**New Directions in Air Quality Model Evaluation:
Probabilistic Model Evaluation (2.6)**

Kristen Foley, Robert Pinder, Sergey Napelenok

Collaborators: Christopher Frey (NC State University), Halûk Özkaynak (NERL/HEASD)

When weighing the societal benefits of different air quality management strategies, policy-makers need quantitative information about the relative risks and likelihood of success of different options to guide their decisions. A key component in such a decision support system is an air quality model that can estimate not only a single "best estimate" but also a credible range of values to reflect uncertainty in the model predictions. Probabilistic evaluation is a relatively new component of AMAD's model evaluation framework. This aspect of the evaluation of the Community Multiscale Air Quality (CMAQ) model seeks to answer these questions: How do we quantify our uncertainty in model inputs and parameterizations? How do we propagate this uncertainty to the predicted model outputs? How do we communicate our level of confidence in the model-predicted values in a way that is valuable and useful to decision-makers?

To address these questions, we have deployed a combination of deterministic air quality models and statistical methods to derive probabilistic estimates of air quality. For example, an ensemble of deterministic simulations is frequently used to account for different sources of uncertainty in the modeling system (e.g., emissions or meteorological inputs, boundary conditions, parameterization of chemical or physical processes). A challenge with ensemble approaches is that chemistry-transport models require significant amounts of input data and computational resources to complete a single simulation. We have applied the Decoupled Direct Method (CMAQ-DDM-3D) to generate high-member ensembles while avoiding the major computational cost of running the regional air quality model multiple times. We have also used the Bayesian Model Averaging (BMA) statistical technique to postprocess the ensemble of model runs based on observed pollutant levels. Maximum likelihood estimation is used to fit a finite-mixture statistical model to simulated and observed pollutant concentrations. The final predictive distribution is a weighted average of probability densities, and the estimated weights can be used to judge the performance of individual ensemble members, relative to the observations. Daily ozone data from a set of Air Quality System (AQS) monitoring stations in the Southeast United States are used to select a set of weighted ensemble members that, when compared to observations, have minimum spread but still capture the observed variability.

The results from this research provide an estimated probability distribution of pollutant concentration at any given location and time in the study domain. The full probability distribution can be used in several ways, such as estimating a range of likely, or "highly probable," concentration values, or estimating the probability of exceeding a given threshold value of a particular pollutant. This information can be used to quantitatively compare the relative risks and rewards of air quality control options and to select the emissions control strategy that has the largest probability of success. The information contained in the ensemble of model simulations can also be used as a model evaluation tool by uncovering cases and locations that cannot be explained using a realistic range of model inputs.