AQMEII: A New International Initiative on Air Quality Model Evaluation

S. Trivikrama Rao[1], Kenneth Schere[1], Stefano Galmarini[2], Douw Steyn[3]

[1]Atmospheric Modeling and Analysis Division, National Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, NC, USA

[2]Joint Research Centre, Institute for Environment and Sustainability – European Commission, Ispra, ITALY

[3]The University of British Columbia, Vancouver, BC, CANADA

Abstract We provide a conceptual view of the process of evaluating regional-scale three-dimensional numerical photochemical air quality modeling systems, based on an examination of existing approaches to the evaluation of such systems as they are currently used in a variety of applications. A framework for model evaluation is introduced to provide a context for the evaluation process. The objectives of the model evaluation process include: determining the suitability of a modeling system for a specific application; distinguishing between the performance of different models through confidence-testing of model results; and guiding further model development. The evaluation framework includes methods for *operational, diagnostic, dynamic, and probabilistic* model evaluation. Also discussed is a new effort, the Air Quality Model Evaluation International Initiative (AQMEII), in which some of the new ideas in model evaluation are applied to air quality modeling systems being used in North American and Europe, to assess the utility of the techniques and to compare and contrast model evaluation results among different models on both sides of the Atlantic Ocean. An AQMEII Workshop was conducted in April 2009 in Stresa, Italy to discuss model evaluation concepts and establish collaborative model application and evaluation projects.

1. Introduction

Regional-scale three-dimensional numerical photochemical air quality simulation models (AQMs) are being used for air quality management decisions and for short-term forecasting of air quality in many nations around the globe. To build confidence in the model estimates, a model must be critically evaluated to assess whether it is properly simulating the spatial and temporal features on the scales resolved by the model. The evaluation also assesses whether the physical and chemical processes are simulated correctly in the model, leading to proper model response to changes in meteorology and emissions, the principal classes of input

data required by AQMs. To this end, a renewed examination is needed to establish the best methods for assessing the performance of regional-scale AQMs. Motivated by discussions at a model evaluation Workshop in 2007, sponsored by U.S. EPA and the American Meteorological Society (AMS), a new framework for regional-scale air quality model evaluation is introduced.

2. Model Evaluation Framework

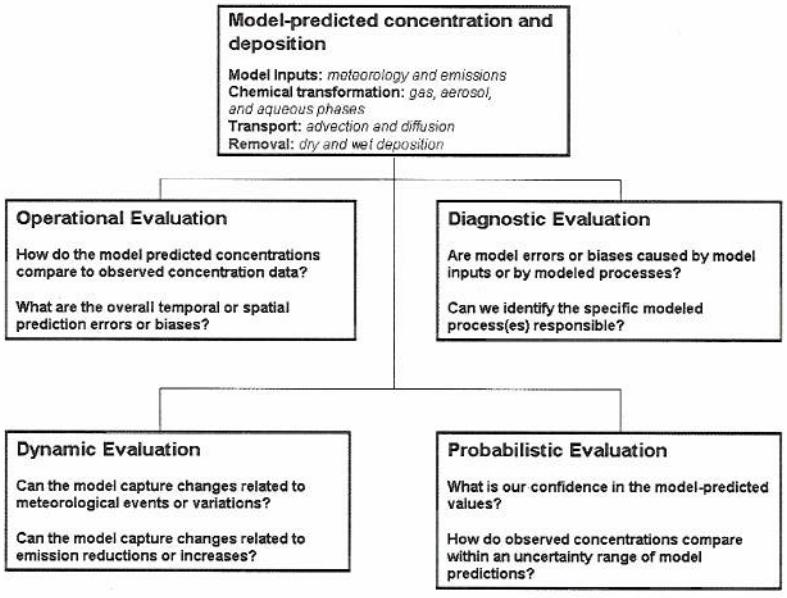

Fig. 1. The framework components of air quality model evaluation.

As shown in Figure 1, the framework consists of four components for regional-scale air quality model evaluation. **Operational Evaluation** characterizes how well model predictions compare to observations for specific time periods and conditions. Key questions to consider as part of operational evaluation include: 1) what statistical metrics and graphical depictions are most useful in assessing air quality model performance? and 2) how can point measurements and volume-averaged model predictions be reconciled in terms of spatial and temporal scales? In **Diagnostic Evaluation,** we determine how the performance of the model could be improved through complementary process-based analysis of modeled and measured values. Specific model processes or data are evaluated with the goal of

attribution of model errors. Figure 2 is an example of a diagnostic study that used inverse modeling of ammonia (NH_3) to update the emissions inventory used in the Community Multiscale Air Quality (CMAQ) model to improve model predictions of aerosol nitrate, NO_3 (Gilliland et al., 2003).

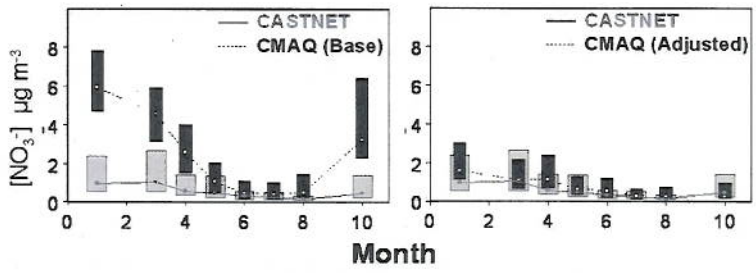

Fig. 2. CMAQ model ambient $PM_{2.5}$ nitrate estimates and CASTNET observations before (left) and after (right) top-down estimates of NH_3 emissions.

Dynamic Evaluation characterizes how well the model captures observed changes in air quality induced by changes in emissions and/or meteorology. Some of the challenges involved in dynamic evaluation include distinguishing between meteorological and emissions signals in the model results and determining the relevant space/time scales on which to conduct these evaluations. Assessing the emissions signal is particularly challenging, as anthropogenic source emissions tend to change slowly over time (years), except for weekday/weekend activity differences, and emissions controls implemented over a short period of time. **Probabilistic Evaluation** transforms deterministic model predictions into probabilistic form, helping to build confidence in the use of air quality models in policy setting. This approach, for example, may characterize the probability of success of an emissions control option in meeting a given air quality objective. One key issue is determining the role of ensemble modeling in probabilistic model use and evaluation (Pinder et al., 2009).

3. Air Quality Model Evaluation International Initiative (AQMEII)

Inspired by the emergence of the model evaluation framework and the discussions held at the August 2007 EPA/AMS Workshop, the Air Quality Model Evaluation International Initiative (AQMEII) is now proposed. To start this new collaborative project, a Workshop was held during 27-29 April 2009 in Stresa, Italy, hosted by the European Commission's Joint Research Centre and attended by 50 scientists from North America and Europe. Workshop discussions covered the types of model evaluations contained in the framework and were motivated by key

questions (see Figure 3). The goals of the Workshop included exchanging expert knowledge in regional air quality modeling, identifying knowledge gaps in the science, establishing model evaluation methods to increase knowledge about relevant processes and support the use of models for policy development, and initiating a coordinated research project on model evaluation/intercomparison. A major outcome from the Workshop was the plan for a near-term North American/European air quality model intercomparison using modeling platforms from both continents, and including aspects of all four types of model evaluations contained in the framework including the demonstration of new methods of evaluation. In addition to this activity, a plan is being developed to establish a long-term vision for maintaining international collaboration in model evaluation and rapidly advancing the science in air quality models. Detailed information on AQMEII and its initiative is available at the website: http://aqmeii.jrc.ec.europa.eu.

OPERATIONAL EVALUATION	DYNAMIC EVALUATION	PROCESS ORIENTED (DIAGNOSTIC) EVALUATION	PROBABILISTIC EVALUATION
M. Moran (Chair), J. Brandt (Rapporteur), G. Kallos (Provocateur)	P. Builtjes (Chair); K.H. Schluenzen (Rapporteur), C. Hogrefe (Provocateur)	M. Beekmann (Chair), M. Schaap (Rapporteur), K. Schere (Provocateur)	D. Steyn (Chair), R. Vautard (Rapporteur), S. Galmarini (Provocateur)
When evaluating AQ models, how do we determine, represent and present uncertainty?			
How best should we deal with uncertainty in model input and observations when evaluating AQ models?			
What kind of evaluation exercise would you perform in this context that has not yet been performed? Suggestions for new evaluation activities?			
What evaluation results will show that a model is not applicable to a case? Can we opt for falsification rather than validation criteria?			
How can we determine whether an AQ model is appropriate for the context in which it is to be applied?			

Fig. 3. AQMEII Workshop sessions and facilitators and major questions for discussion.

Disclaimer: Although this paper has been reviewed by EPA and approved for publication, it does not necessarily reflect EPA's policies or views.

References

Gilliland, A.B., R.L. Dennis, S.J. Roselle, T.E. Pierce, 2003: Seasonal NH_3 emission estimates for the Eastern United States using ammonium wet concentrations and an inverse modeling method. *J. Geophys. Res.-Atmos.*, **108**, doi: 10.1029/2002JD003063.

Pinder, R.W., R.C. Gilliam, K.W. Appel, S.L. Napelenok, K.M. Foley, A.B. Gilliland, 2009: Efficient probabilistic estimates of surface ozone concentration using an ensemble of model configurations and direct sensitivity calculations. *Environ. Sci. Technol.*, **43**, 2388-2393.

Discussion

Question: Is AQMEII solely focused on regional air quality models?
Response: Initially yes, the focus will be on regional-scale models. Eventually the program may broaden to multi-scale modeling, including regional-to-urban and global-to-regional scale model domains.

Question: Is the participation in AQMEII closed?
Response: No, the participation in AQMEII is not closed. In fact, we are encouraging members of the air quality modeling and monitoring communities to participate in AQMEII.

Question: Is the IPCC framework for model intercomparison and representation of uncertainty one that our community should do more to emulate?
Response: The IPCC has done an excellent job regarding global climate model evaluation and intercomparison. We need to adapt some of the global modeling community's evaluation techniques to our regional air quality models.