4.12

EVALUATION OF THE NATIONAL AIR QUALITY FORECAST SYSTEM (NAQFS) DEVELOPMENTAL LARGE DOMAIN MODEL: SUMMARY OF THE AIR QUALITY FORECASTERS FOCUS GROUP WORKSHOP

William F. Ryan
The Pennsylvania State University, University Park, Pennsylvania

Paula Davidson
Office of Science and Technology, National Weather Service, Silver Spring, Maryland

Paul Stokols
Office of Climate, Weather and Water Services, National Weather Service, Silver Spring, Maryland

1. INTRODUCTION

The National Oceanic and Atmospheric Administration (NOAA), in association with the United States Environmental Protection Agency (USEPA), have embarked on a National Air Quality Forecasting (NAQF) program. The development of a national air quality forecast capability was directed by Congress (Energy Policy Act of 2002). The vision of the NAQF program is to provide ozone (O_3), fine particulate matter ($PM_{2.5}$) and other pollutant forecasts with sufficient accuracy and advance notice to allow actions to be taken to prevent or reduce adverse health effects. The strategy to achieve this vision calls for NOAA to work cooperatively with EPA and state and local air quality agencies to develop end-to-end air quality forecasting capabilities. This cooperative relationship recognizes the responsibility of state and local governments to develop pollution control strategies, based on EPA guidance, and their traditional role as provider of health warnings. In this case, NOAA will be providing a tool for state and local forecasters to provide accurate and timely health warnings. The implementation plan calls for the initial development and implementation of a 1-day O_3 forecast model for the northeastern US by the fourth quarter of FY 2004. This model will then be extended to the entire US by FY 2009. In the longer term, 5-10 years, a forecast model for $PM_{2.5}$) will be developed and the forecast lead-time will be extended to two days or beyond, as accuracy and resources permit.

In pursuit of the development of an air quality forecast capability, a prototype air quality forecast (AQF) model was tested in the northeastern US in the summer of 2003. As part of the evaluation and development of the model, a focus group was convened to review and comment on test results: both for accuracy in their respective forecast areas and utility as forecast guidance. The National Weather Service (NWS) convened a two day workshop for the focus group and model system developers in September, 2003 to discuss the results. A summary of the Focus Group Workshop in 2003 was presented at the 84th AMS Annual Meeting (Ryan, et al., 2004). The success of the focus group as a forum for the interchange of ideas between operational forecasters and model developers, and continuing model development efforts, led to its expansion in 2004. This paper reports on the activities of the focus group in 2004.

* *Corresponding author address*: William F. Ryan, Pennsylvania State University, Department of Meteorology, 409 Walker Building, University Park, PA, 16802, email: wfr1@psu.edu.

2. THE FOCUS GROUP

The AQF focus group is composed of air quality forecasters from state and local air quality agencies, as well as academic researchers and private sector forecasters who contributed their knowledge and experience with local air quality forecasting issues. A list of members and their affiliations is given in Table 1.

The focus group members provided daily feedback to NOAA and EPA on model performance and utility. The responses ranged from very local considerations (metropolitan area forecasts), to regional (e.g., New England), to domain wide. Members of the focus group also provided visualization products to assist in the discussion. The members of the focus group met in Silver Spring Maryland on September 8-9, 2004 with EPA and NOAA model developers to discuss results and recommend further actions.

3. THE AIR QUALITY FORECAST MODEL

Details of the air quality forecast model and its evaluation will be provided in other papers (e.g., Davidson et al., 2005; Pleim and Mathur, 2005). A brief description of the developmental model system as analyzed by the focus group and the improvements made during 2004 will be provided here. The focus group was charged with analyzing forecast guidance provided by the developmental version of the NAQF model system. The *developmental* version of the model was different in certain respects from the *experimental* version that was run both in 2003 and 2004 and was recently (September, 2004) approved for deployment into NWS operations (see, Davidson et al., 2005). The two models, developmental and experimental, use a similar suite of model components. Both versions use NWS weather observations and the NCEP mesoscale numerical weather prediction model (Eta-12) to provide the meteorological drivers for the EPA Community Model for Air Quality (CMAQ). While there are additional differences in model configuration discussed below, the key difference between model versions is the expanded domain used by the developmental model (Figure 1). The size of the domain is increased so that roughly three times as many grid cells are contained in the developmental version. As a result, the developmental model is often referred to as the "3x" model. Forecasts for both versions were run twice daily, initialized at 1200 UTC (primary forecast) and run for 48 hours, and at 0600 UTC to support morning updates of current forecasts and provide initial fields for the 1200 UTC runs.

As originally intended, CMAQ is used to analyze historical pollution episodes and is run without operational time constraints. In a forecasting application, time constraints are imposed by the need to use the most recent weather forecast data and provide next-day guidance output to state and local forecasters in a timely manner. Specifically, the system uses the 1200 UTC Eta-12 forecast cycle output to drive the air quality prediction modules and provides air quality forecast output no later than 1730 UTC. For operational forecasting use, CMAQ had to be simplified and optimized to decrease run-time. Modifications to the model system required a trade off between increased computational speed and decreased generality and flexibility. The key additions and modifications included: PREMAQ – a module that pre-processes emissions inventory data with Eta-12 forecast weather data, and simplifications and optimization of CMAQ to meet the run-time requirements. These included a variety of changes to the chemistry model including: simplification of the Carbon Bond-4 chemical mechanism, disabling aerosol formation modules, streamlining the biogenic module (BEIS) to hardwire chemical speciation, dropping transport terms for fast reacting radicals, and pre-calculation of mobile emissions coupled with a temperature adjustment. In addition, augmentation of NCEP's central computing system (IBM-SP, a massively parallel platform) was necessary to provide the run-time window needed for air quality modeling.

Based on an analysis of the 2003 air quality forecast model performance, that showed a tendency to over-predict O_3, particularly in the southeastern US (McQueen et al., 2004; Ryan, et al., 2004), a number of changes and improvements were made to the forecast systems for both the experimental and developmental models. First, an analysis of the Eta-CMAQ linkage exposed systematic errors in the specification of land use as well as temperature interpolation errors that affected boundary layer heights. These errors were corrected and retrospective analyses showed that over-prediction was decreased significantly when corrections were applied (Figure 2). Improvements and upgrades were also made to area (2001 National Emissions Inventory, Version 3) and point source emissions with the Mobile 6 mobile source emissions model implemented to create retrospective mobile emissions fields. Biogenic emissions estimates were developed using BEIS 3.12. Changes were also made to boundary layer and turbulence modules. Planetary boundary layer (PBL) heights are now determined directly from Eta-12 output and a new scheme for the specification of minimum K_z was added. The K_z values now vary depending on the urban land use fraction in each cell. This mimics the urban heat island effect during the nighttime hours.

Beyond the change in domain size, the other distinguishing difference in the developmental (3x) model is the specification of the lateral boundary conditions (LBC). A "clean", or low O_3, default vertical) O_3 profile is used in the experimental model while the developmental model utilized a LBC O_3 profile from the NCEP GFS model intitialized with Solar Backscatter Ultra-Violet (SBUV-2) satellite observations. This approach sought to simulate variations in O_3 driven by dynamic conditions. In practice, the GFS-based profiles were higher in O_3 than the default profiles and translated to systematically higher biases in the 3x domain. As a result of preliminary evaluations, the 3x LBC was changed on June 30 to limit the influence of GFS-based profiles to heights above 6 km. The seasonal LBC was about 3-5 ppbv higher than the default (40 ppbv) through the depth of the boundary layer. The default value was used as the boundary layer LBC for the experimental (1x) model domain. Comparison of prediction biases for the overlap region in the developmental (3x) and experimental (1x) domains revealed a systematic bias that was 3-5 ppbv higher for the 3x domain – about the same magnitude as the difference in boundary layer LBC – suggesting a connection. To reduce the 3x bias, its boundary layer profle was changed to match the default (1x LBC) on August 3rd. Note that in all cases, the GFS-based profiles were used above 6 km.

4. FOCUS GROUP ACTIVITIES: VISUALIZATION AND EVALUATION

The evaluation undertaken by focus group members was of the developmental (3x) version of the NAQF model initialized at 1200 UTC verifying on the following day (i.e., 12-36 hour forecast). The 1200 UTC forecast output was typically available at 1730 UTC on the NCEP-NWS server. Output files contained surface O_3 concentrations in standard NWS gridded binary (GRIB) format.

To assist the focus group, model output was provided, in the form of hourly graphical images and animations, by NCEP's Environmental Modeling Center (EMC) at a password protected site. As in 2003, additional images and test results were provided at a Pennsylvania State University (PSU) website. At the PSU site, a set of Eta-12 forecast images accompanied the O_3 forecast. Sub-domain images were provided for the larger urban areas in North Carolina and along the I-95 Corridor in the northeast along with time series of O_3 for stations selected by the forecasters. Examples are provided in Ryan et a al., 2004. The goal of the PSU web site was to provide forecasters with the underlying meteorological forecast supporting the air quality forecast. The Eta-12 images were focused on boundary layer processes and included vertical time series of wind, potential temperature and relative humidity at selected stations as well as domain-wide fields of precipitation, 950 mb winds and temperature. Comparison plots of forecasted and observed O_3 were provided for the focus group by EPA's AIRNow data management center, with support from Sonoma Technology. These images allowed forecasters a glance at recent model biases.

Using these forecast images, focus group members were able to provide daily feedback on forecast model performance. An online worksheet was provided by NWS to allow forecasters to easily enter daily model feedback. The spreadsheet allowed for input on data availability and timing, a brief identification of key weather elements on both the synoptic scale and the mesoscale, as well as forecast and observed O_3 concentrations. In addition, comments on model performance were included. The feedback forms were collated by NOAA personnel and provided to focus group members to assist in group discussion at the September workshop.

5. METEOROLOGICAL CONDITIONS DURING THE EVALUATION PERIOD

Like the summer of 2003, the summer of 2004 featured extremely low O_3 compared to recent climatology. This was particularly true during July and August – the "heart" of the O_3 season. For the period June 1-September 30, the Mid-Atlantic Regional Air Managers Association (MARAMA), which collects O_3 data for the mid-Atlantic region, identified only 18 days (139 monitors) with O_3 concentrations in excess of the 85 ppbv 8-hour O_3 standard. Only three days (6 total monitors) were in excess of 125 ppbv for the 1-hour standard. By way of comparison, the single week of July 7-13, 2002 experienced 180 8-hour exceedances and 32 1-hour exceedances. In the Philadelphia metropolitan area, no days exceeded the 1-h O_3 health standard (125 ppbv) throughout the entire summer and only 7 days exceeded 85 ppbv for an 8-hour average. This represented the lowest frequency of high O_3 cases in Philadelphia since the modern monitoring network was installed in the early 1980's and likely for many decades prior.

The historically low O_3 concentrations in 2004 were driven by unusual summer season weather patterns. Throughout the summer months, a persistent large-scale circulation anomaly placed an upper level low over the Great Lakes (Figure 3). The presence of persistent low pressure aloft corresponded to frequent periods of rain and cooler than normal temperatures across most of the forecast domain (Figure 4). The combination of precipitation, cloud cover and intrusions of cool Canadian air was not conducive to O_3 formation.

6. FOCUS GROUP WORKSHOP AND EVALUATION SUMMARY

The AQF focus group convened in Silver Spring, MD on September 8-9, 2004 to hear a series of presentations by NOAA and EPA model developers on recent changes to the air quality forecast models and results from preliminary forecast evaluations. Focus group forecasters presented their observations on model performance in the local operational forecast arena and all attendees took part in a wide-ranging discussion of model performance and future directions for the NOAA air quality forecast program.

The NOAA Model Development Laboratory (W. Shaffer) and the EPA Air Resources Laboratory (B. Eder) presented preliminary model performance evaluations. Shaffer presented a comparison of 2003 and 2004 forecast model performance. As the 3x model was not in place in 2003, this comparison was limited to the experimental (1x) model domain. The 2004 results showed a large decrease in both bias and mean absolute error (MAE) (Figures 5 and 6). Comparing the 1x and 3x models, results were quite similar with the 3x (developmental) model showing less over-prediction in the nighttime hours and a slightly higher bias during the peak mid-afternoon hours. These differences are most likely due to the changes, noted above, in O_3 LBC's as well as slight variations between models in the minimum K_z settings. Contingency table metrics for color code forecast thresholds were also presented showing good probability of detection (POD) of 8-hour O_3 concentrations in excess of 85 ppbv (Code Orange or Unhealthy for Sensitive Groups) but at the cost of a high false alarm rate (FAR). The combination of high POD and FAR led to a low Threat Score (0.20).

Similar results were shown by Eder et al. In particular, the experimental (3x) model tends to over-predict in the very low O_3 cases (8 hour O_3 < 40 ppbv). This low end bias is common with standard statistical forecast guidance and typically reflects poor performance in precipitating or overcast conditions. For 8 hour O_3 concentrations in excess of 40 ppbv, the 3x model showed a bias of +5.4 ppbv (monthly averages) with an $r^2 = 0.48$ and an average rms error of 12 ppbv. Normalized mean error averaged monthly ranged from 15-19% and a high FAR was also noted. Although over-prediction in the southeastern US, a serious problem in 2003, was considerably reduced, it still remains present, particularly in warm temperature cases. A case study for August 4-12, a period featuring frontal passage followed by the slow modification of a continental polar (cP) air mass, showed overall good performance except in the vicinity of frontal boundaries and heavy cloud cover where significant over-prediction occurred.

Feedback from the operational air quality forecasters confirmed the results of the statistical evaluations and added additional insights. First, significant improvements in forecast performance for the 2004 season compared to 2003 was noted. Persistent over-prediction in the southeastern US was greatly reduced. Over-prediction did occur in the region but tended to be more episodic in nature. Two situations where over-prediction was noted by local forecasters in the southeastern US were easterly flow cases associated with cool air wedges and warm weather cases with widespread cumulus development. Problems with O_3 titration by excessive NO_x concentrations along the I-95 Corridor seen in 2003 (Figure 7) were also greatly reduced in 2004. This may reflect changes in the emissions inventory and changes to PBL height determinations.

Most forecasters noted that the developmental model showed good success in locating the plume of highest O_3 within their forecast area. Although skill in predicting the magnitude of peak O_3 within the plume varied, forecasters could rely on plume placement. This reflects skill in the meteorological drivers of the model and, knowing the location of the plume with some degree of confidence, forecasters could use local experience to make decisions as to the expected magnitude.

Related to the spatial skill of the model, several forecasters noted good results in stagnation cases. An example for the Philadelphia area is given below. Stagnation cases are very difficult for current statistical models to accurately resolve and increased skill in this type of case is of great value to operational air quality forecasters.

With the developmental model showing reasonable skill and accurately locating the high O_3 plume, the focus group forecasters expressed an interest in future development of post-processed products to assist in the interpretation of model output. In particular, short and long period model bias correction information and the development of a model output statistic (MOS) product. Although the AQF model will not likely be “frozen” for some period of time sufficient to support a standard MOS, the development of an updateable MOS might be possible.

7. APPLICATION TO LOCAL FORECASTS

Both model evaluations and forecaster feedback noted significant improvement in model performance in 2004. In this section, forecast performance in the Philadelphia (PHL) metropolitan area is analyzed to assess the utility of the model in operational use. The Philadelphia air quality forecast area includes the City of Philadelphia and portions of southeastern PA, southern NJ and northern DE and MD (Figure 8). In 2003, the air quality forecast model showed reasonably good skill in this area although there were recurrent problems with reduction in O_3 concentrations due to NO_x titration along the I-95 Corridor as well as unrealistically high O_3 concentrations (“bullseyes”) near land-sea boundaries (Ryan et al., 2004).

In 2004, the developmental model showed good skill in the PHL area and exhibited decreased influence of NO_x titration and local over-predictions. Six monitors within the PHL were tracked to investigate model skill (Figure 9). With the exception of the Ancora monitor in southern NJ, results were quite good with r^2 ranging from 0.51-0.65 and small biases. Domain peak O_3, a measure commonly used to verify local forecasts, was quite well forecast and close in skill to the current statistically based forecasts.

Although the summer of 2004 was very low in O_3, with no severe multi-day regional episodes, there were several high O_3 cases featuring stagnation or re-circulation. Historically, stagnation cases have proven very difficult to forecast due, in part, to their short duration, which limits the information that persistence can provide, and uncertainty with respect to plume placement. In the July 2-3 case, the PHL area sat in a featureless pressure gradient field between two weak frontal boundaries (Figure 10). Although the 3x model under-predicted peak concentrations, it did capture the upward trend (20 ppbv increase) and accurately located the plume along the I-95 Corridor with concentrations exceeding 100 ppbv (Figure 11). Forecast experience in the PHL area has been that, in the context of higher than average regional O_3 concentrations, stagnation can lead to Cod Orange or higher concentrations. Anticipating a slight under-prediction in peak concentrations from the forecast model and with regression-based forecasts of a similar magnitude, forecasters were able to accurately predict Code Orange 8-hour O_3 concentrations in this case.

The following day offered a more complex situation as a cold front was forecast to drop just south of PHL and stall. This is perhaps the most difficult forecast in this region as O_3 gradients can be quite steep in the vicinity of diffuse, slow moving frontal boundaries. The developmental model showed high O_3 continuing south of the boundary in the Baltimore-Washington Corridor with a slight decrease in the PHL area (Figure 12). The north-south O_3 gradient did occur, with high concentrations continuing a second day in Washington DC.

The July 21-22 short episode of high O_3 was the only episode during 2004 that developed according to the "standard" regional high O_3 scenario. The onset of the episode (July 21), featured sustained westerly transport and was well forecast. July 22nd posed a very common forecast problem for air quality: timing of convection ahead of a cold front and its impact on O_3. Because O_3 has a strong diurnal gradient, and decreases rapidly in the presence of convection, a difference of only a few hours in the onset of convection can lead to large (~ 10's of ppbv) changes in peak O_3. In this case, Eta-12 was reasonably accurate with frontal position but the forecast system did not accurately diagnose the fall in O_3 concentrations as pre-frontal convection and cloud cover moved into the region (Figures 13 and 14). In this case, O_3 was over-predicted in the Baltimore-Washington region but well predicted, although perhaps for the wrong reasons, in Philadelphia.

8. CONCLUSIONS

The deployment by NOAA of an operational numerical air quality forecast model this September marks a major milestone in air quality forecasting. Providing accurate numerical forecast guidance for O_3 is a difficult and demanding task. Not only must the full suite of meteorological outputs be accurately simulated, including moisture, radiation flux and winds, but a chemistry model must be integrated with the meteorology and a variety of O_3 precursor emissions, ranging from automobile exhaust to large power plants, must be adequately modeled.

The large increase in the number of cities and states issuing air quality forecasts has led, over the past decade, to the development of a cadre of experienced air quality forecasters. These forecasters, organized in a focus group under the auspices of NOAA, have provided valuable feedback to NOAA and EPA model developers on the skill of the model and on the products that will be of most utility for forecast preparation.

In 2004, the focus group expanded its membership to reflect the expanded domain of the new developmental forecast model. The focus group provided daily feedback to model developers and presented their results, along with NOAA and EPA modelers, at the second Focus Group Workshop in September, 2004. The focus group concluded that forecast model performance was greatly improved in 2004 compared to the previous year. Spatial orientation of the high O_3 plume was generally quite good and, in specific cases, led to greater confidence in forecasts issued.

ACKNOWLEDGEMENTS

The authors acknowledge the valuable feedback provided by all members of the focus group (Table 1) as well as NOAA, NWS and EPA personnel that provided presentations to the workshop including Brian Eder, Will Shaffer, Jeff McQueen, Jon Pleim and Rohit Mathur, their colleagues and co-authors.

REFERENCES

Davidson, P. et al., 2005: National air quality forecasting capability: Initial operational capability, 7th Conference on Atmospheric Chemistry, 4.10, American Meteorological Society, San Diego.

Davidson, P. et al., 2004: National air quality forecasting capability:First steps towards implementation, Air Quality in Megacities (Joint Symposium on Planning, Nowcasting and Forecasting in the Urban Zone and 6th Conference on Atmospheric Chemistry), J2.13, American Meteorological Society, Seattle.

McQueen, J. et al., 2004: Development and evaluation of the NOAA/EPA prototype air quality model prediction system, Air Quality in Megacities (Joint Symposium on Planning, Nowcasting and Forecasting in the Urban Zone and 6th Conference on Atmospheric Chemistry), J2.16, American Meteorological Society, Seattle.

Pleim, J. and R. Mathur, 2005: Diagnostic evaluation, sensitivity analysis, and new developments in the Eta/CMAQ air quality forecast system, 7th Conference on Atmospheric Chemistry, 4.7, American Meteorological Society, San Diego.

Ryan, W. F. et al., 2004: Evaluation of the National Air Quality Forecast System (NAQFS): Summary of the Air Quality Focus Group Workshop, Air Quality in Megacities (Joint Symposium on Planning, Nowcasting and Forecasting in the Urban Zone and 6th Conference on Atmospheric Chemistry), J2.13, American Meteorological Society, Seattle.

Table 1. Members of the air quality forecasters focus group.

Member	Affiliation
Joanne M. Alexandrovich	Vanderburgh County Health Department
Jerry Beasley	Mississippi Department of Environmental Quality
Robert Brawner	State of Tennessee, Air Pollution Control Division
George M. Bridgers	North Carolina Division of Air Quality
Ken Carey	Mitretek Systems
Joe Cassmassi	South Coast Air Quality Management District
Neal Conatser	Michigan Department of Environmental Quality
David Conroy	U.S. EPA, Region 1
Paula Davidson	National Weather Service
Tim Dye	Sonoma Technology, Inc.
Tammy Eagan	Florida Department of Environmental Protection
Pamela Frazier	State of Tennessee, Division of Air Pollution Control
Colleen Farrell	Environment Canada - Dartmouth
Mike Gilroy	Puget Sound Clean Air Agency
Alan Hansen	Electric Power Research Institute
James G. Haywood	Michigan Department of Environmental Quality
Bryan Lambeth	Texas Commission on Environmental Quality
Lisa Landry	State of New Hampshire, Dept of Environmental Services
Paul Martin	South Carolina Department of Health and Environmental Control
Anne McWilliams	U.S. EPA Region 1
Bill Murphey	Georgia Environmental Protection Division
Sean Nolan	Pennsylvania Department of Environmental Protection
William F. Ryan	The Pennsylvania State University
Dan Salkovitz	Virginia Department of Environmental Quality
Scott Southwick	Alabama Department of Environmental Meteorology
Paul Stokols	NWS Office of Climate, Weather and Water Services
Richard A. Wayland	US EPA/OAR/OAQPS (AirNOW)
Martha Webster	Maine Department of Environmental Protection
Dan White	Texas Commission on Environmental Quality
John E. White	US EPA/OAR/OAQPS (AirNOW)
Nick Witcraft	North Carolina Division of Air Quality
Jeanne Worthen	Connecticut Department of Environmental Protection

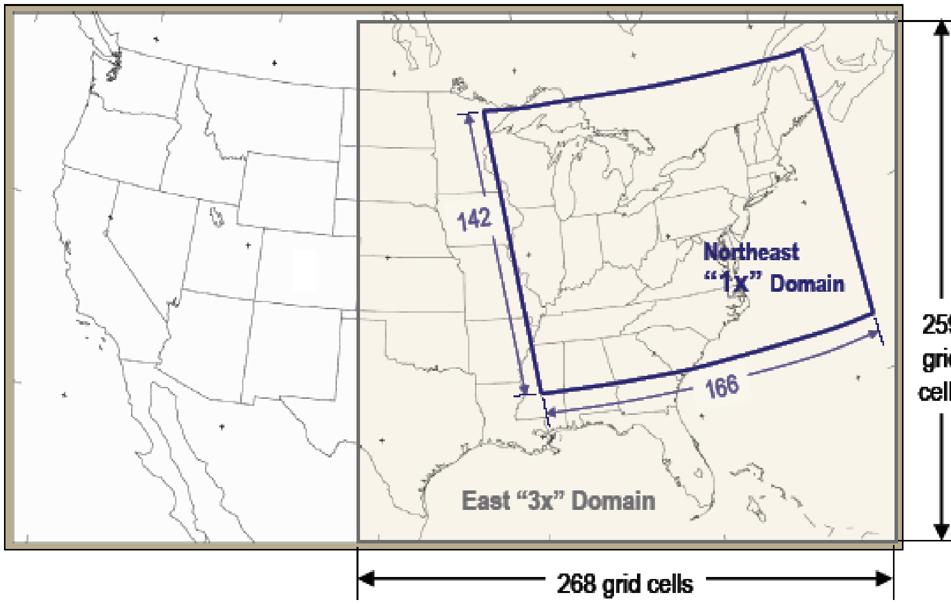

Figure 1. Domain of the developmental ("3x") NOAA Air Quality Forecast model as used in 2004. This domain contains approximately three times more grid cells than the experimental model domain .

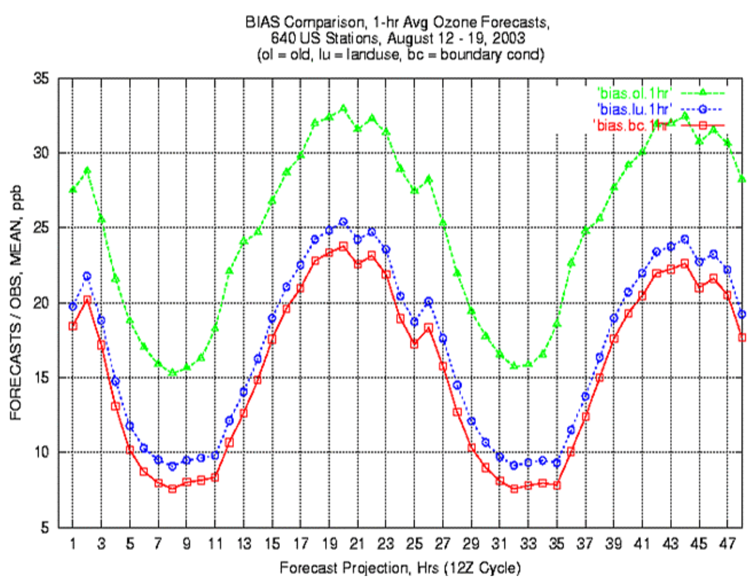

Figure 2. Bias comparison of three versions of the AQF experimental model for the period August 12-19, 2003 by forecast hour. Results from the 2003 model configuration are shown in green triangles, 2004 in blue circles and 2004 with GFS-based LBC in red squares. Figure courtesy of the NOAA Meteorological Development Laboratory (W. Shaffer,et al.).

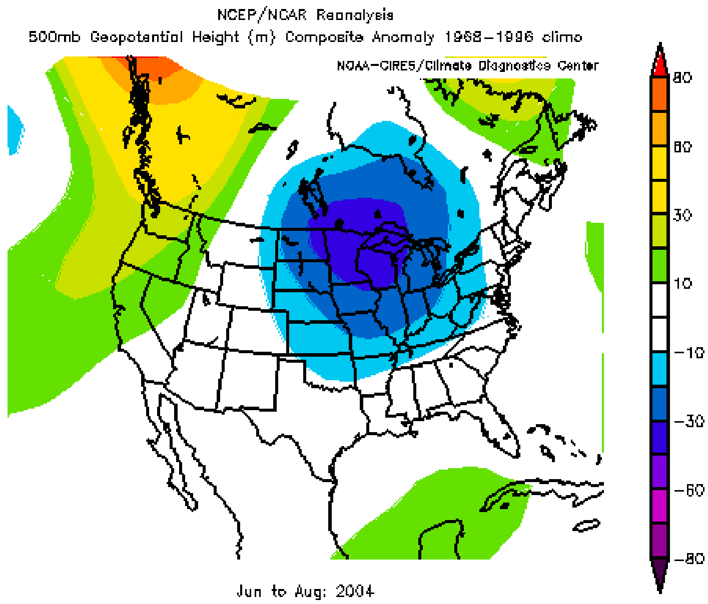

Figure 3. 500 mb geopotential height anomaly for June-August 2004 compared to recent climatology using the NCEP/NCAR re-analysis data. A significant negative height anomaly is present over the Great Lakes. Figure courtesy of NOAA-CIRES, Climate Diagnostics Center.

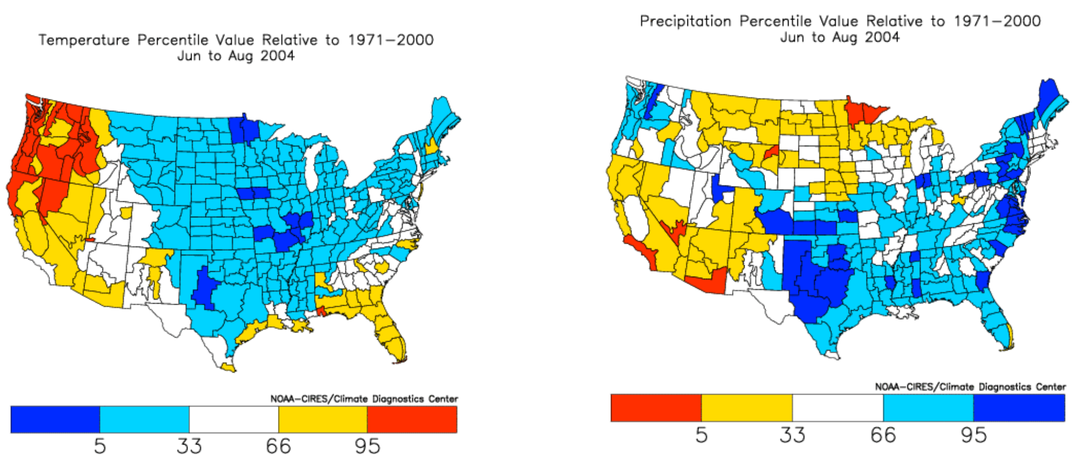

Figure 4. Percentile values of temperature (left panel) and precipitation (right panel) for June-August 2004 compared to recent (1971-2000) climatology. Figure courtesy of NOAA-CIRES, Climate Diagnostics Center.

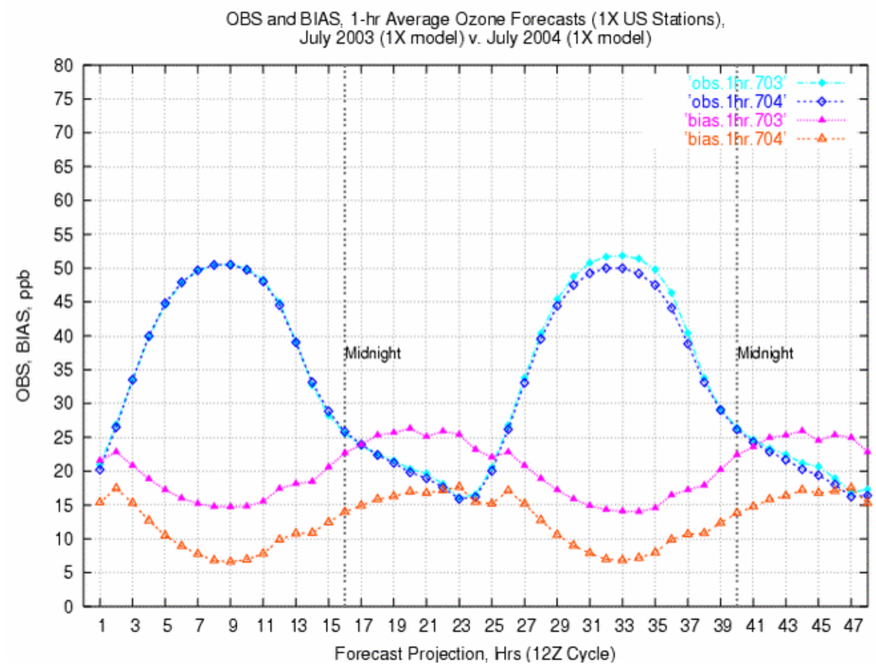

Figure 5. Comparison of experimental NOAA AQF model performance for July 2003 and July 2004 by forecast hour. Observed O_3 concentrations given by diamonds (light blue, 2003; dark blue, 2004) and forecast bias by triangles (purple, 2003; orange, 2004). Figure presented at the 2004 NOAA Focus Group Workshop by W. Shaffer et al., NOAA Model Development Laboratory.

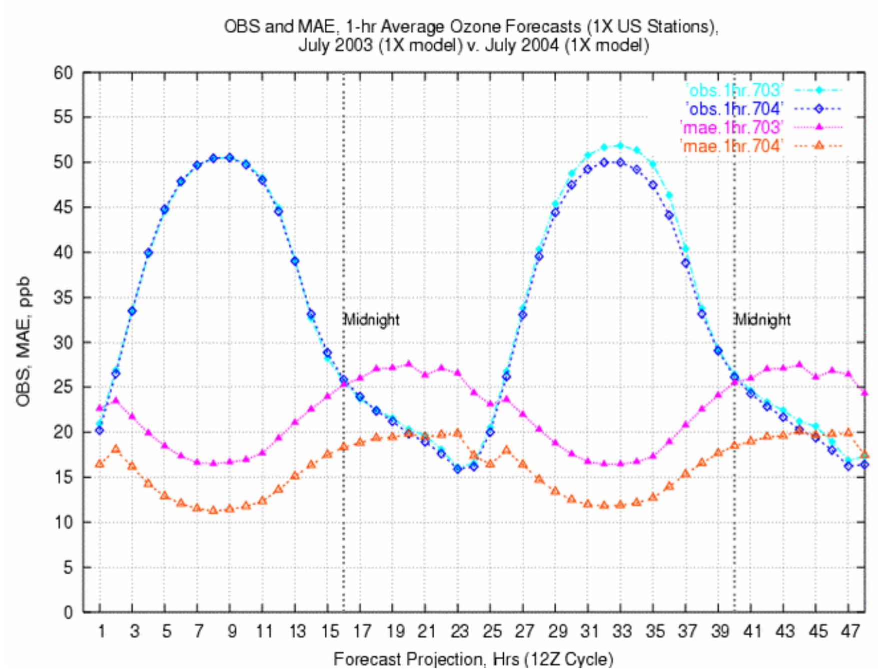

Figure 6. As in Figure 5 except that triangles give mean absolute forecast error (MAE).

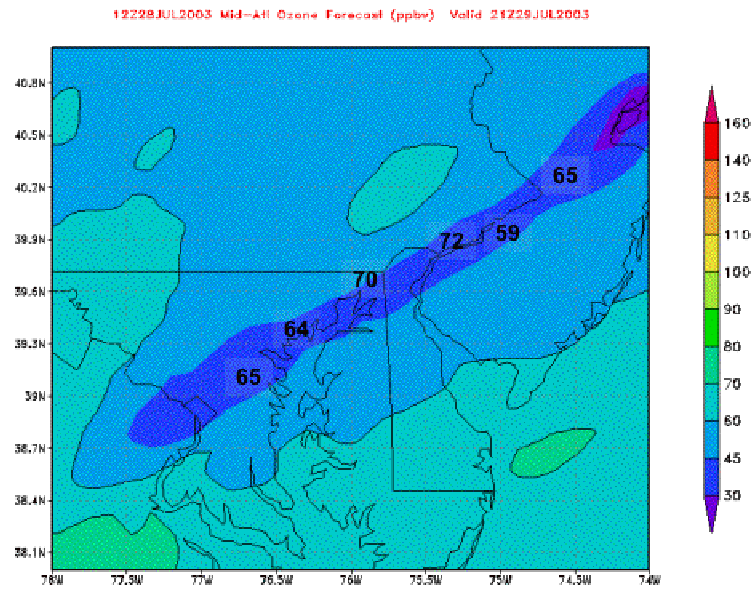

Figure 7. NAQF (experimental) forecast for July 28, 2003. Observed O_3 concentrations are shown in numeric values (ppbv). Contours for forecast O_3 are on right sidebar (ppbv).

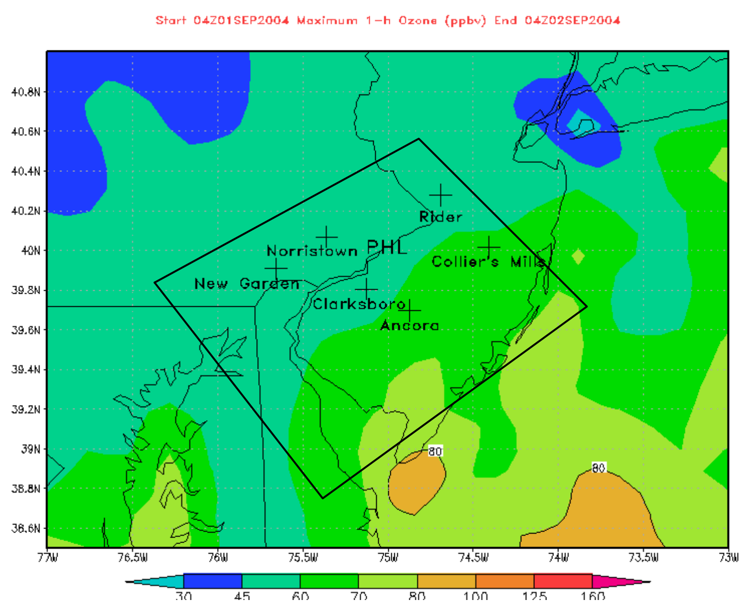

Figure 8. Philadelphia metropolitan air quality forecast area (enclosed in black box) and set of six monitors used to evaluate model performance.

Preliminary Results for Key PHL Monitors
Peak 1-hour O_3 (June 14-August 22)

	Bias (ppbv)	Normalized Bias	r^2
Collier's Mills	-0.6	-0.8%	0.55
Ancora, NJ	-4.5	-7.0%	0.23
New Garden, PA	-3.6	-5.7%	0.56
Norristown, PA	+0.3	+0.5%	0.51
Clarksboro, NJ	-2.2	-1.5%	0.65
Rider, NJ	+2.2	+3.6%	0.62
Maximum Domain O_3:	+1.1	+1.5%	0.67

$$[O_3]_{max} = 2.13 + 0.96[O_3]_{fcst}$$

Figure 9. Forecast performance of the NOAA developmental air quality forecast model for 6 selected monitors in the Philadelphia forecast area during the 2004 forecast season along with maximum domain O_3.

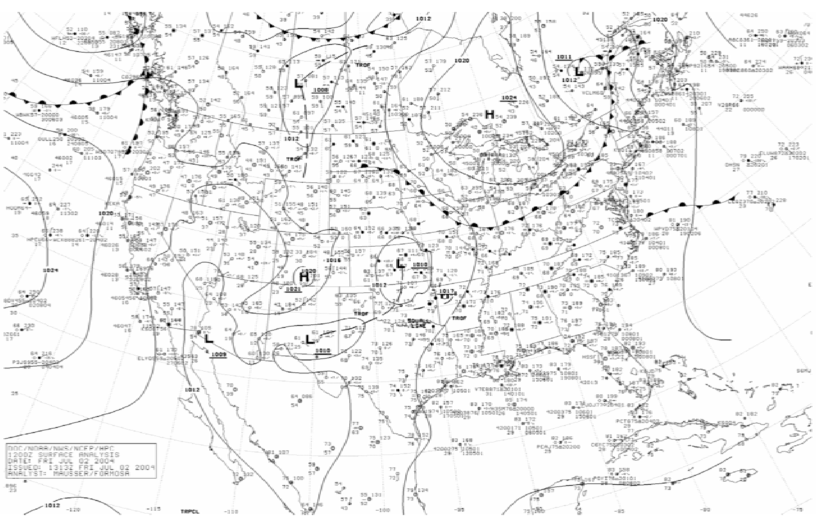

Figure 10. NCEP surface analysis for 1200 UTC July 2, 2004.

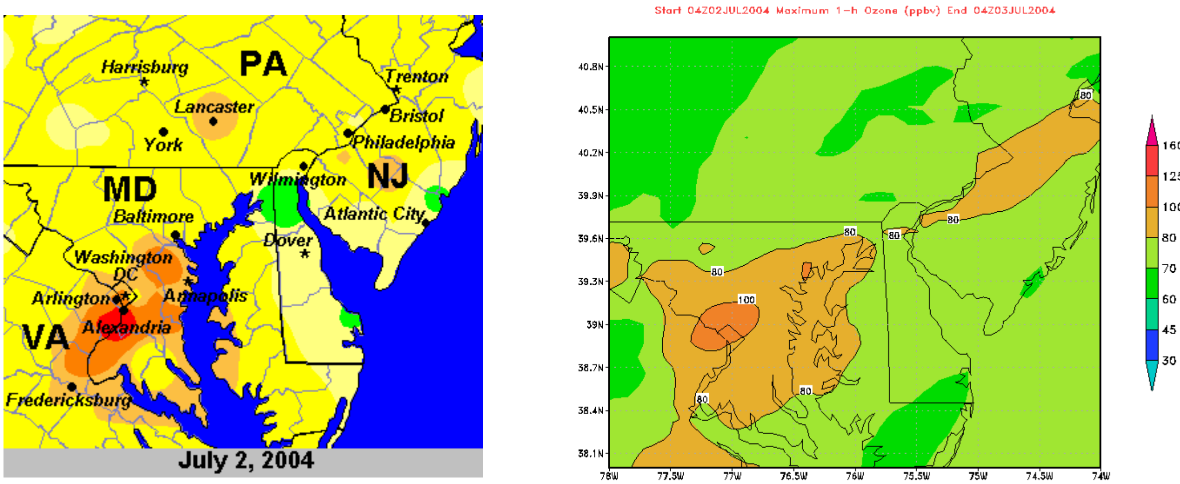

Figure 11. Peak 1-hour observed O_3 for July 2, 2004 (left panel) and NOAA developmental model forecast (right panel). Observed O_3 figure courtesy of EPA AirNow. Observed concentration contours are: yellow (90-99 ppbv), light orange (100-109 ppbv), dark orange (110-124 ppbv), red (≥ 125 ppbv). Forecast O_3 contours follow color bar at extreme right.

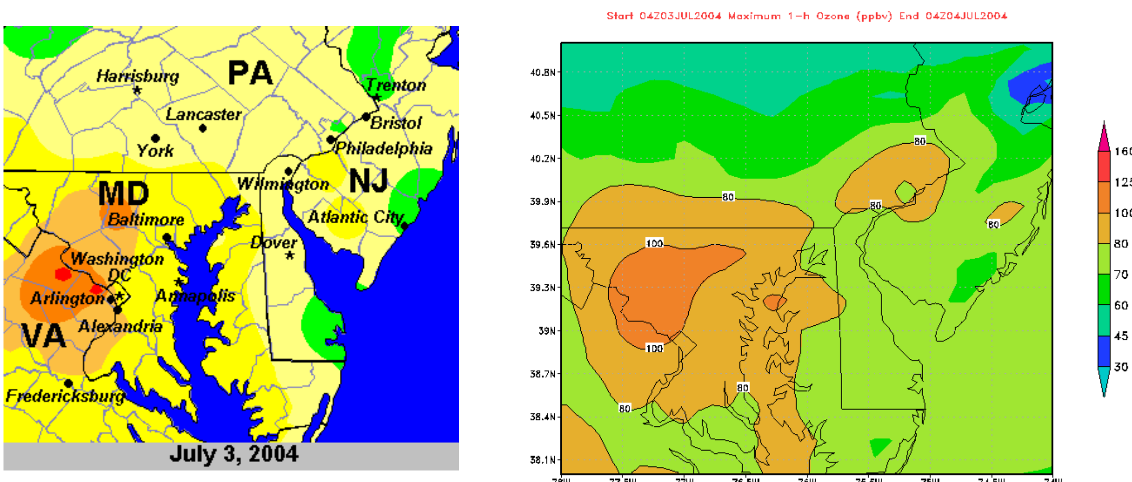

Figure 12. As in Figure 11 but for July 3, 2004.

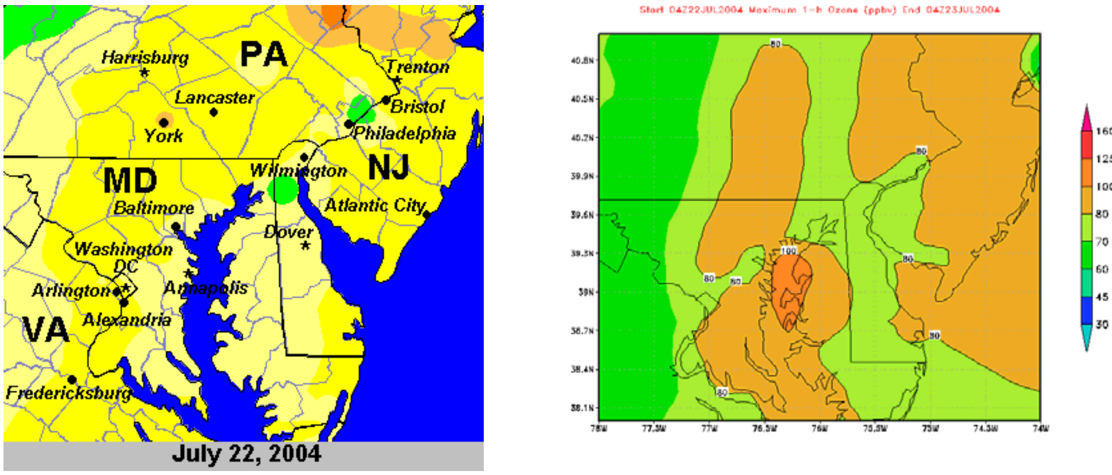

Figure 13. As in Figure 11 but for July 22, 2004.

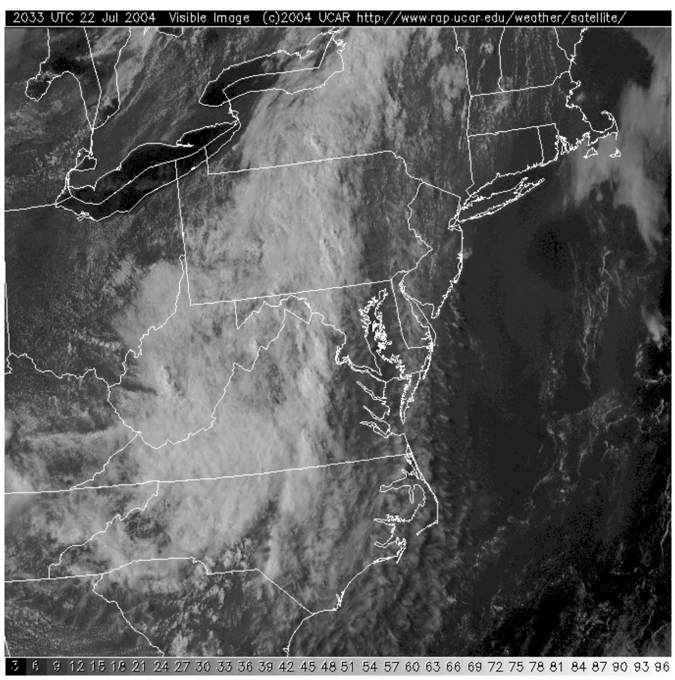

Figure 14. GOES visible image for 2033 UTC on July 22, 2004.