Evaluation of O_3 Predictions from the National Air Quality Forecasting System to Support Transition of Air Quality Modeling Research to Operations

Brian Eder, Rohit Mathur, Jonathan Pleim, Tanya Otte, Kenneth Schere, Jeffrey Young, George Pouliot

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

Environmental Issue

Although air quality has improved significantly since the passage of the Clean Air Act and its Amendments, there are still many areas in the United States where the public is exposed to unhealthy levels of air pollutants, most notably ozone (O_3) and fine particulate matter ($PM_{2.5}$).

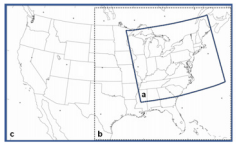
The cost of poor air quality to the US from pollution-related illnesses has been estimated by EPA to exceed 150 billion dollars annually.

For many citizens, especially those who suffer from respiratory problems, the availability of air quality forecasts, analogous to weather forecasts, could make a significant difference in how they plan their daily activities and in turn improve the quality of their lives.

Accordingly, EPA entered into a partnership with NOAA to develop the real-time National Air Quality Forecast System (NAQFS).

Background

Initially deployed during **2004,** the NAQFS provided next-day forecasts of surface-level O_3 concentrations (both 1 and 8-hour averages) at a 12 km resolution for the northeast quadrant of the US (Domain "a" in figure below). This initial System utilized the Eta meteorological model (Black, 1994; Rogers et al., 1996) coupled with EPA's CMAQ model (Byun and Schere, 2006) as described in Otte et al. (2005).



2005 An updated version of the NAQFS began providing forecasts over an expanded domain (b).

2007 The Eta meteorological model was replaced with the NAM meteorological model run of the Nonhydrostatic Mesoscale Model (NMM) core of the Weather Research and Forecasting (WRF) system (Janjic, 2003). Additionally, the domain was expanded to encompass the contiguous US and bordering areas (c).

Research Objectives

The purpose of this evaluation research has been twofold:

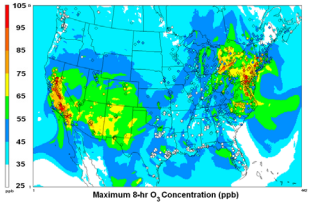
1) to identify problem areas with both the meteorological and CMAQ models in O_3 prediction, thereby providing critical information to the model developers; and

2) to support, in part, the NAQFS's transition from research to operations during each major upgrade and expansion.

Model Evaluation Approach

The NAQFS's performance was closely monitored and evaluated on a *daily basis* during each forecast season. Such "real-time" evaluation of the forecasts, under constantly varying dynamical and chemical conditions, helped identify model bias/error and related model deficiencies.

The evaluations used a stringent, direct matching scheme between each EPA AIRNow monitor (nearly 1100 for the contiguous United States (CONUS)) and NAQFS grid cell. An example of the System's performance for 13 June, 2008 is provided below with monitors denoted as diamonds.



Weekly compilations of the evaluations were then created and discussed in briefings held between EPA and NOAA.

The evaluations utilized a suite of statistical metrics, focusing primarily on *discrete forecasts* (observed versus forecast maximum 8-hour O_3 concentrations) provided by the System, including the correlation coefficient (r), Root Mean Square Error (RMSE), Normalized Mean Error (NME), Mean Bias (MB) and Normalized Mean Bias (NMB)

Each metric was calculated throughout the evaluation period and was summarized for various temporal (i.e., daily, monthly and seasonal) and spatial (CONUS and subregions) scales.

The Results and Discussion that follow will focus primarily on the latest evaluation, for the summer of 2008.

Results and Discussion

✣ Spatial Evaluation Results

As seen in the table below, the NAQFS generally performed better, in terms of correlation, over eastern regions, with poorer performance over the RM and PC subregions.

2008 Regional Analysis of the Maximum 8-hour O_3 Forecasts

Domain	Obs (ppb)	Mod (ppb)	RMSE (ppb)	NME (%)	MB (ppb)	NMB (%)	r
Contiguous States	47.6	51.6	12.6	20.3	4.0	8.4	0.67
Northeast (NE)	49.9	53.8	10.8	16.7	3.8	7.7	0.72
Southeast (SE)	47.8	54.5	12.5	20.8	6.7	14.4	0.73
Upper Midwest (UM)	47.7	51.5	10.1	16.4	3.4	7.2	0.64
Lower Midwest (LM)	39.8	48.8	14.2	29.9	9.0	22.7	0.71
Rocky Mountain (RM)	55.5	57.0	11.2	15.8	3.0	5.5	0.52
Pacific Coast (PC)	57.7	51.6	17.1	22.6	-3.9	-7.1	0.61

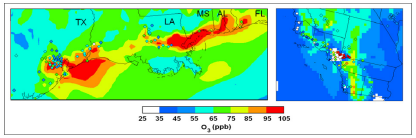
Results and Discussion (cont'd.)

There are two likely reasons for these spatial differences:

- First, the NAQFS has been in operation over the NE domain (over the SE, UM and LM domains) for 5 and 3 years, respectively, allowing iterative refinement; thus, better performance in these areas.
- Second, the RM and PC subregions are dominated by complex terrain, which can be problematic for numerical models, considering that the NAQFS has a 12 km resolution.

Examination of the spatial distribution of the NMB (not shown) also revealed several interesting characteristics:

- The performance is comparable across the NE, UM and RM subregions as the NMB generally averages under 10%.
- The SE and LM subregions exhibit higher NMBs, due mainly to the large overpredictions often found along coastal regions (see example below). Eder et al. (2008) attributed this to



- The negative biases found in the PC subregion were caused by very large and systematic underpredictions in areas adjacent to and especially downwind of the Los Angeles basin (see example above). Eder et al. (2008) found this was due to an overestimation of NO_x emissions on weekends.

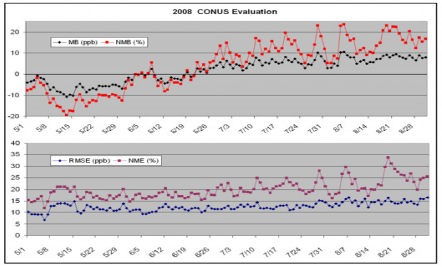
✣ Temporal Evaluation Results

Time series of the CONUS-wide correlations during 2008 (not shown) exhibit a fairly consistent nature, with values exceeding 0.60 over 90% of the days and 0.70 over 55% of the days.

- The few days when the correlations fell below 0.60 were associated with extensive cloud cover and precipitation.

Unlike the correlations, measures of bias and error were *not* temporally consistent (as seen below), rather they exhibited an increasing trend as the forecast season progressed.

- These increases are attributable, in part, to poor temperature



Conclusions

Results of the numerous seasonal evaluations indicate that the NAQFS has and continues to perform well through its numerous refinements and expansions. Mean CONUS-wide correlations typically exceed 0.70. Values of NMB and NME are generally within 10% and 20%, respectively.

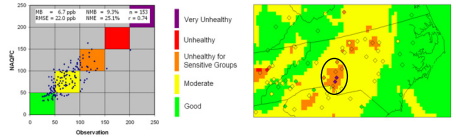
Closer examination of the NAQFS' performance over finer spatial and temporal scales has enabled model developers to identify and subsequently address systematic deficiencies with the various NAQFS configurations.

When compared to the evolution of numerical weather prediction's incorporation into weather forecasting, that took decades, the success of this program, which involves numerical O_3 forecasts based on the WRF-CMAQ model, has been realized in a remarkably short time.

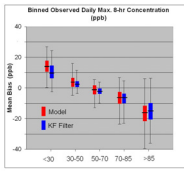
Impact and Future Directions

Forecasts from the NAQFS (disseminated via NOAA's website: www.nws.noaa.gov/aq) have provided both State and local air quality forecasters, as well as the general public, with timely, accurate forecasts of O_3 concentrations.

- States and local agencies have used the forecast as guidance in the issuance of Ozone Action Days based on EPA's **Air Quality Index**. (An example AQI forecast and evaluation is shown below for Charlotte, NC.)



- Additionally, a post-processing, bias-adjustment technique (based on the Kalman Filter [KF]), was developed that adjusts forecasts with recent observations (Kang et al., 2008). This technique, when applied to the 2008 O_3 season, significantly reduced model bias as seen below.



AMAD's commitment to this project has come to a successful fruition, as future development and evaluation of the System has been handed off to NOAA and EPA operations.

Contributors/Collaborators

Daiwen Kang, Shaocai Yu, Daniel Tong, Hsin-mu Lin, and Tianfeng Chai from Science and Technology Corporation. Paula Davidson and Jeff McQueen from NOAA, NWS