ evaluation program has been to encourage more advanced model evaluations by the user community. The **Atmospheric Model Evaluation Tool (AMET)** was developed over the past four years and the first public release was in February 2008.

Model Evaluation Framework

In the Model Evaluation Framework, we refer to four types of model evaluation: **Operational, Diagnostic, Dynamic and Probabilistic evaluation** approaches. Since these are not necessarily mutually exclusive, research studies often incorporate aspects from more than one category of evaluation.

❖ Operational evaluation is a fundamental first phase of any model evaluation study. Operational evaluation, as defined here, is a comparison of model predicted and observed concentrations of the end-point pollutant(s) of interest in a general sense. The time and spatial scale(s) of the evaluation can vary to provide more insight into the results. For more examples of analysis approaches for operational evaluation of CMAQ O_3 and $PM_{2.5}$ predictions, see Figure 1, the text box (right column) on the Atmospheric Model Evaluation Tool, and Appel et al. (2007, 2008). Operational evaluation is also an essential part of the National Air Quality Forecasting System. See Poster 2.2 for more details.

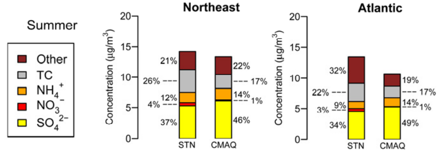

Figure 1. An example comparison of $PM_{2.5}$ species from a CMAQ simulation and the Speciated Trends Network (STN) from Appel et al. (2008).

Recently, more advanced statistical methods have been developed to aid in evaluation by making the best use of the limited monitoring data available, accounting for the differences between point-based measurements (monitors) and grid cell averages (model output), and assessing the model output for grid cells in which no monitors are located (Swall and Foley, in press; Irwin et al., 2008; Swall and Davis, 2006; Davis and Swall, 2006) (e.g. Figure 2).

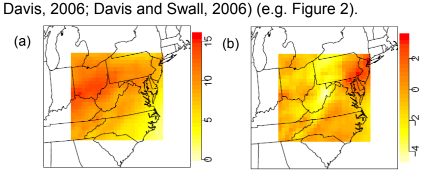

Figure 2. (a) Mean CMAQ sulfate concentration (µg/m³) during July 17 – August 13, 2001 and (b) predicted sulfate from a Bayesian statistical model based on observations from STN during the same time period (Swall and Davis, 2006).

❖ Diagnostic evaluation investigates the processes and input drivers that affect model performance (see Session 1 for many examples of diagnostic evaluation). A unique type of diagnostic evaluation discussed in this session is that of instrumented models to track the contributions and sensitivities of emissions, boundary conditions, transport, etc., and inverse modeling for emission evaluation. (See Posters 2.2 and 2.3). Diagnostic studies often rely on specialized datasets that include many chemical species that are not standard in large networks.

❖ Dynamic evaluation focuses on assessing the model's air quality response to changes in emissions and meteorology (Gilliland et al., 2008; Godowitch et al., 2008), which is central to an air quality model's application in air quality management (see Poster 2.5). Figure 3 summarizes the type of information that is demonstrated from dynamic evaluation, where the model's ozone response to the NOx State Implementation Plan (SIP) Call is evaluated.

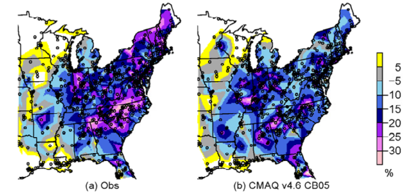

Figure 3. Example of dynamic evaluation of an air quality model-predicted change in ozone concentrations from summer 2002 to 2005 from Gilliland et al. (2008). The results illustrate the relative change in ozone when comparing the ≥95th percentile daily 8-hour maximum levels from the two summers.

❖ Probabilistic evaluation is challenging, given the deterministic nature of models like CMAQ, and the lack of data about the uncertainties in the emission inputs, meteorological inputs, and the compounding effect of numerous uncertainties. To explore ways to estimate uncertainty, ensemble modeling studies are being conducted using CMAQ, instrumented CMAQ-Decoupled Direct Method (CMAQ-DDM), and Bayesian Model Averaging techniques (see Poster 2.6 and Figure 4).

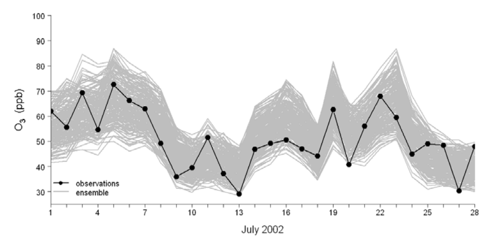

Figure 4. Time series of daily maximum 8-hour O3 concentrations (ppb) for July 2002 at a monitoring site located in the Birmingham, Alabama metropolitan area. Gray lines are results from individual members of a 200-member CMAQ model ensemble; black line/symbols are observed data from the monitor. From Pinder et al. (in review).

Atmospheric Model Evaluation Tool

Since operational model evaluation is a fundamental step for initiating comprehensive model evaluation, it is important to have the right analysis and graphical tools. The Atmospheric Model Evaluation Tool (AMET) was developed by the Division and released as a Beta version in February 2008. Framed around a relational database and the R (http://www.R-project.org) statistical package, AMET provides graphical and statistical evaluation results for air quality and meteorological model output. The intent is that additional R routines can be added by members of the Division and the research community as AMET development continues. Below are a few examples of AMET graphs. (http://www.cmascenter.org /conference/2008/slides/appel_amet_cmas08.ppt)

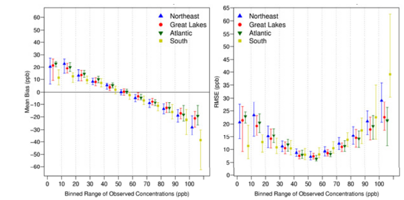

Figure 5. Median and inner-quartile ranges of mean bias and RMSE for daily maximum 8-hr ozone for four geographic regions in the eastern United States.

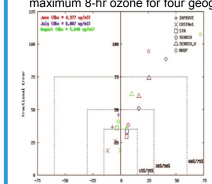

Figure 6. A "soccer goal" plot showing fractional bias vs fractional error for CMAQ sulfate predictions versus observational networks.

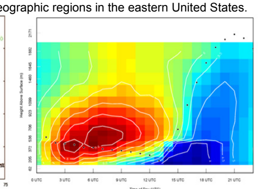

Figure 7. Predicted mean model wind speed. Warmer colors represent faster wind speeds, and cooler colors slower wind speeds. The points represent the wind profiler data.

Impact

The Model Evaluation Framework has been presented at an international conference (Gilliland et al., 2008) and in a review article currently being peer-reviewed (Dennis et al., 2008, in review) to provide leadership to the research community on conducting comprehensive evaluation programs.

The EPA Office of Air Quality Planning and Standards (OAQPS) has been using the AMET since 2006. AMET was publicly released in February 2008, and more than 200 users have downloaded AMET to date.

External Collaborators

Christian Hogrefe, NY Department of Environmental Conservation; Steve Porter, University of Idaho; Sharon Phillips, Office of Air Quality Planning and Standards; Jerry Davis, NC State University; John Irwin, Irwin and Associates