J2.13

EVALUATION OF THE NATIONAL AIR QUALITY FORECAST SYSTEM (NAQFS): 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

Kenneth Carey
Center for Science and Technology, Mitretek Systems, Falls Church, Virginia

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 initial 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 is presented here.

* *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 was 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 9-10, 2003 to discuss results and recommend further actions.

3. THE AIR QUALITY FORECAST MODEL

More complete details of the forecast model will be provided in other papers. A basic description of the model system is provided here. The modeling system linked the NCEP Eta-12 meteorological model with the EPA Community Multi-Scale Air Quality (CMAQ) model. Some details of the models are provided in Table 2. The model domain is shown in Figure 1. Significant modifications to the air quality model were needed to adapt it to operational use, driven by the Eta-12. CMAQ was designed and intended for use as an assessment tool for developing air pollution control strategies. 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 3.10) 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.

In the course of the development of the coupled model system, significant changes were made to the Eta-12 model. These changes were effective July 8, 2003 and, in some cases, impacted the chemistry model. (http://wwwt.emc.ncep.noaa.gov/ mmb/tpb.spring03/tpd.htm). The results discussed here include primarily model results for periods after these changes became effective. The changes of interest to air quality forecasts include modifications to cloud microphysics and radiation modules that affect fractional coverage of convective clouds including shallow convective clouds. The recent changes tend to increase the occurrence of partly cloudy skies and the frequency of overcast skies. An inadvertent error in the output land-use categories was corrected on September 8, 2003. For a period of two months, this systematic error resulted in CMAQ reading ocean surface for a large portion of the domain. Analyses of systematic impact indicate that this error contributed an average of 5-10 ppbv excess in predicted O_3 concentrations (Kenneth Schere, personal communication).

4. FOCUS GROUP ACTIVITIES: VISUALIZATION AND EVALUATION

The forecast production cycle consisted of two daily runs. The primary forecast, initialized at 1200 UTC, was run for 48 model hours. A secondary forecast, initialized at 0600 UTC, was run for 30 model hours and provided initial conditions for the 1200 UTC run. The evaluation undertaken here focuses on the 1200 UTC model run 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 and, later in the season, on the NWS Telecommunications Operations Center (TOC) 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. Additional images and test results were provided at a Pennsylvania State University (PSU) website, again for the focus group evaluations. 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 Figures 2-4. 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. Later in the forecast season, 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. Several examples of the comparison plots are shown later in this paper.

Using these forecast images, focus group members were able to provide daily feedback on forecast model performance. The feedback was typically provided on a standardized Excel worksheet and sent to focus group coordinators at NWS. 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

The summer of 2003 featured extremely low O_3 compared to recent climatology. This was particularly true during July and August – the "heart" of the O_3 season in the northeastern US. For the period July 6-September 13, the Mid-Atlantic Regional Air Managers Association (MARAMA), which collects O_3 data for the mid-Atlantic region, identified 45 total cases of O_3 monitors in excess of the 85 ppbv 8-hour O_3 standard. Only three cases 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, only 2 days exceeded the 1-h O_3 health standard (125 ppbv) throughout the entire summer and only 13 exceeded the 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.

The historically low O_3 concentrations in 2003 were driven by unusual summer weather patterns. In July, a persistent large-scale circulation anomaly placed a deep upper level trough over the eastern US with a corresponding strong ridge over the western US (Figure 5). With an upper level low over the region, rain and frontal passages were frequent. The combination of precipitation, cloud cover and intrusions of cool Canadian air was not conducive to O_3 formation. While temperatures returned to near normal levels in August, the air mass along the eastern seaboard was often maritime tropical (mT) in nature. This led to low background O_3 levels, as the mT air mass is clean with respect to O_3 precursors, and frequent rain showers. Many stations across the mid-Atlantic (e.g., PHL, ILG, BWI) recorded measurable precipitation on 50% or more of July and August days.

6. FOCUS GROUP WORKSHOP AND EVALUATION SUMMARY

The AQF Focus Group convened in Silver Spring, MD on September 9-10, 2003 to hear a series of presentation on model performance, discuss their findings, exchange ideas, and provide analysis and recommendations to NOAA and EPA. The systematic error in the input land-use data had not been corrected prior to workshop evaluations. A sample of forecasts, during the period August 12-19, was re-run with corrected land-use data. The correction reduces predicted O_3 concentrations by 5-10 ppbv on average during this period of relatively elevated O_3 concentrations. This correction is not included in workshop discussions and evaluations.

A statistical analysis of model performance was provided by the Meteorological Development Lab (Shaffer, W., et al.; Office of Science and Technology - NWS). A total of 640 US monitors was surveyed (Figure 6). The mean absolute error, determined hourly, of the model ranged from 16-27 ppbv (Figure 7) with a peak error in the pre-dawn hours and a minimum error in the late afternoon/early evening. A high bias (over-prediction), in the range of 15-25 ppbv, was noted. Results from a threshold contingency table showed a probability of detection (POD) of instances of 8-hour average O_3 in excess of the 85 ppbv of 0.52 (0.63 for August only) but at the cost of a very high (~0.9) false alarm rate. As noted earlier, the land-use correction will likely reduce the bias by a significant amount.

An evaluation carried out by the Atmospheric Modeling Division of the NOAA Air Resources Laboratory (Eder, B., et al.) showed similar results. The overall correlation for maximum 1-hour forecast and observed O_3 was 0.65 ($[O_3]_{fcst} = 34.9 + 0.65[O_3]_{obs}$). Spatial analysis was included in this presentation showing that the poorest overall correlation, and highest over-prediction, occurred along the spine of the Appalachians west of the I-95 Corridor and in the Piedmont of western NC, western SC and northern GA (Figure 8).

Problems with forecast over-prediction were also noted by local forecasters. An example of over-prediction in the NC Piedmont is given in Figures 9-10. The over-prediction in this case occurred in the context of warm, unstable conditions (Figure 11). Over-prediction was also found on occasion in the NO_x-rich Ohio River Valley region (Figures 12-13). In this case, conditions were favorable for O_3 production with a large upper level ridge over the region. A case of over-prediction along the spine of the Appalachians west of I-95 of is shown in Figure 14. In this case, a frontal boundary stalled along the I-95 Corridor. Stalled boundaries can often be the site of rapid and localized rises in O_3. In this case, however, the air mass east of the front was of tropical origin and unstable resulting in considerable cloud cover and locally heavy rain (Figure 15)

Additional, more localized, problems were also identified by focus group members. In the vicinity of bay-land interfaces, there was a tendency for localized "bullseyes" of high O_3. An example near Cape May, NJ is shown in Figure 16. A similar feature was also observed over the southern Chesapeake Bay and affected forecast performance in the Norfolk region. Very low O_3 , presumably due to NO_x titration from motor vehicle sources, was observed along the I-95 Corridor from Washington

DC to Philadelphia (Figure 17). The effect was typically stronger near PHL than the other cities.

The forecasters noted that the spatial location of the maximum O_3 plume was often well forecast by the model. This suggests that the meteorological model was successful in providing low level, and fine scale, wind forecasts. This led to a discussion of the merits of a dynamic MOS model to accompany the forecast guidance that might adjust the model bias while retaining spatial skill.

Several hypotheses for the modeled over-prediction and plans for sensitivity runs to determine the source of the problem were presented (J. Pleim, Atmospheric Science Modeling Division, NOAA ARL). Three major areas of concern were noted. First, planetary boundary layer mixing schemes may contribute to over-prediction. Vertical concentration profiles of ground-based emissions show large gradients in the convective boundary layer with highest concentrations in the lowest layer. This could lead to over-prediction in the lowest layer in some cases (e.g., areas with high emissions of isoprene) as well as titration and under-prediction in others (e.g., areas with large NO_x sources). Second, photolysis related issues include problems with the CMAQ cloud cover algorithm as well as short comings in the photolysis model itself. Finally, the emissions model may contain errors that led to systematic model biases. As over-prediction was most enhanced in the southeastern US, it is possible that over-prediction of isoprene, whose concentrations are quite high in this region, is a problem. Mobile source NO_x may be an issue with respect to O_3 titration along the I-95 Corridor.

The initial set of sensitivity runs showed that Increasing vertical mixing and modifying the eddy diffusion coefficient did reduce NO_x titration effects but had no large effect overall on the over-prediction problem. Model predicted O_3 showed a low sensitivity to variations in isoprene emissions but a high sensitivity to NO_x reductions. A variety of future investigations of model performance are possible including the use of an eddy diffusion coefficient directly from the Eta model, adoption of a non-local PBL model, and a full evaluation of the photolysis module including modifying or replacing the cloud cover algorithm and updating the radiation code. The use of dynamic boundary conditions will be investigated as well as more analysis of sensitivity to emissions and further evaluation of the effect of the land surface error noted above.

7. CONCLUSIONS

The deployment by NOAA of an operational numerical air quality forecast model will mark 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, can provide valuable feedback to the model developers on the skill of the model and on the products that will be of most utility for forecast preparation. The air quality forecasters focus group workshop described here represented the initial step towards developing a effective feedback cycle from forecast users to developers.

ACKNOWLEDGEMENTS

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

REFERENCES

Draxler, R.R. and Rolph, G.D., 2003. HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY Website (http://www.arl.noaa.gov/ready/hysplit4.html). NOAA Air Resources Laboratory, Silver Spring, MD.

Rolph, G.D., 2003. Real-time Environmental Applications and Display sYstem (READY) Website (http://www.arl.noaa.gov/ready/hysplit4.html). NOAA Air Resources Laboratory, Silver Spring, MD.

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

Member	Affiliation
Joanne M. Alexandrovich	Vanderburgh County Health Department
Robert Brawner	State of Tennessee, Air Pollution Control Division
George M. Bridgers	NC Division of Air Quality
Richard Burkhart	U.S. EPA, Region 1
Ken Carey	Mitretek Systems
Jennifer Carfagno	The Weather Channel
Joe Cassmassi	South Coast Air Quality Management District
Neal Conatser	Michigan Department of Environmental Quality
David Conroy	U.S. EPA, Region 1
Paul Dallavalle	National Weather Service
Paula Davidson	National Weather Service
Phil Dickerson	US EPA
Tim Dye	Sonoma Technology, Inc.
Pamela Frazier	State of Tennessee, Division of Air Pollution Control
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
Robert Brawner	State of Tennessee, Air Pollution Control Division
Anne McWilliams	U.S. EPA Region 1
Sean Nolan	Pennsylvania Department of Environmental Protection
William F. Ryan	The Pennsylvania State University
Dan Salkovitz	Virginia Department of Environmental Quality
Will Shaffer	NWS Office of Science and Technology
Paul Stokols	NWS Office of Climate, Weather and Water Services
Richard A. Wayland	US EPA
Dan White	Texas Commission on Environmental Quality
John E. White	US EPA
Lian Xie	North Carolina State University

Table 2. Summary of air quality modeling systems with links providing further details.

Meteorological Model	
Eta-12	60 vertical layers Stepped mountain vertical coordinate Arakawa-E staggered grid
Web sites of interest:	http://www.emc.ncep.noaa.gov/mmb/mesoscale.html http://wwwt.emc.ncep.noaa.gov/mmb/tpb.spring/tpb.htm http://www.meted.ucar.edu/nwp/pcu2
Chemistry Model	
CMAQ	22 vertical layers SMOKE emissions model (MOBILE 6 and BEIS 3.10 sub-modules) 1999 EPA emissions inventory
Web sites of interest:	http://www.epa.gov/asmdnerl/models3

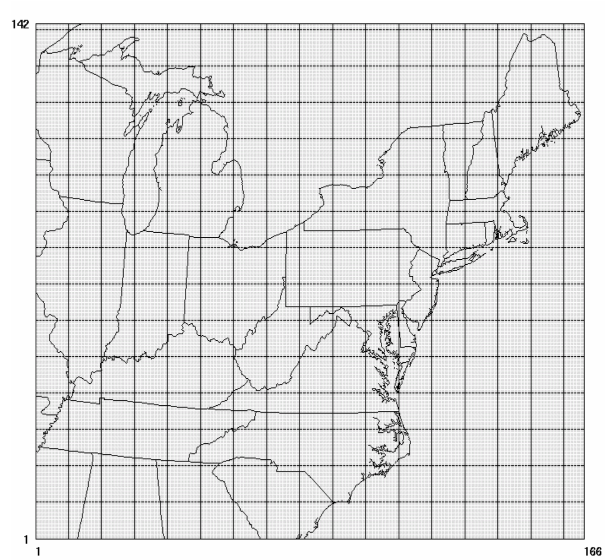

Figure 1. Domain of the NOAA Air Quality Forecast model.

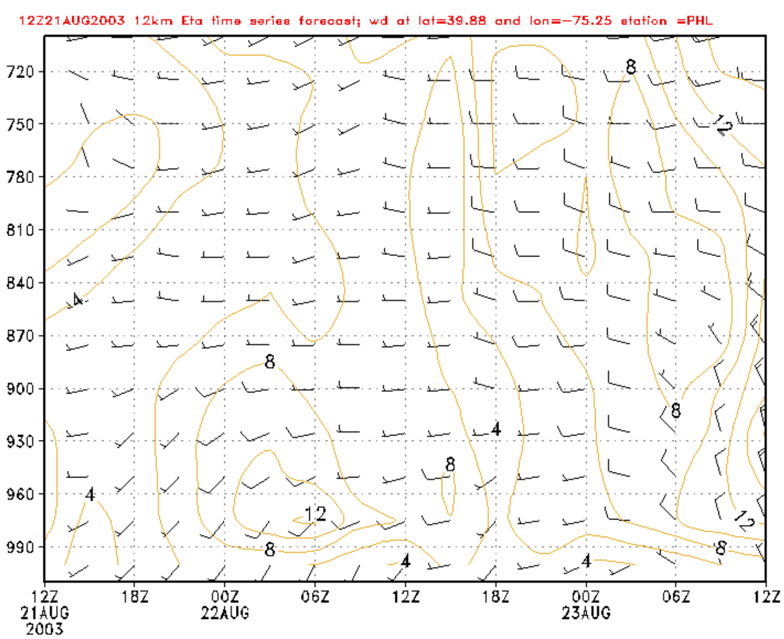

Figure 2. Vertical time series forecast of wind (in ms^{-1}, full barbs = 10 ms^{-1}) from the NCEP Eta-12 model initialized at 1200 UTC August 21, 2003.

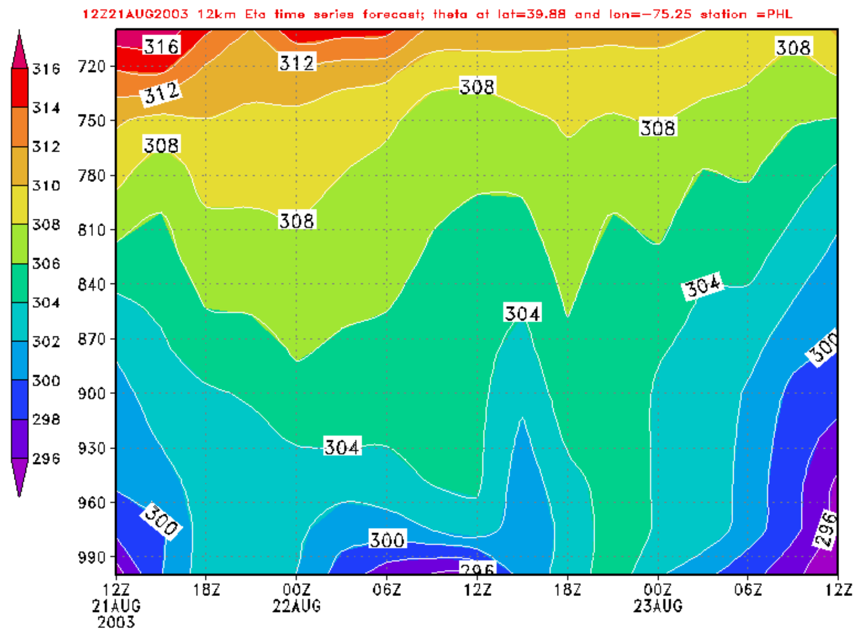

Figure 3. Vertical time series as in Figure 2 but for potential temperature (in degrees K).

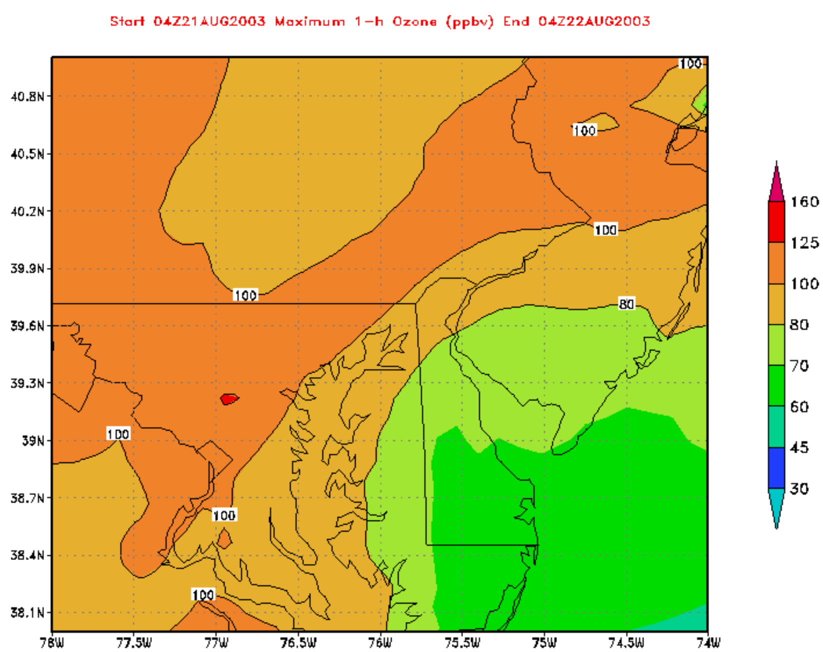

Figure 4. Maximum 1-hour O_3 concentrations from the AQF model for the DC-PHL sub-region valid August 21, 2003 from forecast initialized at 1200 UTC August 20, 2003..

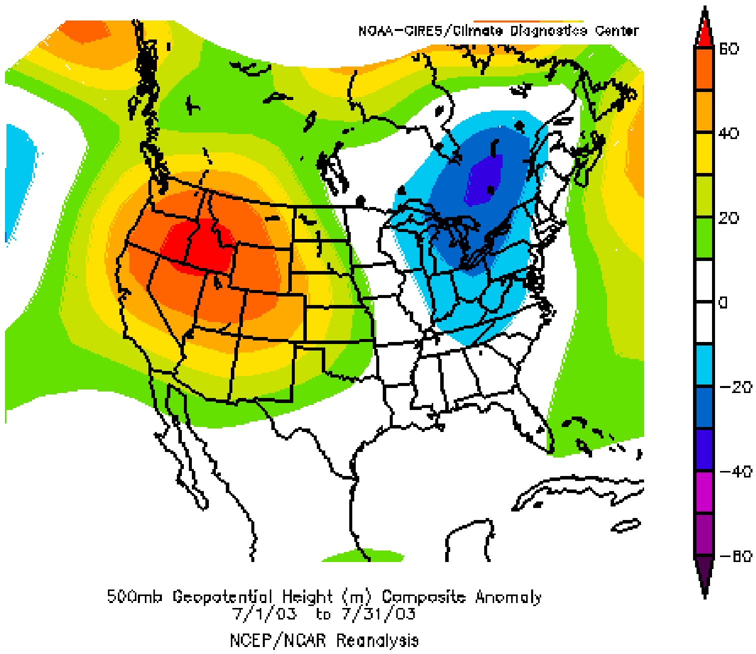

Figure 5. 500 mb height anomaly for July, 2003 compared to recent climatology using NCEP/NCAR re-analysis data. Figure courtesy of NOAA-CIRES, Climate Diagnostics Center.

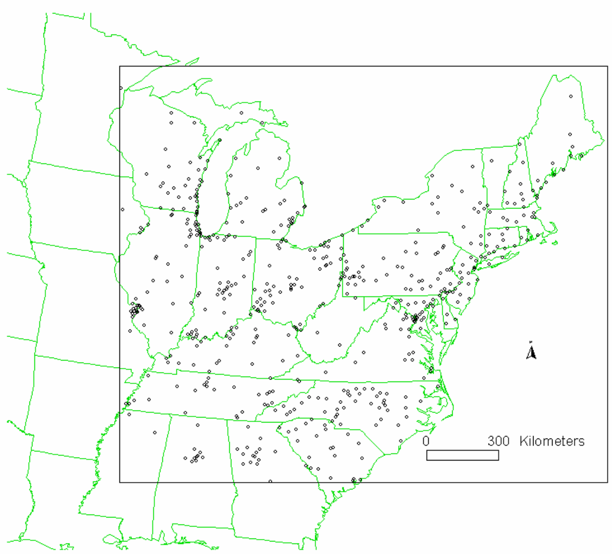

Figure 6. Ozone monitors in the northeastern United States used for model evaluation. Figure courtesy of the NWS Meteorological Development Laboratory (W. Shaffer, M. Schenk, J. Gorline and V. Dagostaro).

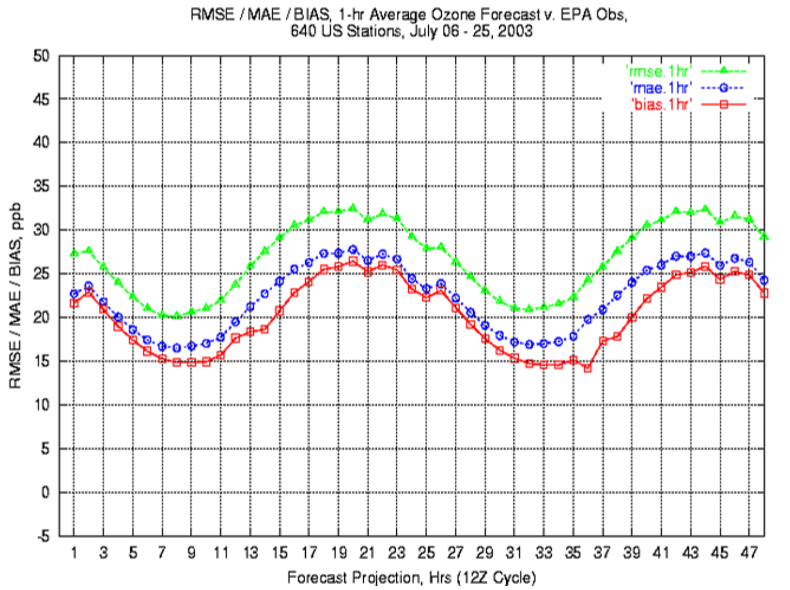

Figure 7. Performance of the AQF model by forecast hour. Root mean square error (rms) is given by the green line, mean absolute error by the blue line and bias by the red line. Figure courtesy of the NWS Meteorological Development Laboratory (W. Shaffer, M. Schenk, J. Gorline and V. Dagostaro).

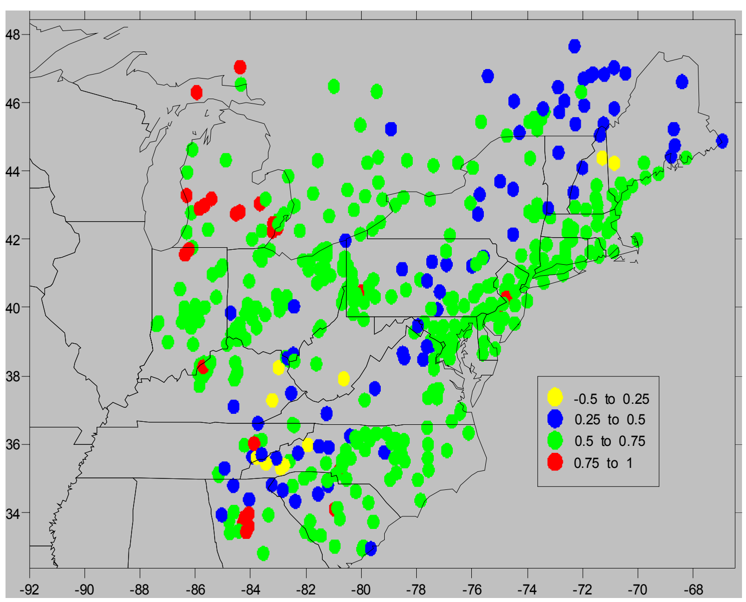

Figure 8. Correlation of hourly O_3 forecasts and observations. Figure courtesy of the Atmospheric Modeling Division, Air Resources Laboratory, NOAA (B. Eder, D. Kang, K. Schere, and J. Pleim).

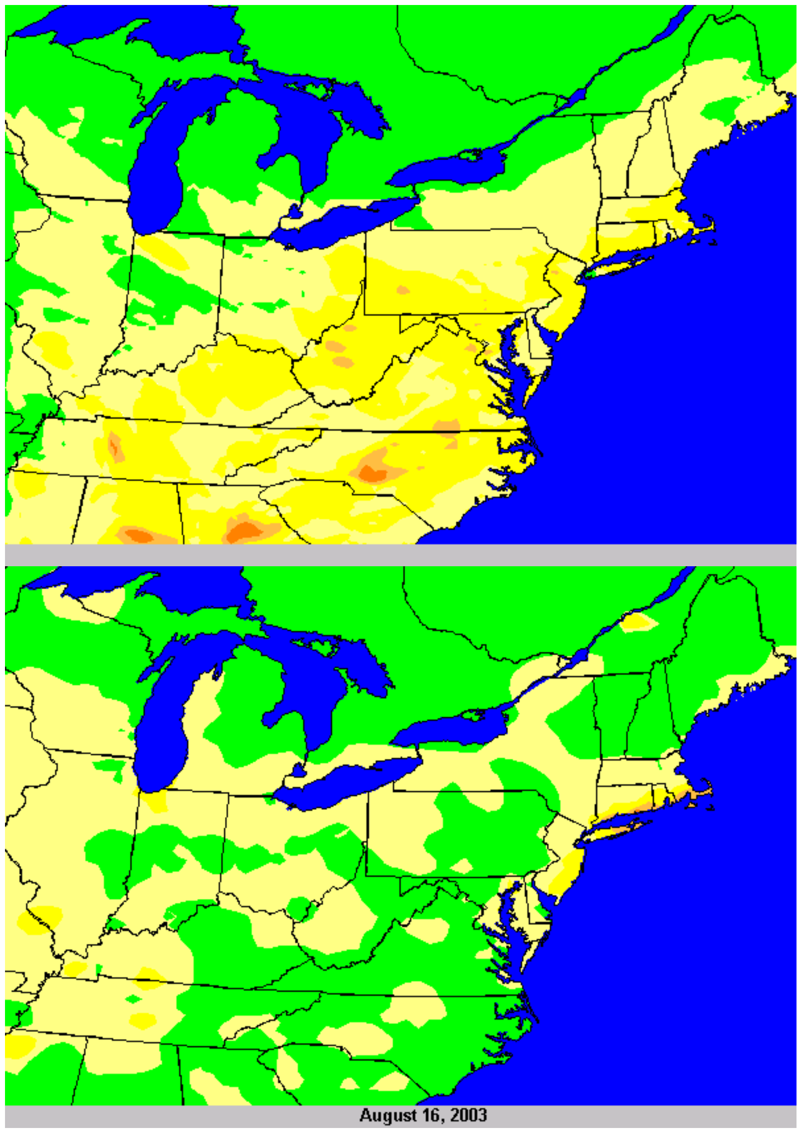

Figure 9. Forecast (top panel) and observed (bottom panel) peak 1-hour O_3 for August 16, 2003. Figure courtesy of EPA AIRNow and Sonoma Technology, Inc. Color contours are as follows: Green (0-60 ppbv), light yellow (61-79 ppbv), yellow (80-99 ppbv), light orange (100-110 ppbv), orange (111-124 ppbv) and red (> 124 ppbv).

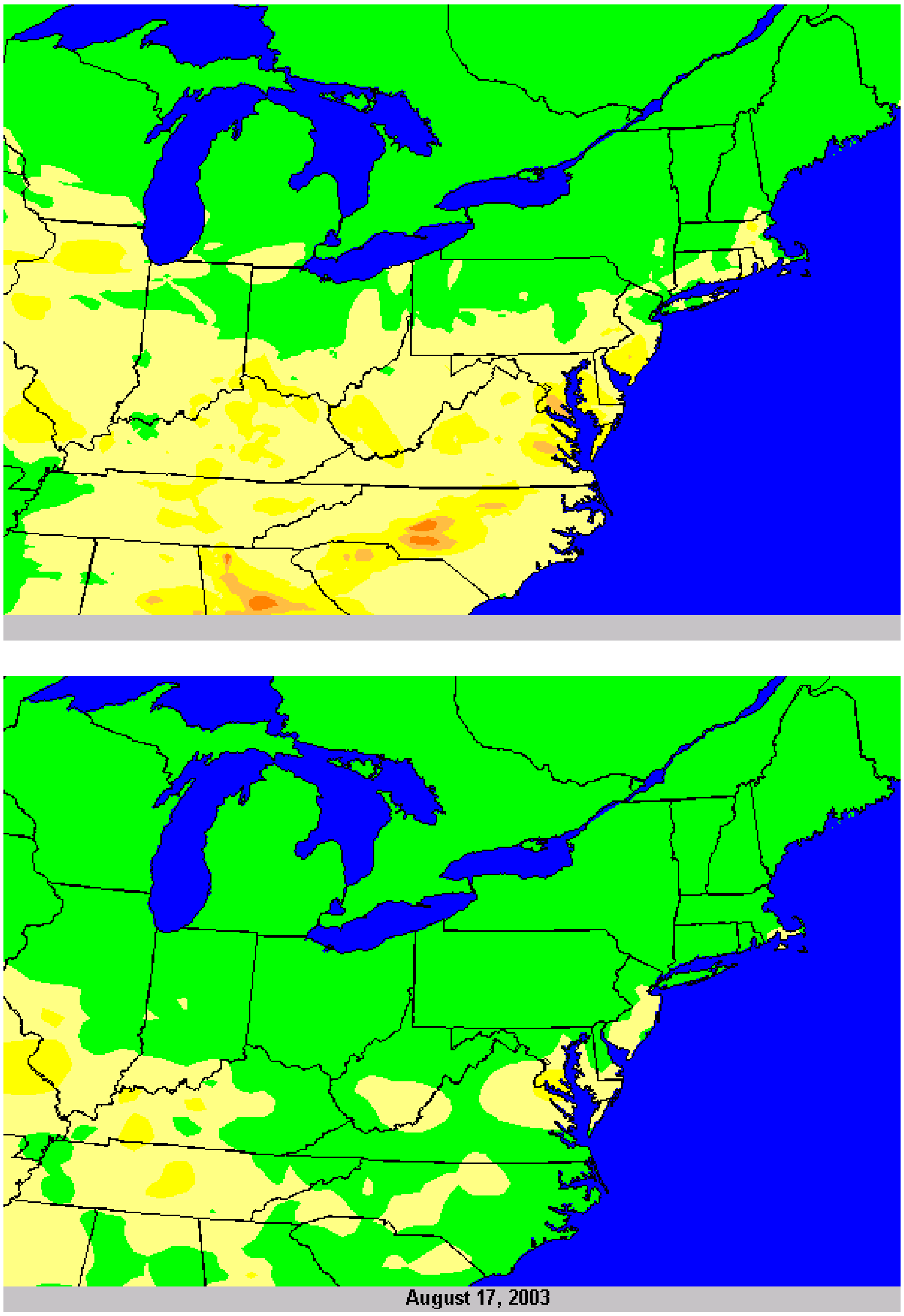

Figure 10. As in Figure 9 but for August 17, 2003.

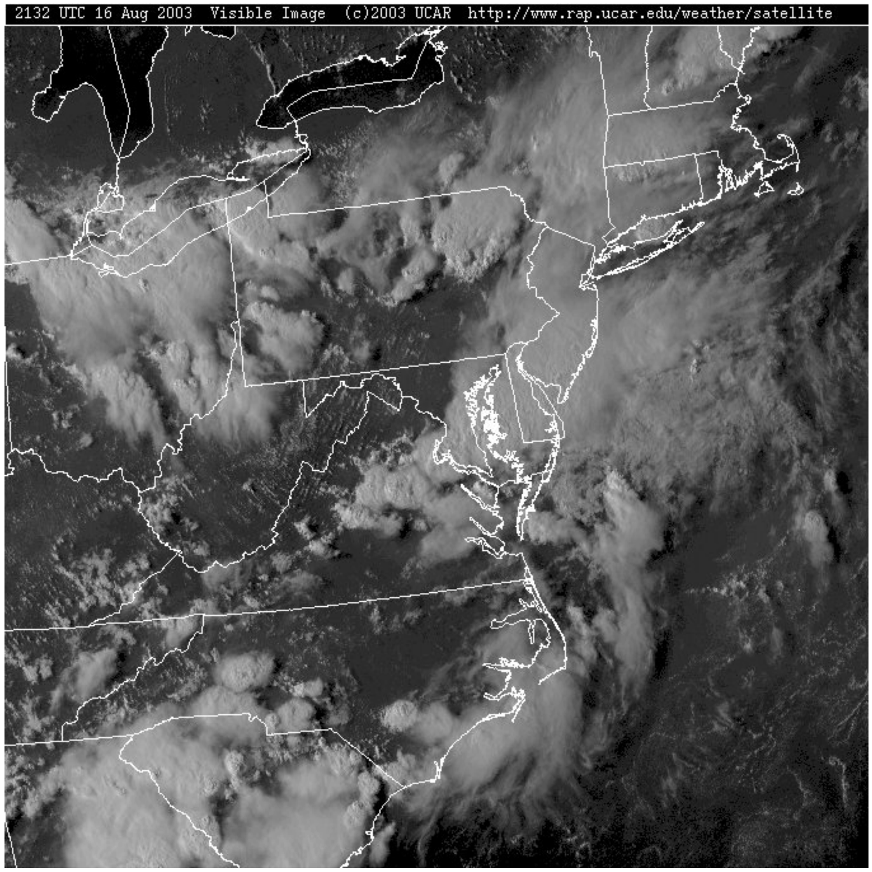

Figure 11. High resolution GOES visible image for 2132 UTC on August 16, 2003. Ozone observations and forecasts are given for this day in Figure 9.

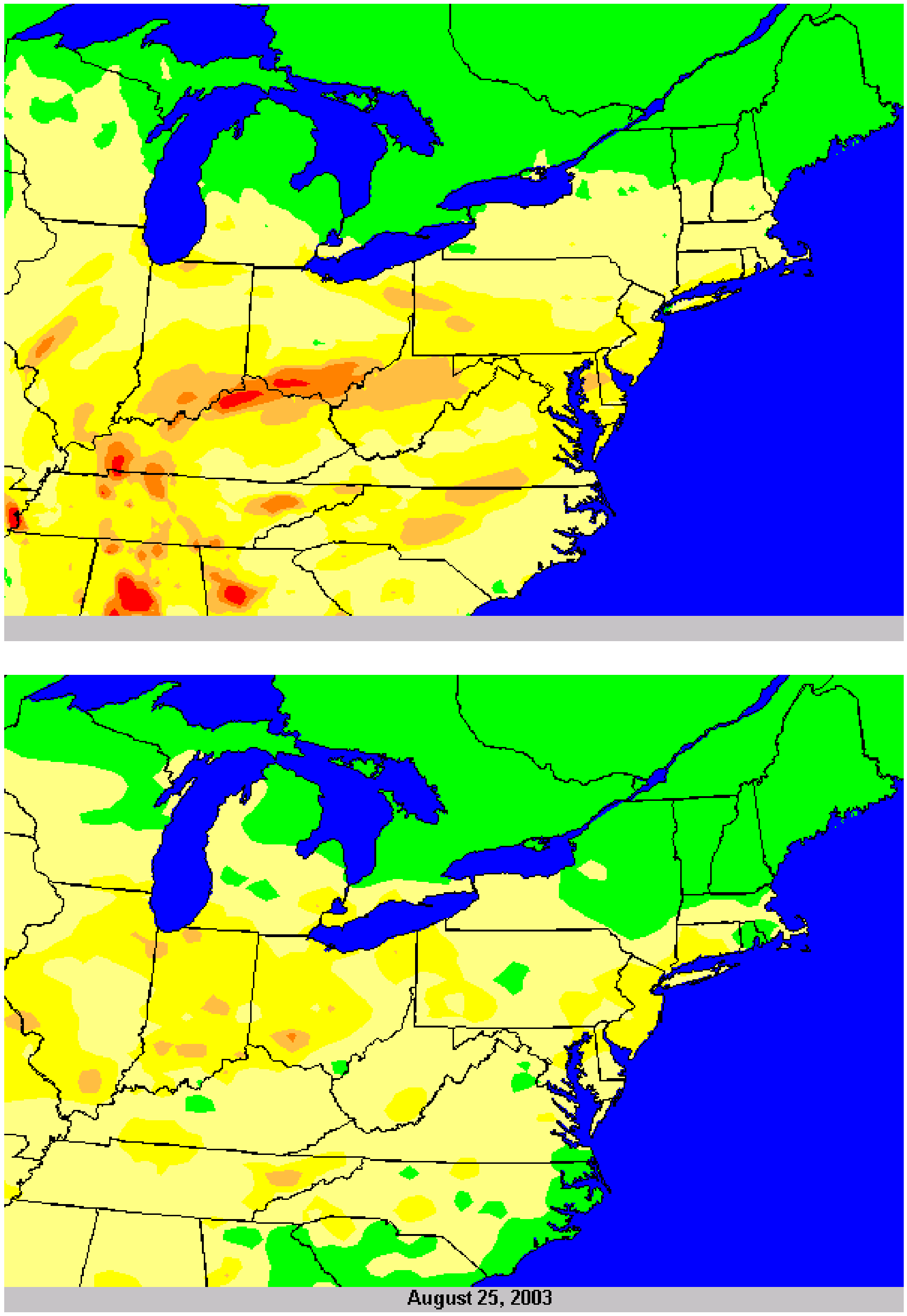

Figure 12. As in Figure 9 but for August 25, 2003.

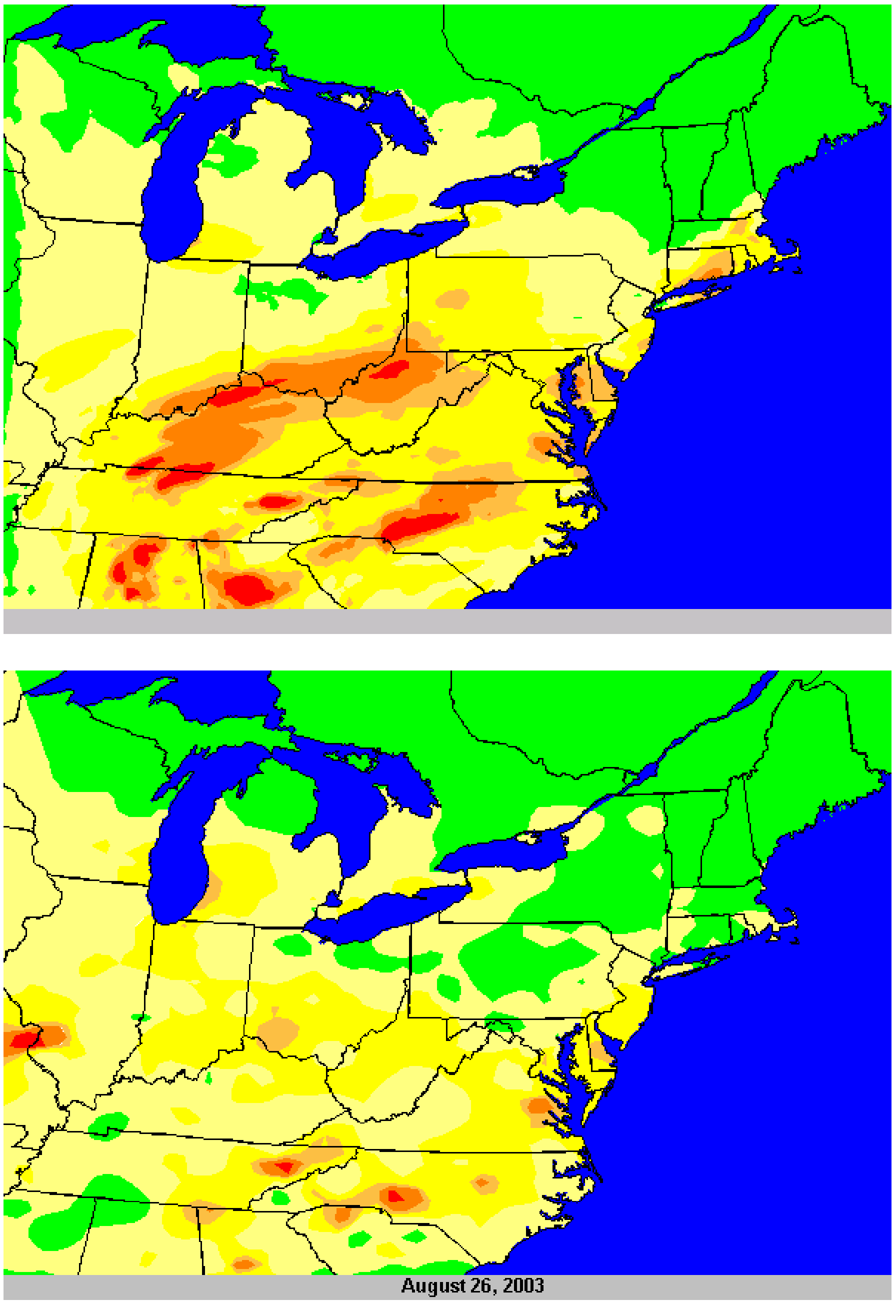

Figure 13. As in Figure 9 but for August 26, 2003.

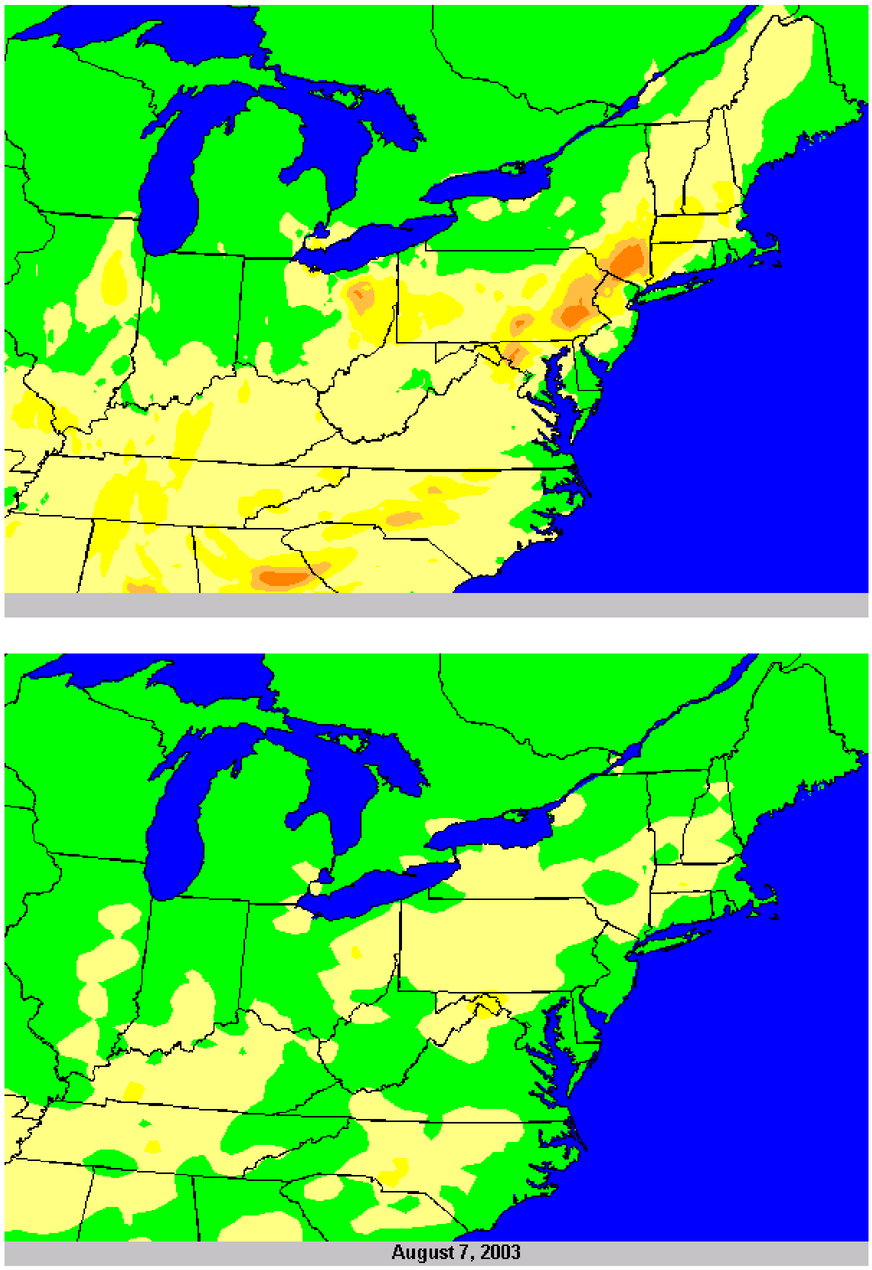

Figure 14. As in Figure 9 but for August 7, 2003.

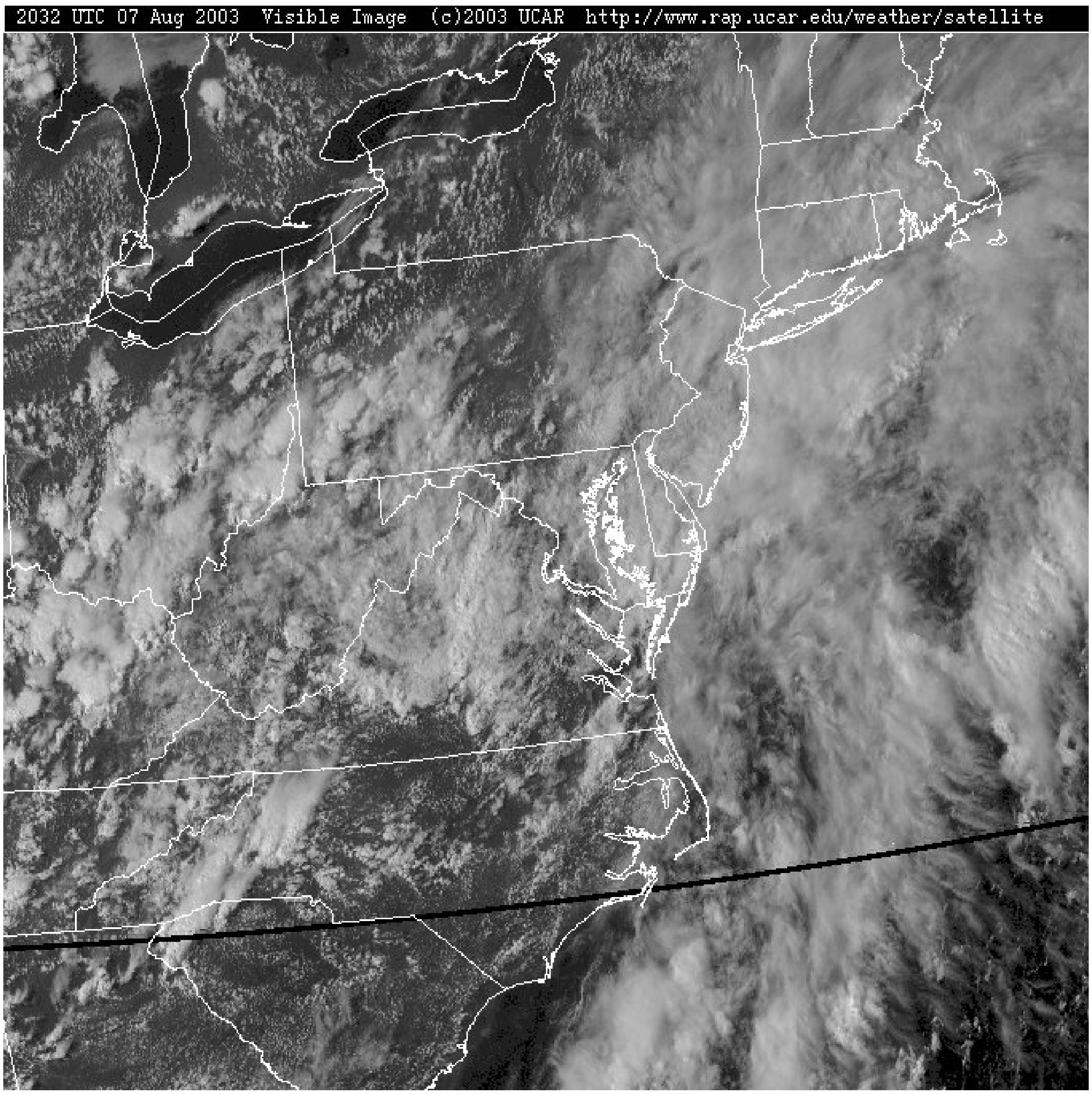

Figure 15. High resolution GOES visible image for 2132 UTC on August 7, 2003. Ozone concentrations and forecast for this day are given in Figure 15.

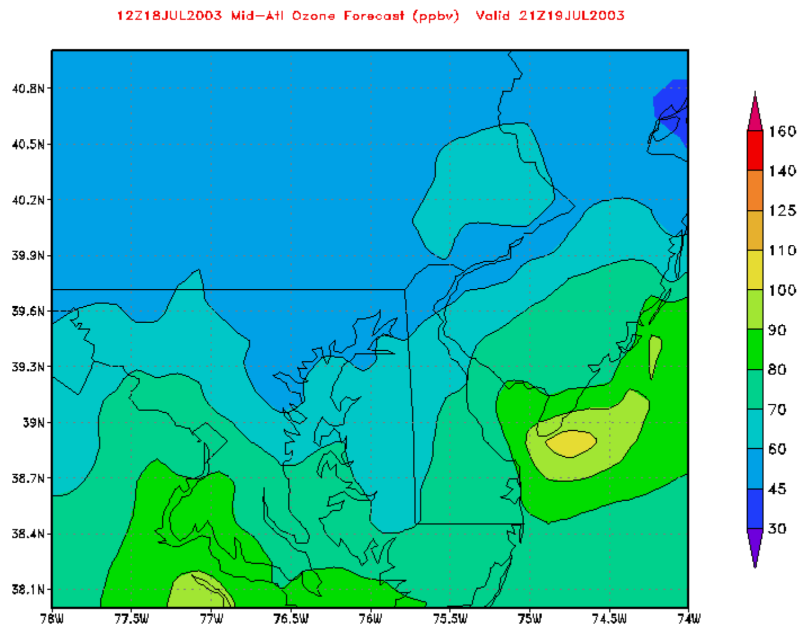

Figure 16. Forecast 1-hour O_3 concentrations for 2100 UTC on July 19, 2003. The AQF model was initialized at 1200 UTC on July 18, 2003.

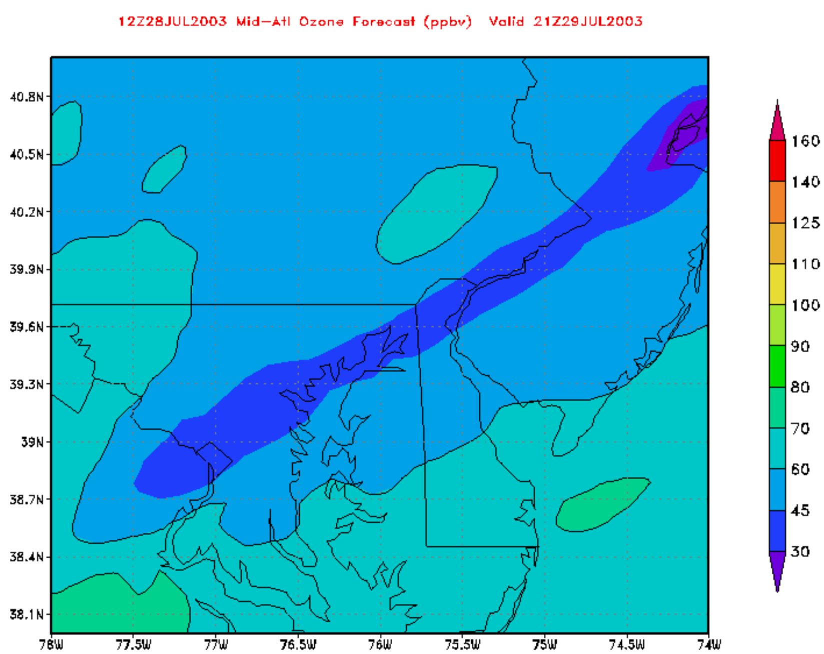

Figure 17. As in Figure 17 but for 2100 UTC on July 29, 2003.