4.19 NEAR-REALTIME SIMULATION OF O_3 AND $PM_{2.5}$ OVER THE NORTHEASTERN UNITED STATES: RESULTS FOR SUMMER 2004

C. Hogrefe[1,2,*], G. Sistla[1], W. Hao[1], J.-Y. Ku[1], R.S. Gaza[1], L. Sedefian[1], K. Schere[3], A. Gilliland[3], and R. Mathur[3]

[1]New York State Department of Environmental Conservation, Albany, New York
[2]Atmospheric Sciences Research Center, University at Albany, Albany, New York
[3]Atmospheric Sciences Modeling Division, National Oceanic and Atmospheric Administration, On assignment to National Exposure Research Laboratory, U.S. Environmental Protection Agency, RTP, NC

1. INTRODUCTION

The application of grid-based photochemical modeling systems to provide real-time air quality forecasts has been a fairly recent development and has been mostly restricted to the prediction of O_3 (i.e. McHenry et al., 2000, 2004; Chang and Cardelino, 2000, Cai et al., 2002; Vaughan et al., 2004; Mathur et al., 2004). Since June 2003, the National Weather Service (NWS) / National Centers for Environmental Prediction (NCEP) have been performing grid-based numerical O_3 forecasts in partnership with the National Oceanic and Atmospheric Administration (NOAA) and the U.S. Environmental Protection Agency (EPA) (Davidson et al., 2004; McQueen et al., 2004). For 2004, experimental $PM_{2.5}$ forecasts were added for selected domains (McQueen et al., 2004). On the other hand, many operational air quality predictions for O_3 and $PM_{2.5}$ by federal, state and local agencies are still based on a combination of weather predictions, statistical analyses, and expert judgment (Ryan et al., 2000; Dye et al., 2000). While numerical models can potentially provide air quality forecasts at higher spatial and temporal resolution than the traditional methods and also can provide forecasts for regions that do not have the resources to develop and apply statistical forecasting tools, it is critical to perform on-going evaluation of such predictions before photochemical models are more widely used by various agencies for real-time air quality predictions. This paper presents the initial assessment of a pilot study designed to simulate both O_3 and $PM_{2.5}$ over the Northeastern United States on a near-realtime basis for July 1 – September 30, 2004. The simulations were performed in cooperation between the New York State Department of Environmental Conservation (NYSDEC), NOAA, and EPA, utilizing resources from the operational NWS/NOAA/EPA air quality forecasts.

In this paper, we measured model performance utilizing both discrete and categorical statistical metrics. Discrete metrics such as bias, root mean square error, or absolute error quantify differences between observations and model predictions regardless of observed or predicted concentration levels, while categorical metrics are used to measure forecasting skill in terms of correctly/incorrectly predicting concentration levels above/below a certain threshold (U.S. EPA, 1999). A list of the categorical metrics used in this study is presented in Table 1. To assess the ETA/CMAQ modeling system's usefulness as a forecasting tool, we also compared its forecast statistics against those computed for the routine air quality forecasts issued by NYSDEC based on traditional techniques. Finally, to address the role of errors in the forecasted meteorology on the predictions of air quality, we present results from the re-simulation of a high $PM_{2.5}$ concentration event using meteorological fields that were closer in time to the ETA initialization time than possible in our current pilot study.

2. MODEL DESCRIPTION AND DATA BASE

The forecasting system developed by NWS/NOAA/EPA (Davidson et al., 2004; McQueen et al., 2004) and also utilized in the NYSDEC/NOAA/EPA pilot study presented here consists of operational weather forecasts from the National Weather Services (NWS) ETA model (Black, 1994) at a horizontal resolution of 12 km, the PREMAQ emissions and meteorology pre-processor (Otte et al., 2004), and the Community Multiscale Air Quality (CMAQ) model (Byun and Ching, 1999). As discussed by Mathur et al. (2004), the emission inventories used by the ETA/CMAQ system were updated to represent the 2004 forecast period. NO_x emissions from point sources were projected to 2004 relative to a 2001 base inventory and area source emissions were based on the 2001 National Emissions Inventory, version 3. Since MOBILE6 is computationally expensive and inefficient for real-time applications, mobile source emissions were estimated using approximations to the MOBILE6 model as discussed by Pouliot et al. (2003), while BEIS3.12 (Pierce et al., 2002) was used to estimate the biogenic emissions. Further details on the model setup for this air quality forecasting pilot study can be found in Pouliot et al. (2003), Pouliot (2005), Pleim and Mathur, (2005), Otte et al. (2004), and Mathur et al., (2004). In contrast to the operational air quality forecasts performed by NWS/NOAA/EPA, time-invariant climatological chemical boundary conditions were used in the NYSDEC/NOAA/EPA simulations presented here.

Furthermore, we obtained the archived official next-day air quality forecast issued by the NYSDEC for O_3 and $PM_{2.5}$ for the month of August. These NYSDEC next-day forecasts are issued on weekdays for eight regions in New York State based on statistical

* *Corresponding author contact information: Christian Hogrefe, BAQAR, New York State Department of Environmental Conservation, 625 Broadway, Albany, NY 12233-3259, Phone (518) 402 8402, Fax (518) 402 9035, Email chogrefe@dec.state.ny.us*

techniques, climatology, and expert judgment. Finally, for the purpose of model evaluation, preliminary hourly ozone and $PM_{2.5}$ observations for the summer of 2004 were obtained from the EPA's AIRNOW system. Hourly total $PM_{2.5}$ mass was measured by TEOM instruments. Additional $PM_{2.5}$ measurements from filter-based networks such as STN, IMPROVE or CASTNet that provide both total and speciated $PM_{2.5}$ mass were not yet available at the time the analysis presented in this paper was performed. Figures 1a-b show a map of the modeling domain with the locations of all O_3 and $PM_{2.5}$ monitors used in this analysis, and a map of the eight forecast regions in New York State, respectively.

3. MODEL SIMULATION SET-UP

Each ETA/CMAQ air quality simulation was performed for 48 hours utilizing 48-hr ETA forecasts initialized with 12:00 UTC analysis fields. These operational ETA forecasts are provided by the NWS between roughly 08:00 Eastern Daylight Savings Time (EDST) and 10:00 EDST. ETA fields are then interpolated vertically and horizontally on NWS machines to match the CMAQ grid structure, and these files are transferred to EPA's National Environmental Scientific Computing Center ($NESC^2$) between 18:00 and 21:00 EDST to avoid slower network connections during work hours. NYSDEC then utilizes $NESC^2$ computing facilities to generate CMAQ-ready meteorological and emission fields by executing the PREMAQ processor between 21:00 and 22:00 EDST. These input files are then transferred via ftp to NYSDEC where daily CMAQ simulations are performed on a 10 CPU PC cluster operating under Linux. Each day, CMAQ pollutant concentration fields are initialized with results from the previous day's ETA/CMAQ simulation. The CMAQ simulations typically are performed between 01:00 and 05:00 EDST, and post-processing is complete within 15 minutes of the completion of the CMAQ run. By this time, the first 21 hours of the 48-hr ETA/CMAQ forecast initialized at 12:00 UTC / 08:00 EDST have already passed so that the 'same-day' O_3 and $PM_{2.5}$ forecasts posted daily at 05:15 EDST are essentially 'next-day' forecasts from a standpoint of ETA model initialization because they are based on ETA forecast hours 16 – 40, i.e. 04:00 – 04:00 GMT or 00:00 – 00:00 EDST. In other words, the first 16 hours of the ETA/CMAQ 48-hour forecast initialized at 12:00 GMT on the previous day are discarded to generate the 'same-day' O_3 and $PM_{2.5}$ forecast posted by 05:15 EDST.

4. RESULTS AND DISCUSSION

4.1 Evaluation of O_3 and $PM_{2.5}$ Forecasts July – September, 2004

Figures 2a-b present time series of observed and predicted daily maximum 8-hr ozone and daily average 24-hr PM2.5 concentrations for July 1 – September 30, 2004. These concentrations were averaged over all O_3 and $PM_{2.5}$ monitors in the modeling domain as shown in Figure 1a; model predictions were extracted for the grid cells that correspond to each monitoring location. It is evident from both figures that the ETA/CMAQ forecasts track well with observations, with the correlation between observed and predicted time series being 0.84 for ozone and 0.65 for $PM_{2.5}$. However, it is also evident that domain-averaged predicted 8-hr daily maximum ozone concentrations do not show the general downward trend visible in the observations after August 15. On the other hand, the under-predictions of domain-average 24-hr $PM_{2.5}$ concentrations steadily decreases from July through September, with very good agreement between observations and predictions for September. A possible contributor to the under-prediction of $PM_{2.5}$ concentrations during July and August may be the impact of forest fires in Alaska on observed concentrations during this time period; this effect was not accounted for in the model predictions. While Figures 2a-b are useful to examine the modeling system's ability to capture pollution episodes and temporal trends in the observations, they provide no information about potential regional differences in model performance. Therefore, Figure 3a-c and 4a-c present maps of observed and predicted ozone and $PM_{2.5}$ concentrations averaged between July 1 – September 30, 2004 as well as the differences between the average observed and predicted concentrations. Figures 3a-c illustrate that both observed and predicted 3-month average ozone concentrations are generally higher in the southern two-thirds of the modeling domain than in the northern portions. These figures also illustrate that average ozone concentrations are over-estimated by ETA/CMAQ in all parts of the modeling domain with the exception of some stations along the western domain boundary. As evident from Figure 2a, most of this over-estimation is occurring during the later months of the forecasting period. For $PM_{2.5}$, Figures 4a-c illustrate that observed concentrations were highest in a belt from Indiana through Pennsylvania and in the Southeast. ETA/CMAQ captures the relatively low $PM_{2.5}$ concentrations during this time period in the Northeast, but strongly under-estimates $PM_{2.5}$ concentrations in other parts of the modeling domain. While the time series of domain-wide observed and predicted $PM_{2.5}$ averages shown in Figure 2b suggests very good model performance for September 2004, additional maps of average observed and predicted $PM_{2.5}$ concentrations for September 2004 (not shown here) indicate that this is a result of canceling errors, with a tendency for CMAQ to over-predict $PM_{2.5}$ in the Northeast and under-predict $PM_{2.5}$ in the South during September, though the under-prediction in the South is less severe than for other months. Table 2a presents a summary of discrete model evaluation statistics computed over all monitors for the entire simulation period from July 1 – September 30, 2004 for both ozone and $PM_{2.5}$. The table confirms the general tendency for this ETA/CMAQ simulation to over-predict 8-hr daily maximum ozone concentrations and under-predict 24-hr average $PM_{2.5}$ concentrations. It is also noteworthy that the absolute error for ozone (14.4 ppb) represents a 33% error when normalized by the average observed value while the relative absolute error for $PM_{2.5}$ is 63%. The better model performance for O_3

compared to $PM_{2.5}$ is consistent with continued uncertainties in the simulation of primary and secondary fine particulate matter in current-generation photochemical modeling systems.

When evaluating model performance using the categorical metrics listed in Table 1a-b, we applied two separate thresholds each for ozone and PM2.5, corresponding to AQI values of 50 (transition from good to moderate) and 100 (transition from moderate to unhealthy for sensitive groups). The results are presented in Table 2b. Depending on pollutant and threshold, the false alarm rates range from 25.6% to 85.4%, accuracy (which measures both correctly-predicted exceedances and non-exceedances) ranges from 71.9% to 98.7%, the probability of detection ranges from 6.5% to 58%, and the critical success index ranges from 4.7% to 39.8%. For comparison, in a past evaluation study for regression-type forecasting in California, Dye et al., 2000 reported values of ~85-90% for accuracy, ~70% for the probability of detection, and ~40% for the false alarm rate. McHenry et al. (2004) reported a probability of detection of 49%, a false alarm rate of 13%, an accuracy of 80%, and a critical success index of 34% for MAQSIP-RT forecasts at 67 monitors in New England during a pollution episode from August 1 – 10, 2001.

Overall, the results presented in this section indicate that the performance of the ETA/CMAQ modeling system for ozone forecasts during the summer of 2004 is within the range of other numerical forecasting systems. Model performance for $PM_{2.5}$ was worse than for ozone, but no $PM_{2.5}$ forecasts from other modeling systems were available for comparison to establish whether this reflects potential problems with the particular model setup used in this study or a general uncertainty stemming from the current state-of-science. It should also be noted that the time period analyzed in this study was marked by cooler and wetter than average conditions. As discussed by Mathur et al. (2004), the tendency of the operational ETA/CMAQ system utilized by NWS/NOAA/EPA to over-estimate the frequently-observed low ozone concentrations as indicated in Figure 2b might point to uncertainties in the specification of boundary conditions and the representation of cloud mixing and the effects of clouds on attenuation of photolysis rates. Since the setup of the modeling system in this study is identical to the operational system utilized by NWS/NOAA/EPA except for the use of different boundary conditions, this provides the opportunity for future research to study the impact of boundary conditions on predicted ozone concentrations by comparing the two sets of simulations.

4.2. Comparison between different forecast methods for New York State for August 2004

In this section, we compare the ETA/CMAQ O_3 and $PM_{2.5}$ forecasts for August 2004 to three other forecast methods for New York State. The first such method is the official next-day air quality forecast that is issued on weekdays by the NYSDEC for O_3 and $PM_{2.5}$ for eight regions in New York State based on statistical techniques, climatology, and expert judgment. The other two forecast methods investigated here are climatology and persistence. We followed the same 8-region approach as the official NYSDEC forecasts for the ETA/CMAQ forecasts, persistence and climatology, and only days in which all forecasts were available were included in the analysis, thereby excluding Sundays and Mondays for which no official next-day forecasts were available. Figures 5a-b show the mean absolute error for the different forecast methods for each of the eight forecast regions for both ozone and $PM_{2.5}$. For ozone, both ETA/CMAQ and the NYSDEC next-day forecast have a lower mean absolute gross error than forecasts based on climatology or persistence in most regions. For $PM_{2.5}$, the ETA/CMAQ forecasts have a lower mean absolute gross error than climatology for seven out of eight regions, while the NYSDEC forecasts have a lower error in all regions. Compared to the persistence forecast, the ETA/CMAQ forecasts have a lower error in six out of eight regions, while the NYSDEC forecasts have a lower error in four out of eight regions. The highest error for ETA/CMAQ $PM_{2.5}$ forecasts is observed in region 2, i.e. the New York City metropolitan area where the ETA/CMAQ strongly over predicts observed concentrations. This may point to potential problems with the treatment of primary aerosols, secondary organic aerosols, and the treatment of vertical mixing in this region of strong land/water contrasts in many adjacent grid cells. These issues will be the subject of future investigations. It should also be reiterated that this analysis was only based on one month of data, and that the summer of 2004 was cooler and wetter than usual. Therefore, it is instructive to compare the NYSDEC forecast performance for August 2004 to historic records. While the mean absolute error for the next-day NYSDEC forecasts for August 2004 ranged from 7 ppb to 16 ppb across the eight regions (Figure 5a), the error was 8-12 ppb for the summer of 2002 and 7-11 ppb for the summer of 2003.

In summary, while both NYSDEC and ETA/CMAQ forecasts clearly show better forecast skill than climatology and persistence for ozone, the picture is less clear for $PM_{2.5}$ forecasts. This probably is indicative of the longer experience of NYSDEC forecasters with ozone predictions compared to $PM_{2.5}$ predictions and the continued need to improve the state-of-science in simulating primary and secondary fine particulate matter in current-generation photochemical modeling systems.

4.3. Re-Simulation of a Pollution Event in July 2004

As described in Section 3, the routine ETA/CMAQ air quality forecasts rely on ETA forecast hours 16-40. Assuming that errors in forecasted meteorological fields increase with increasing forecast duration, we performed an experiment to assess the effects of forecast errors on air quality predictions. To this end, we re-simulated the period from July 2 – July 25 utilizing only forecast hours 1-24 from each ETA forecast cycle to perform subsequent 24-hr CMAQ simulations. As discussed in Section 3, the ETA forecast hours are almost passed by the time the CMAQ simulations are performed and, therefore, are of little interest from a forecasting perspective . Of specific interest in this re-simulation experiment is the time period from July 20 – 24 which was characterized by high observed ozone and especially $PM_{2.5}$ concentrations over large portions of the Northeastern U.S., including the western portions of New York State. Analysis of the original ETA/CMAQ forecasts presented above showed that predicted $PM_{2.5}$ concentrations were well below observed levels for this region during the episode. Figure 6 shows time series of observed and predicted hourly $PM_{2.5}$ concentrations averaged over 5 monitors in western New York State. It becomes obvious that neither the original forecast utilizing ETA forecast hours 18-42 nor the re-simulation utilizing ETA forecast hours 1-24 captured the amount of $PM_{2.5}$ buildup that occurred between July 20 and July 23. Overall, the differences in predicted $PM_{2.5}$ between these two simulations are very small. This is further confirmed by the discrete and categorical evaluation statistics for both simulations presented in Tables 3a-b. Although no test for statistical significance was performed, the similarity of model performance between both simulations suggests that possible errors in forecasted meteorological fields were a relatively minor contributor to overall model error for $PM_{2.5}$ air quality forecasts during this episode.

5. SUMMARY

In this study, O_3 and $PM_{2.5}$ predictions from the ETA/CMAQ air quality forecasting system applied in a pilot study between NOAA, EPA and NYSDEC were compared to available observations and other forecast methods for the summer of 2004. Results indicate that the performance of the ETA/CMAQ modeling system for ozone forecasts during the summer of 2004 is within the range of other numerical forecasting systems. Model performance for $PM_{2.5}$ was worse than for ozone, but no $PM_{2.5}$ forecasts from other modeling systems were available for comparison to establish whether this reflects potential problems with the particular model setup used in this study or a general uncertainty stemming from the current state-of-science. Furthermore, while both non-model based NYSDEC and ETA/CMAQ forecasts clearly show better forecast skill than climatology and persistence for ozone, the picture is less clear for $PM_{2.5}$ forecasts. This probably is indicative of the longer experience of NYSDEC forecasters with ozone predictions compared to $PM_{2.5}$ predictions and the continued need to improve the state-of-science in simulating primary and secondary fine particulate matter in current-generation photochemical modeling systems. A re-simulation of a high $PM_{2.5}$ concentration event suggests that possible errors in forecasted meteorological fields were a relatively minor contributor to overall model error for $PM_{2.5}$ air quality forecasts during this episode. The results of this study illustrate that photochemical modeling systems can be useful tools for air quality forecasting, but additional work is needed to improve model performance for $PM_{2.5}$. In addition, outputs from real-time air quality models such as ETA/CMAQ could potentially be useful for other objectives. For example, the potential ability of the photochemical models to provide air quality characterization at higher spatial and temporal resolution than possible with the current monitoring network might be of interest for research projects studying the linkages between air quality and human health, especially if advanced statistical methods are applied to combine the ambient monitoring data with such air quality model outputs.

6. DISCLAIMER

The research presented here was performed under the Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number DW13921548. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views. Furthermore, the views expressed in this paper do not necessarily reflect the views or policies of the New York State Department of Environmental Conservation.

7. REFERENCES

Black, T., 1994: The new NMC mesoscale Eta Model: description and forecast examples. *Wea. Forecasting*, **9**, 265-278.

Byun, D.W. and Ching, J.K.S. (eds.), 1999. Science algorithms of the EPA Models-3 Community Multiscale Air Quality Model (CMAQ) modeling system. EPA/600/R-99/030, U. S. Environmental Protection Agency, Office of Research and Development, Washington, DC 20460.

Cai, C., C. Hogrefe, J. Biswas, S.T. Rao, N. Seaman, A. Gibbs, G. Kallos, P. Katsafados, C. Walcek, J. Chang, and K. L. Demerjian, 2002: An experimental Air Quality Forecast Modeling System (AQFMS) for the northeast United States: a demonstration study. Preprints, *Fourth Conference on Atmospheric Chemistry: Urban, Regional, and Global-Scale Impacts of Air Pollutants*, Orlando, FL, Amer. Meteor. Soc.

Cardelino, C. A., M. Chang, J. St John, B. Murphey, J. Cordle, R. Ballagas, L. Patterson, K. Powell, J. Stogner, S. Zimmer-Dauphinee., 2001: Ozone predictions in Atlanta, Georgia: analysis of the 1999 ozone season, *J. Air Waste Manag. Assoc.*, **51**, 1227-1236.

Chang, M. E., and C. A. Cardelino, 2000: An application of the Urban Airshed Model to forecasting next-day peak ozone concentrations in Atlanta, Georgia, *J. Air Waste Manag. Assoc.*, **50**, 2010 – 2024.

Davidson, P.M., N. Seaman, K. Schere, R.A.Wayland, J.L. Hayes and K.F. Carey, 2004: National Air Quality Forecasting Capability: First Steps toward Implementation. Preprints, 6th Conference on Atmospheric Chemistry: Air Quality in Megacities. Seattle, WA, Jan 11-15, 2004.

Dye, T., C. P. MacDonald, and C. B. Anderson, 2000: Air quality forecasting for the spare the air program in Sacramento, California: summary of four years of ozone forecasting, Preprints, *11th Joint Conference on the Applications of Air Pollution Meteorology with the A&WMA*, Long Beach, CA, Amer. Meteor. Soc, 396 – 402.

Grell, G. A., J. Dudhia, and D. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5), *NCAR Technical Note*, NCAR/TN-398 + STR.

Mathur, R. J. Pleim, T. Otte, K. Schere, G. Pouliot, J. Young, B. Eder, D. Kang, S. Yu, H.-M. Lin, J. McQueen, P. Lee, M. Tsidulko, and D. Wong, 2004: Adaptation and Applications of the Community Multiscale Air Quality (CMAQ) Modeling System for Real-Time Air Quality Forecasting During the Summer of 2004, presented at the Models-3 conference, October 18-20, Chapel Hill, NC, available online at http://www.cmascenter.org/html/2004_workshop/abstracts/Forecasting/mathur_abstract.pdf

McHenry, J. N., N. Seaman, C. Coats, D. Stauffer, A. Lario-Gibbs, J. Vukovich, E. Hayes, and N. Wheeler, 2000: The NCSC-PSU numerical air quality prediction project: initial evaluation, status, and prospects, Preprints, *11th Joint Conference on the Applications of Air Pollution Meteorology with the A&WMA*, Long Beach, CA, Amer. Meteor. Soc, 403 – 408.

McHenry, J.N., W.F. Ryan, N. Seaman, C. Coats, J. Pudykiewicz, S. Arunachalam, and J. Vukovich, 2004: A real-time eulerian photochemical model forecast system: Overview and initial ozone forecast performance in the Northeast U.S. Corridor, Bull. Amer. Met. Soc., **85**, 525-548

McQueen, J., P. Lee, M. Tsidulko, G. DiMego, R. Mathur, T. Otte, J. Pleim, G. Pouliot, D. Kang, K. Schere, J. Gorline, M. Schenk, and P. Davidson, Update of the Eta-CMAQ forecast model run at NCEP operations and its performance for the summer 2004 presented at the Models-3 conference, October 18-20, Chapel Hill, NC, available online at http://www.cmascenter.org/html/2004_workshop/abstracts/Forecasting/mcqueen_abstract.pdf

Otte, T.L., G. Pouliot, and J.E. Pleim, 2004: PREMAQ: A New Pre-Processor to CMAQ for Air Quality Forecasting, presented at the Models-3 conference, October 18-20, Chapel Hill, NC, available online at http://www.cmascenter.org/html/2004_workshop/abstracts/Forecasting/otte_abstract1.pdf

Pleim, J.E., and R. Mathur, 2005: Diagnostic evaluation, sensitivity analyses, and new development in the Eta/CMAQ air quality forecast system, *Seventh Conference on Atmospheric Chemistry*, San Diego, CA, Amer. Meteor. Soc.

Pierce, T., C. Geron, G. Pouliot, E. Kinnee, and J. Vukovich, 2002: Integration of the Biogenic Emission Inventory System (BEIS3) into the Community Multiscale Air Quality Modeling System, Preprints, *12th Joint Conf. on the Apps. Of Air Pollut. Meteor. with the A&WMA*, Amer. Meteor. Soc., Norfolk, VA, 20-24 May 2002, J85- J86.

Pouliot, G., S. He, and T. Pierce, 2003: Recent advances in the modeling airborne substances, presented at the Models-3 conference, Research Triangle Park, NC, October 27-29, 2003, available online at http://www.cmascenter.org/2003_workshop/session2/pouliot_abstract.pdf

Pouliot, G., 2005: The emission processing system for the ETA/CMAQ air quality forecast system, *Seventh Conference on Atmospheric Chemistry*, San Diego, CA, Amer. Meteor. Soc.

Ryan, W. F., C. Piety, and E. D. Luebehusen, 2000: Air quality forecasts in the Mid-Atlantic region: current practice and benchmark skill, *Weather and Forecasting*, **15**, 46 – 60.

U.S. EPA, 1999: Guideline for developing an ozone forecasting program, EPA-454/R-99-009, United States Environmental Protection Agency, Research Triangle Park, NC 27711.

Vaughan, V., B. Lamb, C. Frei, R. Wilson, C. Bowman, C. Figueroa-Kaminsky, S. Otterson, M. Boyer, C. Mass, M. Albright, J. Koenig, A. Collingwood, M. Gilroy, and N Maykut, 2004: A numerical daily air quality forecast system for the Pacific Norhwest, Bull. Amer. Met. Soc., **85**, 549-561

Table 1a. Scheme used for defining the quantities “A” – “D” used in Table 1b.

		Predictions	
		No exceedance	Exceedance
Observations	No exceedance	A (Model correctly predicted no exceedance)	B (Model predicted an exceedance that did not occur)
	Exceedance	C (Model failed to predict an exceedance that occurred)	D (Model correctly predicted an exceedance)

Table 1b. Forecast evaluation metrics as defined in EPA (1999)		
Accuracy (%)	Percent of forecasts that were correct	100 * (A+D)/(A+B+C+D)
False Alarm Rate (FAR) (%)	Percent of forecasted exceedances that did not occur	100 * B/(B+D)
Probability of Detection (POD) (%)	Percent of observed exceedances that were forecasted correctly	100 * D/(C+D)
Critical Success Index (CSI) (%)	Measures how well high ozone events are predicted (not influenced by number of correct non-exceedance forecasts)	100 * D/(B+C+D)

Table 2a. Domain-wide discrete evaluation statistics for ETA/CMAQ forecasts July – September, 2004					
	Observed Average	CMAQ Average	Bias	RMSE	Absolute Error
8-hr Daily Maximum Ozone (ppb)	43.9	52.2	8.3	11.3	14.4
24-hr $PM_{2.5}$ (TEOM) (µg/m3)	14.7	11.2	-3.6	6.8	9.3

Table 2b. Domain-wide categorical evaluation statistics for ETA/CMAQ forecasts July – September, 2004				
	False Alarm Rate (%)	Accuracy (%)	Probability of Detection (%)	Critical Success Index (%)
8-hr Daily Maximum Ozone, threshold 64 ppb (AQI 50)	66.7	85.7	58.0	26.9
8-hr Daily Maximum Ozone, threshold 84 ppb (AQI 100)	82.6	98.7	40.8	13.9
24-hr $PM_{2.5}$ (TEOM), threshold 15 µg/m3 (AQI 50)	25.6	71.9	46.1	39.8
24-hr $PM_{2.5}$ (TEOM), threshold 40 µg/m3 (AQI 100)	85.4	97.6	6.5	4.7

Table 3a. Domain-wide discrete evaluation statistics for ETA/CMAQ standard (ST) and re-run (RE) forecasts July 2 - 24										
	Observed Average	CMAQ Average		Bias		RMSE		Absolute Error	
		ST	RE	ST	RE	ST	RE	ST	RE
8-hr Daily Maximum Ozone (ppb)	48.8	52.8	52.9	4.0	4.1	9.2	8.7	12.0	11.4
24-hr $PM_{2.5}$ (TEOM) (µg/m3)	16.4	9.9	10.1	-6.5	-6.3	8.1	8.0	10.9	10.6

Table 3b. Domain-wide categorical evaluation statistics for ETA/CMAQ standard and re-run forecasts July 2 – 24								
	False Alarm Rate (%)		Accuracy (%)		Probability of Detection (%)		Critical Success Index (%)	
	OP	ST	OP	ST	ST	RE	ST	RE
8-hr Daily Maximum Ozone, threshold 64 ppb (AQI 50)	48.4	45.6	86.5	87.4	54.5	59.4	36.1	39.7
8-hr Daily Maximum Ozone, threshold 84 ppb (AQI 100)	60.5	57.5	98.1	98.3	45.9	45.9	26.9	28.3
24-hr $PM_{2.5}$ (TEOM), threshold 15 µg/m3 (AQI 50)	16.5	17.3	64.7	65.3	30.9	32.9	29.1	30.7
24-hr $PM_{2.5}$ (TEOM), threshold 40 µg/m3 (AQI 100)	58.8	57.1	96.8	96.8	7.6	6.5	6.9	6.0

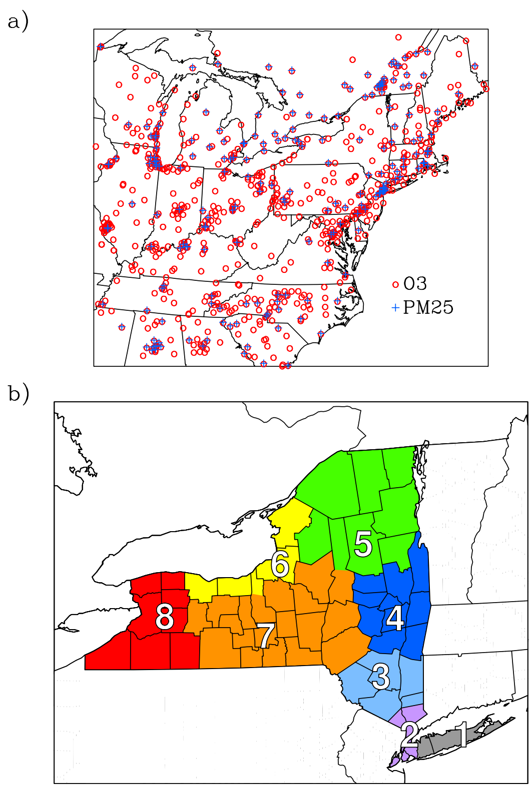


Figure 1: (a) Map of the ETA/CMAQ forecast domain with the locations of all AIRNOW O_3 and $PM_{2.5}$ monitors used in this analysis, (b) map of the eight forecast regions in New York State used in Section 4.2.

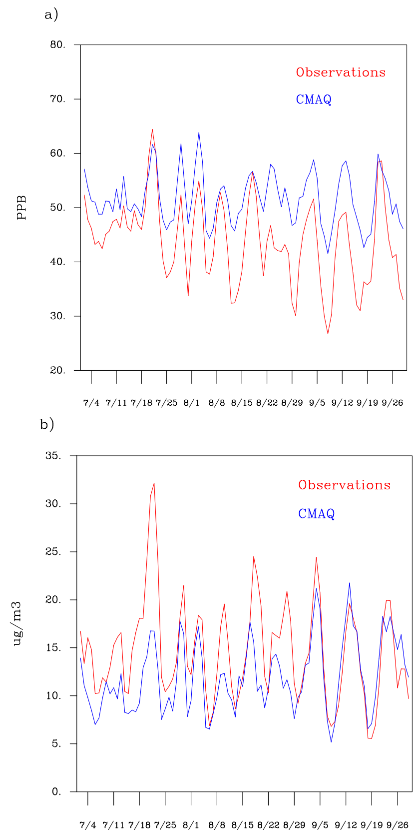


Figure 2: Time series of observed and predicted domain-averaged (a) daily maximum 8-hr ozone and (b) 24-hr average $PM_{2.5}$ concentrations for the time period from July 1 – September 30, 2004.

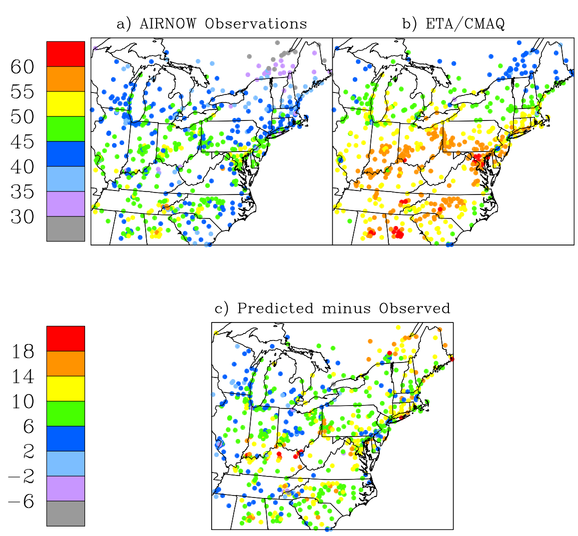


Figure 3: Maps of (a) observed and (b) predicted daily maximum 8-hr ozone concentrations averaged from July 1 to September 30, 2004, and (c) differences between the average predicted and observed concentrations. All values are given in ppb.

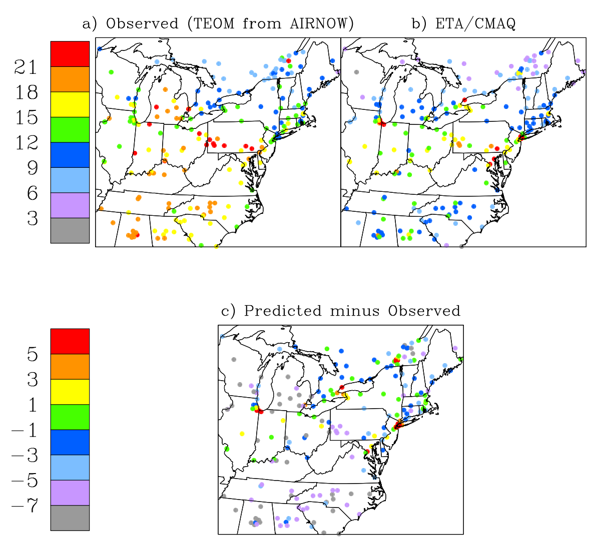
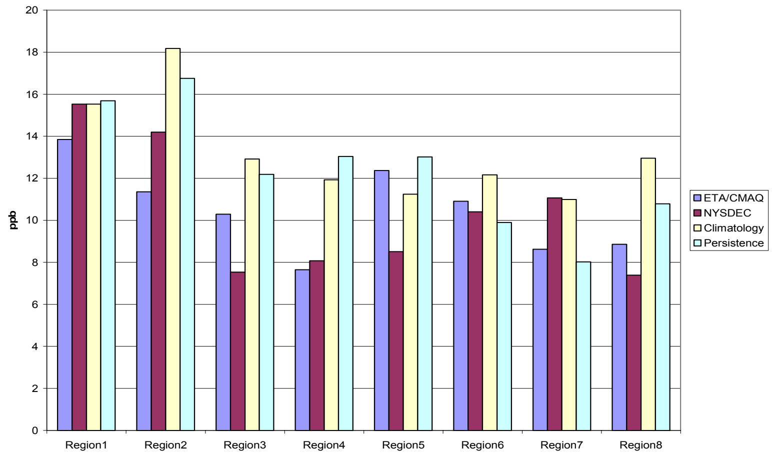


Figure 4: Maps of (a) observed and (b) predicted 24-hr average $PM_{2.5}$ concentrations averaged from July 1 to September 30, 2004, and (c) differences between the average predicted and observed concentrations. All values are given in μg/m^3.

a)



b)

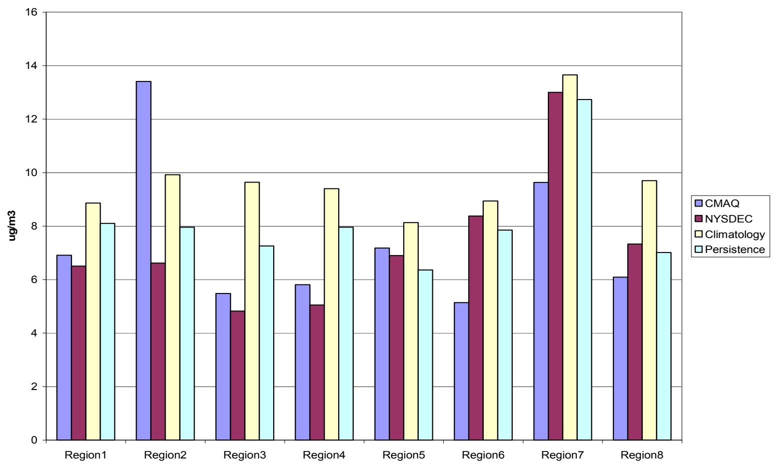


Figure 5: Mean absolute gross error (MAGE) for the four different forecast methods discussed in Section 4.2 for the month of August 2004 for each of the eight New York State forecast regions shown in Figure 1b. (a) MAGE for predicted daily maximum 8-hr ozone concentrations in ppb, (b) MAGE for predicted 24-hr $PM_{2.5}$ concentrations in μg/m^3.

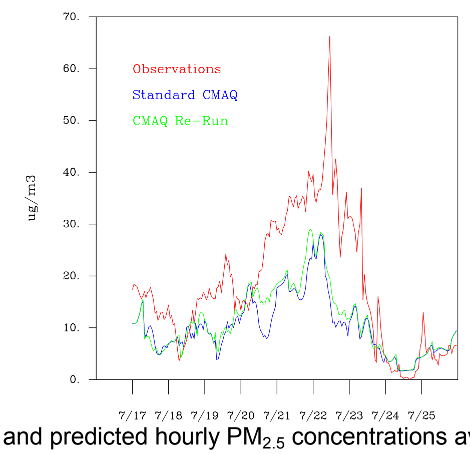


Figure 6: Time series of observed and predicted hourly $PM_{2.5}$ concentrations averaged over 5 monitors in western New York State. The two lines showing predicted $PM_{2.5}$ concentrations represent the standard ETA/CMAQ forecast (blue line) and the CMAQ re-simulation (green line) utilizing ETA forecast hours 1-24 as described in Section 4.3.