4.2 AN EXPERIMENTAL AIR QUALITY FORECAST MODELING SYSTEM (AQFMS) FOR THE NORTHEAST UNITED STATES: A DEMONSTRATION STUDY

C. Cai[1], C. Hogrefe[1], J. Biswas[1], S.T. Rao[1], N. Seaman[2], A. Gibbs[2], G. Kallos[3], P. Katsafados[3], C. Walcek[1], J. Chang[1], and K. L. Demerjian[1*]
[1]Atmospheric Sciences Research Center, University at Albany, Albany, NY
[2]Penn State University
[3]University of Athens

1. INTRODUCTION

A prototype air quality forecast modeling system (AQFMS) was developed and run to examine whether reliable real-time 18-hr air quality forecasts can be made for the Northeast United States. The experimental forecasts were in part, envisioned to provide assistance in planning and operational deployments in the PMTACS-NY Supersite Summer 2001 Field Intensive (http://www.asrc.cestm.albany.edu/pmtacsny/). The goal was to provide daily 18-hr air quality forecasts on the web for the region and the New York metropolitan area specifically, from July 1 through August 31, 2001.

The prototype AQFMS was designed to operate with forecasted meteorological fields from either of two mesoscale meteorological models, the Penn State/NCAR Mesoscale Model MM5 (Grell, et al., 1994) or the University of Athens ETA-SKIRON meteorological model (Nickovic, et al., 2001; http://forecast.uoa.gr/charactnew.html). The meteorological fields were used to drive the Comprehensive Air Quality Model with Extensions (CAMx) (Environ, 2000), a photochemical air quality simulation model, to provide two distinct air quality forecasts. Emissions used in the AQFMS are based on revised OTAG anthropogenic emissions and BEIS-II biogenic emissions. The current application did not use prognostic meteorological data to drive affected emissions sources.

2. AQFMS – COMPONENTS

A schematic diagram of the components of the AQFMS is presented in Figure 1. The two-mesoscale meteorological models used to drive CAMx air quality model operated over somewhat different domains and coordinate systems. The ETA/CAMx and MM5/CAMx domains are shown in Figure 2a-b, with the former using a lat-lon and the latter a Lambert coordinate system. Further details of the model systems are provided in Tables 1 and 2. Initial conditions for the CAMx air quality simulation are obtained from the previous day's forecast. If the previous day's forecast is not available, a two-day spin-up simulation starting from clean initial conditions is performed prior to the forecast simulation using weekday/weekend emissions as appropriate over the modeling domain. The emission inventory is a derivative of the 1995 OTAG inventory that was upgraded for New York State emissions. The generic weekday/weekend gridded emissions used in the air quality forecast were based on an episodic case selected in mid-July from the OTAG emissions data set. The emissions from this episode likely represent an upper limit for typical 2001 emissions in this region. The spatial and temporal distribution of UV, ozone column density and turbidity for the CAMx domain, are processed from daily downloaded TOMS data (ftp://toms.gsfc.nasa.gov). A gridded land use inventory comprised of 11 land use categories is used in the calculation of dry deposition and provides albedo input for the calculation of photolytic rate constants. Cloud cover, based on MM5 and ETA coarse domain parameters are processed within CAMX and also incorporated into the calculation of photolytic rate constants.

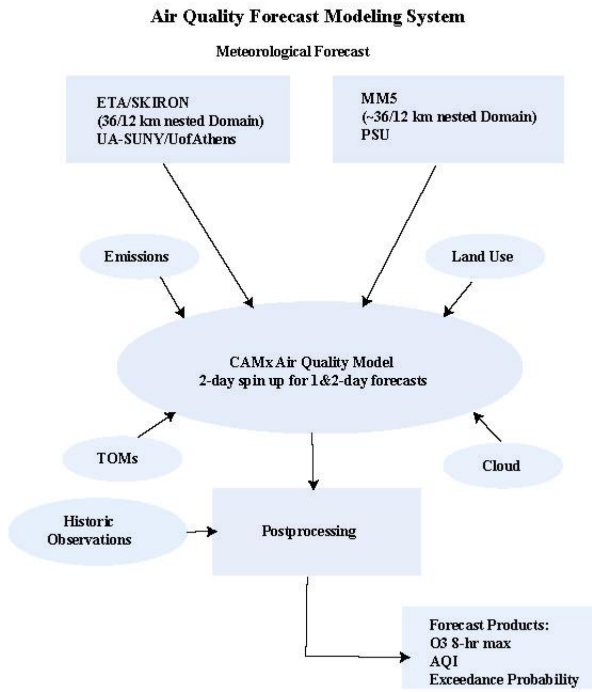

Figure 1. AQFMS – Schematic Diagram

3. OPERATIONAL DETAILS

The MM5 and ETA meteorological forecasts were not available at the same time. In the case of the operational MM5 forecast, Seaman et al., PSU, provided meteorological data sets via ftp to ASRC in two forms. The first data set, a lower resolution 36 km forecast, was received at ~ 3:00am. The CAMx air quality model was run in the nested mode using the MM5 36 km data and a processed interpolated 12 km meteorological data set for the nested domain. As noted in section 2, the air quality model domains for the MM5 and ETA driven CAMx runs are not the same. The initial air quality forecast using the 36km forecast data set was available at ~6:30am for post processing and distribution by 7:30am. A second 12 km forecast data file was received typically by 8:30am. The CAMx model was rerun in the nested mode using 12km meteorological forecast data to produce the final air quality forecast. The 12 km air quality forecast was available at ~9:40am for post processing and distribution by 10:00am.

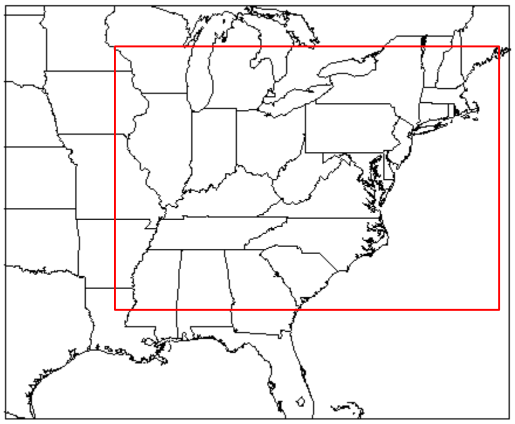

Figure 2a. ETA/CAMx Nested Modeling Domain

Figure 2b. MM5/CAMx Nested Modeling Domain

The ETA operational forecasts were carried out at the ASRC, University at Albany in collaboration with the University of Athens. The ETA model used is based on the SKIRON Version 6.0 meteorological model developed at the University of Athens (ref or website).

The 12 km ETA meteorological forecast data are typically available by 6:30am. The ETA forecast data are post processed and converted to the CAMx format. This involves first transforming the ETA horizontal grid projection (0.12x0.12deg) into a lat/lon projection (0.166x0.111deg), defining the coarse and fine modeling domains (see Table 1 for details) and aggregating grid points within the coarse grid domain. As with the MM5-CAMx simulations, the CAMx is run in a nested mode, but now over a larger spatial domain (see Figures 2a-b). The 12 km ETA-CAMx air quality forecast was available at ~9:30am for post processing and distribution by ~9:45am.

The MM5/CAMx air quality forecasts were carried out on a Sun Microsystems ES6000 with 16, 235mhz processors (operating in a shared environment), while the ETA/SKIRON meteorological forecast and ETA/CAMx air quality forecast were carried out in tandem on a dedicated Sun Microsystems SunBlade Model 1000 with 2, 750mhz processors.

4. FORECAST PRODUCTS

CAMx forecasted ozone concentration fields from the MM5 and ETA driven simulations were post processed to prepare several products for distribution on the web, http://www.asrc.cestm.albany.edu/AQFMS/. The web homepage (see Figure 3) allows the user to select among these products and provides a side-by-side comparison of the two forecasts.

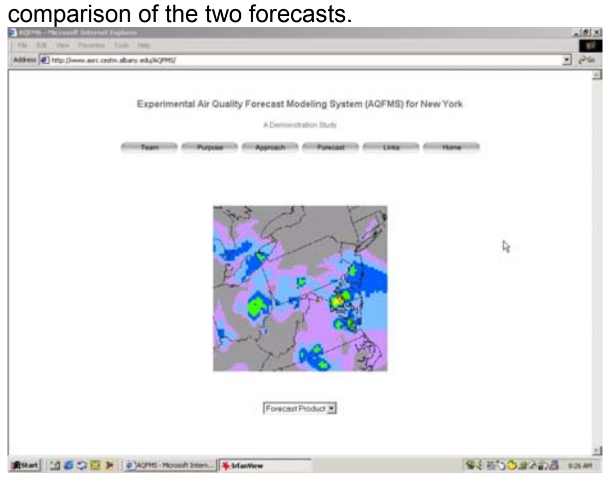

Figure 3. Homepage website: Experimental Air quality Forecast Modeling System

The analyzed products include:
1) Max 8hr O_3 - Model predicted 8-hour averaged maximum ozone concentrations in units of parts per billion (ppb) (see Figure 4 for example MM5/CAMx forecast).
2) Max 8hr AQI – Air Quality Index derived from model predicted peak 8-hour, ground level ozone concentrations. The index is based on EPA's AQI, which associates air pollution levels health concerns (see http://www.epa.gov/airnow/aqibroch/aqi.html#2).
3) Probability of Exceeding AQI – Estimate of the probability (% chance) that the AQI will have a value greater than 100 (AQI of 100 is the break point between moderate and unhealthy for sensitive groups). The probabilities are estimated from model predictions and historic ozone observations (see companion paper Hogrefe, et al., 2002).
4) O_3 Movie – Predicted hourly average ozone concentrations (ppb) from 1:30 to 19:30 eastern day light savings time.

5) CO Movie – Predicted hourly average CO concentrations (ppb) from 1:30 to 19:30 eastern day light savings time.
6) NO_y Movie - Predicted hourly average NO_y concentrations (ppb) from 1:30 to 19:30 eastern day light savings time.

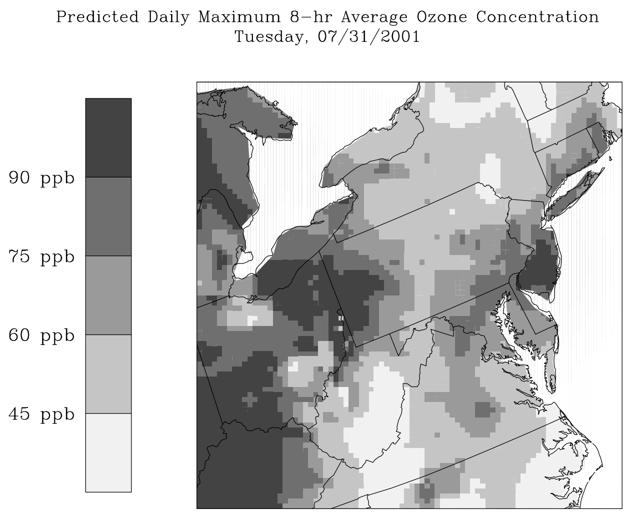

Figure 4. Example of a MM5/CAMx forecast of daily maximum 8-hr average ozone concentrations

5. AIR QUALITY FORECAST RESULTS

Data sets for MM5-CAMx air quality forecasts are available from June 1 to present and for ETA-CAMx from July 17 to present. The present analyses consider data from July 1 – August 31 and selected periods therein. With the exception of ozone, the availability of real-time observational data is quite limited. Therefore the current analyses draw upon preliminary ozone observations EPA's AIRNOW system (R. Wayland, EPA, personal Communication) and air quality observations from monitoring sites for which we have direct access to ozone precursor data. Even here data reduction and quality assurance checks impose time constraints that limit quick turn around and availability of data, so only a partial set of the many chemical parameters monitored are presented. Future performance evaluations will consider additional parameters from these and other sites in the region as they become available.

Evaluation Statistics for Ozone Forecasts

To evaluate ozone predictions from the MM5/CAMx and ETA/CAMx modeling systems, predicted concentrations are compared to preliminary ozone observations from EPA's AIRNOW system. The number of ozone monitors within the 12km modeling domains is 271 and 504 for the MM5/CAMx and ETA/CAMx modeling systems, respectively. For the present study, we compute the mean normalized bias error, mean normalized gross error, and unpaired peak prediction accuracy for the 1-hr and 8-hr averaged daily maximum ozone prediction for each of the modeling systems. The use of these metrics for model evaluation in air quality management studies was stipulated by EPA, and criteria for satisfactory performance of such hind cast simulation as were set forth (U.S. EPA, 1991). In the present study, only observation/predictions pairs for which the observed daily maximum concentration is larger than 20 ppb are considered. Figure 5 provides an example of the spatial distribution of the mean normalized bias error for the 8-hr averaged daily maximum ozone prediction for the MM5/CAMx forecast calculated for the period July 1- August 31 at each monitor within the 12 km domain. A positive bias greater than 35% is present for parts of the Washington-Philadelphia corridor as well as southern Ohio and northern Kentucky.

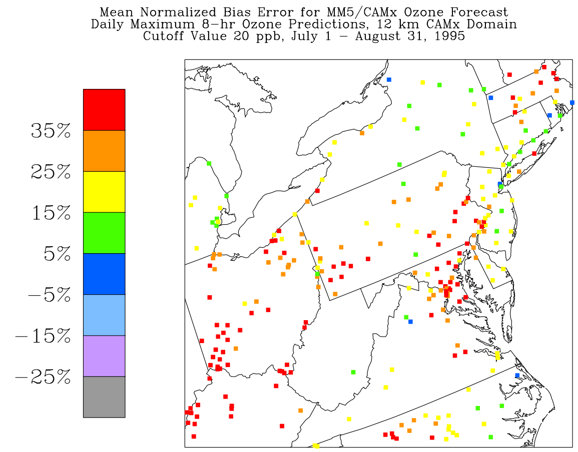

Figure 5. Spatial distribution of the mean normalized bias error for the 8-hr averaged daily maximum ozone prediction for the MM5/CAMx forecast.

Time series of evaluation statistics reporting the mean normalized bias, mean normalized gross error and normalized unpaired peak prediction accuracy for the daily maximum 8-hr averaged ozone prediction over the July-August forecast period are presented in Figures 6a and 6b for the MM5/CAMx and ETA/CAMx respectively. For these time series, the statistics are computed over all monitors within the 12km domains (which are different for the two modeling systems) on any given day. These results indicate that the daily maximum 8-hr averaged ozone predictions are biased high for both forecast modeling systems, over predicting ozone by ~25% ± 20%. While these values are higher than those established by the EPA for hind cast simulations (U.S. EPA, 1991) and reported in hind cast simulation evaluation studies (e.g. Holgrefe et al., 2001, Sistla et al., 2001), these results are nevertheless encouraging, since the forecast simulations analyzed here do not use data assimilation techniques. The potential source of the ozone over prediction is discussed below.

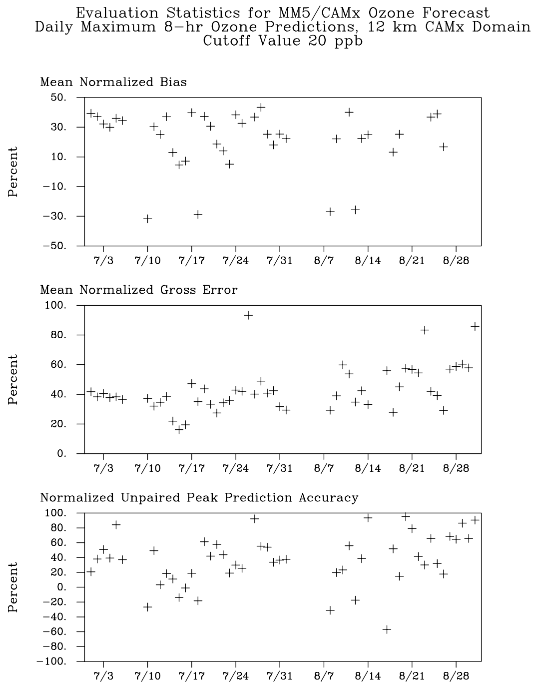

Figure 6a. MM5/CAMx evaluation statistics daily maximum 8-hr average ozone forecast

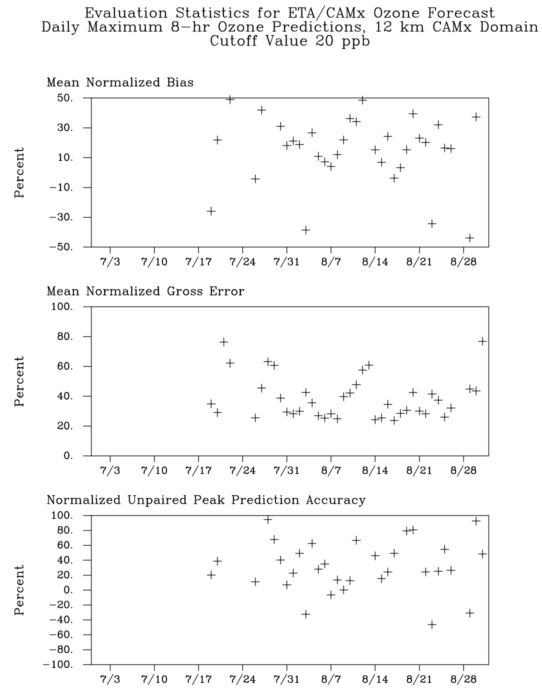

Figure 6b. ETA/CAMx evaluation statistics daily maximum 8-hr average ozone forecast

Time Series comparisons of CO, NO$_y$ and O$_3$

To gain some insight into the basic performance of the AQFM systems, times series of forecasted CO, NOy, and O_3 concentrations for the two models are compared with observations at Whiteface Mountain (44.4N, 73.9W; elevation 1500m) in Figures 7a-c. Where, NOy = NO_x+HNO_3+N_2O_5+RNO_3 + particulate NO_3^- is a nitrogen budget species that is conserved in terms of chemical transformations, but which can be depleted by physical removal processes (i.e., dry and wet deposition). The NO_y removal proceeds predominantly through the loss of nitric acid, HNO_3. The concentration of the oxides of nitrogen in the atmosphere and their chemical partitioning are critical in understanding the oxidizing potential and O_3 formation in the atmosphere.

The results show reasonable agreement in the observed and forecasted temporal patterns of these precursor species. CO forecasted concentrations track observations quite well, with some periods of over and under prediction, while NOy model predictions tend to be bias high in almost all cases. Similar time series comparisons at Pinnacle State Park (42.9N, 77.2W; elevation 200m), not shown here, are consistent with these findings. The matching of temporal pollution patterns suggest that transport is being reasonably represented by the meteorological forecast models. The inability to match the amplitude of these temporal pollution patterns can arise from a variety of sources of uncertainty within the modeling system. The preliminary analyses that follow give some indication of what those sources might be.

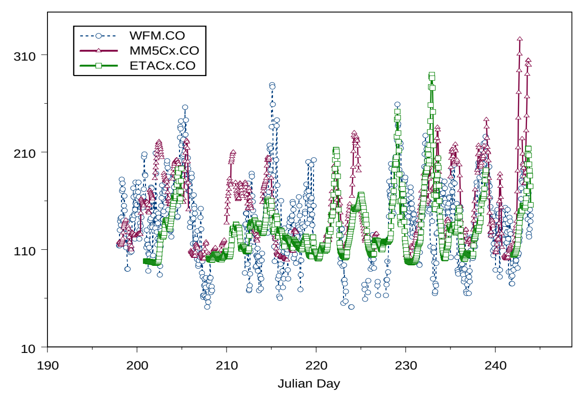

Figure 7a. Time series of MM5/CAMx and ETA/CAMx forecasted CO and observed CO concentrations (ppb) at Whiteface Mountain, NY.

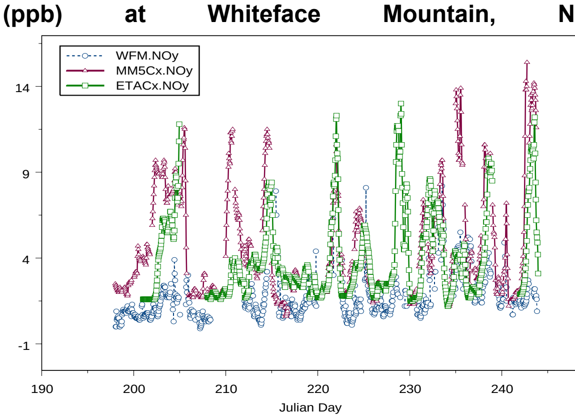

Figure 7b. Time series MM5/CAMx and ETA/CAMx forecasted NOy and observed NOy concentrations (ppb) at Whiteface Mountain, NY.

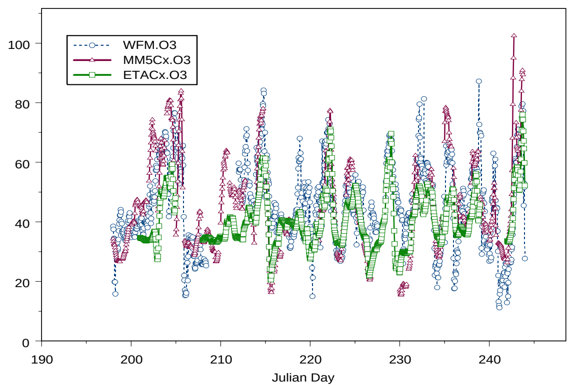

Figure 7c. Time series of MM5/CAMx and ETA/CAMx forecasted O3 and observed O3 concentrations (ppb) at Whiteface Mountain, NY.

Correlation Analyses

Given that CO is a representative surrogate for mobile source hydrocarbons and NO_y for all NOx emission sources, the correlation of NO_y and CO provides an indication of the basic performance of the model in predicting these fundamental precursor relationships.

The correlation of forecasted NO_y vs. CO for the two modeling systems (Figures 8a-b) indicates a CO/NO_y ratio of 15 $\pm$1, while the observations (Figure 8c) indicate a ratio of 27 $\pm$ 2. These results suggest that NOx emission may be over estimated in the inventory. Since the dynamic range of modeled CO for the July – August period is in general agreement with observations, this would suggest that meteorological dynamics are not the likely source of the NO_y over prediction. Another possible explanation for this bias might be that the CAMx removal processes (dry and wet deposition) for HNO_3 are under predicted, but correlation of modeled [NOy-HNO_3] vs. CO, which represents the extreme upper limit for removal, resulted in a CO/NOy ratio of 19 $\pm$ 1, suggesting the NO_x emission inventory in question.

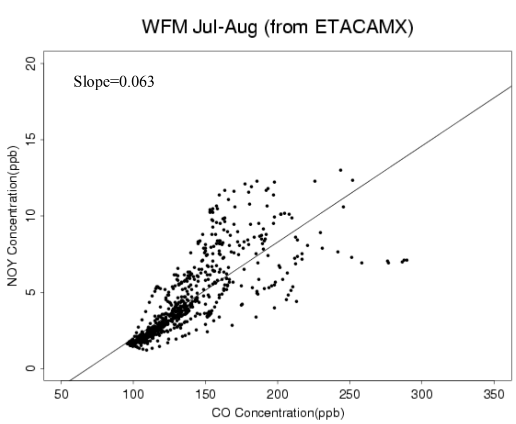

Figure 8a. Correlation of forecasted NO_y vs. CO ETA/CAMx modeling system

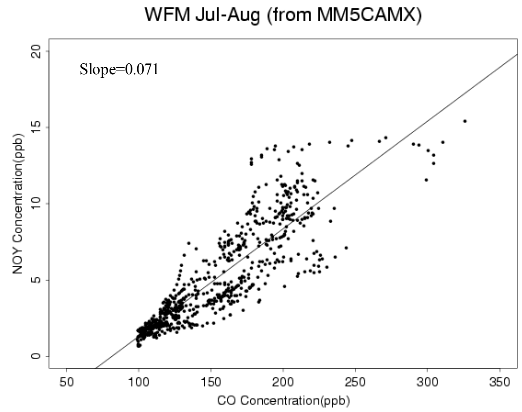

Figure 8b. Correlation of forecasted NO_y vs. CO MM5/CAMx modeling system

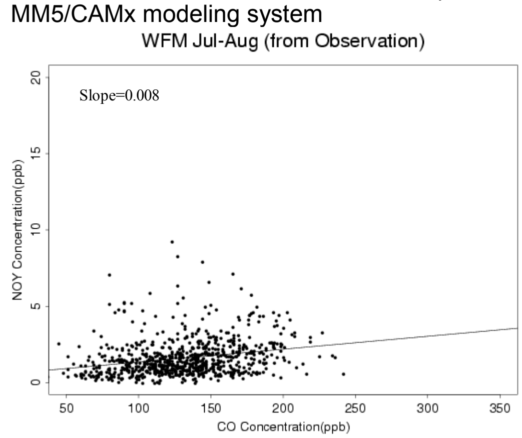

Figure 8c. Correlation of observed NO_y vs. CO from Whiteface Mountain, NY.

6. SUMMARY

An experimental air quality forecast modeling system utilizing two distinct meteorological forecasting models (MM5 and ETA/SKIRON) to drive a photochemical air quality simulation model (CAMx) was designed and operated during the summer of 2001 to provide 18-hour air quality forecasts for distribution on the web. This feasibility study demonstrated that the operation of an AQFMS could be maintained with minimal interruption using available relatively inexpensive high performance computing systems. Initial post analysis evaluations of the air quality forecasts look very promising, indicating ozone evaluation statistics in the expected range compared to reported hind cast results. The correlation analyses of NO_y and CO indicate that the NO_x emission inventory used in the forecast model are likely high, leading to systematically high predictions of O_3 and NO_y as compared with observations. Further analyses considering expanded observational data sets will be performed as these data become available.

7. ACKNOWLEDGEMENT

The work was supported in part as an ASRC/University at Albany internal research initiative in collaboration with the University of Athens, Penn State University and the NYS Department of Environmental Conservation.

8. REFERENCES

ENVIRON, 2000: Comprehensive Air Quality Model with Extension (CAMx), Version 3.00, ENVIRON International Corporation, 101 Rowland Way, Navato, CA 94945-5010.

Grell, G.A., J. Dudhia, and D.R. Stauffer, 1994: A description of the fifth generation Penn State/NCAR mesoscale model (MM5). NCAR Technical Note TN-398, Boulder, CO.

Nickovic, S., G. Kallos, A. Papadopoulos and O. Kakaliagou, 2001: "A model for prediction of desert dust cycle in the atmosphere". J. Geophysical Res. 106(D16):18113-18129.

Hogrefe, C., S.T. Rao and K.L. Demerjian, 2002: Using current–generation meteorological/photochemical modeling systems for real-time ozone forecasting. Preprints, Sixteenth on Probability and Statistics in Atmospheric Sciences, Orlando, FL, Amer. Meteor. Soc.

Hogrefe,C., S.T. Rao, P. Kasibhatla, W. Hao, G. Sistla, R.Mathur, and J.McHenry, 2001: Evaluating the performance of regional-scale photochemical modeling systems: part II- ozone predictions. Atmos. Environ., 35, 4175-4188.

Sistla, G., et al., 2001: An operational evaluation of two regional-scale ozone air quality modeling systems over the eastern United States. Bull. Amer. Meteor. Soc., 82, 945-964.

U.S. Environmental Protection Agency, 1991: Guideline for regulatory application of the Urban Airshed Model, EPA-450/4-91-013, July 1991, U.S. EPA, Research Triangle Park, NC 27711.

Table 1. ETA-CAMx Model Characteristics

ETA Meterological Model Domain	
ETA Domain : 106.7W to 59.5W, 50.4N to 23.6N Horizontal grid increment: 0.12 x 0.12 deg (~12km) Number of ETA grid points in x-y plane: 133 x 213 Number of vertical levels: 32 Origin of initial boundary conditions: NCEP at 10 pressure levels Computation Time for 48hr forecast: 6hr30min	
ETA-CAMx Coarse Grid	**ETA-CAMx Fine Grid**
CAMx Coarse Domain:99.0W to 67W,26N to 47N Horizontal grid increment: 0.500 x 0.333 deg Number of CAMx grid points x-y plane: 64 x 63 Number of vertical levels: 14 Initial conditions: previous day's forecast; if not available 2-day model spin-up with clean conditions to generate initial conditions for the forecast simulation Boundary conditions: Clean boundary conditions characteristic of tropospheric background conditions Computation Time for 24hr forecast: N/A	CAMx Fine Domain: 92.2W to 69.3W, 31.9N to 44.1N Horizontal grid increment: 0.166 x 0.111 deg Number of CAMx grid points x-y plane: 137 x 110 Number of vertical levels: 14 Initial conditions: previous day's forecast; if not available 2-day model spin-up with clean conditions to generate initial conditions for the forecast simulation Boundary conditions: Clean boundary conditions characteristic of tropospheric background conditions Computation Time for 24hr nested forecast: 2hr40min

Table 2. MM5-CAMx Model Characteristics

MM5-CAMx Coarse Grid	**MM5-CAMx Fine Grid**
MM5 Coarse Domain : 123.5W to 60W, 25.4N to 48.0N Horizontal grid increment: 36km Number of MM5 grid points in x-y plane: 133 x 97 Number of vertical levels: 30 Origin of initial boundary conditions: ETA 104 grids Computation Time for 60hr forecast: 3hr47min	MM5 Fine Domain : 86.8W to 69.7W, 37.0N to 42.7N Horizontal grid increment: 12km Number of MM5 grid points in x-y plane: 91 x 91 Number of vertical levels: 30 Origin of initial boundary conditions: Coarse mod. output Computation Time for 27hr forecast: 4hr14min
CAMx Coarse Domain: 98.0W to 66.0W,31.4N to 42.9N Horizontal grid increment: 36km Number of CAMx grid points x-y plane 63 x 61 Number of vertical levels: 14 Initial & Boundary Conditions: Same as ETA above Computation Time for 24hr forecast (including interpolate to 12km domain): 2hr10min	CAMx Fine Domain: 86.1W to 70.3W, 37.6N to 42.5N Horizontal grid increment: 12km Number of CAMx grid points x-y plane: 84 x 84 Number of vertical levels: 14 Initial & Boundary Conditions: Same as ETA above Computation Time for 19hr forecast: 1hr25min