United States
Environmental Protection
Agency

Atmospheric Research and Exposure
Assessment Laboratory
Research Triangle Park NC 27711

Research and Development

EPA/600/S3-89/057 July 1989

EPA

Project Summary

EPA Regional Oxidant Model (ROM2.0): Evaluation on 1980 NEROS Data Bases

Kenneth L. Schere and Richard A. Wayland

The second generation USEPA Regional Oxidant Model (ROM2.0) has been evaluated for the northeastern United States using the 1980 NEROS data bases. The theoretical basis of the model and its structure and organization are described. The data bases available from the summer 1980 period include routine air quality and meteorological monitoring data in addition to data from several extensive field measurement projects conducted during the summer of 1980 in the northeastern U.S. on regional and urban scales. Also, a complete emissions inventory, composed of anthropogenic and biogenic components, was assembled for the 1980 base year for use in air quality modeling. The ROM2.0 evaluation was conducted using quasi-deterministic and diagnostic techniques. Strict temporal and spatial pairing between observations and predictions was not used in the analysis. Model simulation was conducted during the period of July 12 to August 31, 1980. Model performance over the simulation period showed an overall 2% overprediction of the daily surface maximum O_3 concentrations. ROM-predicted concentrations, however, had a narrower range for ambient O_3 with underestimation of highest values and overestimation of lowest values. The spatial extent and concentrations of urban O_3 plumes were generally simulated well, although a bias in the transport direction along the East Coast caused frequent misalignment of the plumes. Model performance analyses using aircraft data showed the model to underpredict the regional O_3 tropospheric burden under episodic conditions, although individual plumes were modeled well.

This Project Summary was developed by EPA's Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

The initial development of a regional (~1000 km) air quality simulation model began in the late 1970's after the realization that photochemical smog often extended beyond individual urban areas to entire sections of the U.S. Interstate transport of O_3 and its precursors was observed during field programs of the 1970's, especially in the Northeastern U.S. Long-range transport of O_3 and precursors and multi-day chemical effects could not be properly treated by existing urban-scale photochemical models. The need became apparent for an appropriate simulation model to test the effectiveness of particular emissions control strategies on O_3 concentrations in urban airsheds as well as region-wide.

The first generation EPA Regional Oxidant Model (ROM1.0) became operational in 1984. It was a test bed for the future production version of the model, the second generation ROM2.0. The

earlier model contained a very condensed chemical kinetic mechanism, could not treat natural hydrocarbons, had limited treatment of vertical mass flux induced by clouds, had constant layer depths, and contained very limited terrain effects. Nevertheless, ROM1.0 was used for extensive testing of various emission control scenarios in the Northeast U.S.

ROM2.0 became operational in 1987. This second generation version of the model was to become the production version. It included a more sophisticated, contemporary chemical kinetic mechanism capable of treating both anthropogenic and natural precursor species. The ROM2.0 system also corrected most of the deficiencies and simplifications of the first generation system, such as cloud-induced mass flux, variable layer depths, and terrain effects. Once operational, ROM2.0 was used for extensive testing of emissions control scenarios within the Northeast U.S. for EPA's Office of Air Quality Planning and Standards (OAQPS), as well as the Vice President's Commission on Clean Coal Technology, and the Congressional Office of Technology Assessment.

A large field project in the Northeast U.S. was planned and conducted concurrently with the ROM model development effort. The purposes of the field program were to gather data to better understand the important processes responsible for photochemical smog on the regional scale so they could be properly simulated by the model and to provide a data base for testing and evaluation of the model. The field program was carried out over two summer periods during 1979 and 1980. Measurements were taken from Ohio and Michigan to the East Coast, both in regional and in local projects. Extensive aircraft measurements supplemented the ground-based network. The data base from 1979 was used primarily in the model development, testing, and evaluation phases of ROM1.0. The 1980 data base was used almost exclusively for model evaluation of ROM2.0.

This project presents the results of the evaluation of ROM2.0 on the 1980 Northeast U.S. data base. Emphasis is placed on the comparison between predicted and observed O_3 concentrations, although NO_x and hydrocarbons are evaluated also, to the extent that the data allow. Ground-based and aircraft data are used in the analysis. The purpose of this evaluation is essentially a *confidence building* exercise in the ROM2.0 for its use as an air quality planning tool.

Model Application

The ROM has been designed to simulate most of the important chemical and physical processes responsible for the production of photochemically produced O_3 on scales of 1000 km, or several days of transport time. These processes include horizontal transport, atmospheric chemistry, nighttime wind shear and turbulence episodes associated with the nocturnal jet, cumulus cloud effects on vertical mass transport and photochemical reaction rates, mesoscale vertical motions induced by terrain and the large scale flow, terrain effects on advection, diffusion, and deposition, subgrid scale chemistry processes, emissions of natural and anthropogenic precursors, and deposition. They are mathematically simulated in the 3-D Eulerian model with 3-1/2 vertical layers including the boundary layer and the capping inversion or cloud layer. Horizontal resolution is 1/4° of longitude by 1/6° of latitude, or about 18.5 km.

The particular application of ROM2.0 used in this evaluation exercise was conducted on a historical data base, with model simulation beginning at 1200 h, local standard time (LST) on July 12, 1980 and continuing through 1200 h, LST on August 31, 1980. The domain of application is that of the Northeast U.S., shown in Figure 1. The simulation was not reinitialized at any time after it began. It was continuous in time, performed in contiguous 3-day segments, with the simulation results from the final step of one segment used as restart conditions for the initial step of the next segment. The simulation period contained several significant O_3 episodes in the Northeastern U.S., with measured concentrations as high as 300 ppb.

The data bases used in this project include meteorological, air quality, and emissions components. The ROM requires all three types of data for simulating regional air quality. The performance evaluation of the model predictions primarily requires the air quality data base. The summer 1980 period was chosen for this effort because it coincided with several major field projects, conducted in the Northeast U.S., designed to study the regional and urban O_3 problems. These projects provided special data bases which supplemented the standard air quality and meteorological measurements archived in EPA's SAROAD (Storage and Retrieval of Aerometric Data) system and collected by NOAA's National Weather Service (NWS), respectively. The model ha been designed to run in an operation mode solely on these routinely collecte data bases. The special study data base are used for model evaluation ar research on model parameterizations.

The SAROAD data, supplemented b monitoring data in southern Ontari Canada obtained from Environme Canada and the Ontario Ministry of th Environment, provided hourly measur ments of O_3, NO_2, and NO_x at fixe monitoring sites within the model domai There were 214 sites where O_3 mea urements were made, 107 sites for NC and 65 sites for NO_x during the summ 1980 period. Most of the monitoring site are within or near urban areas.

Hourly surface meteorological me surements were available from ≈2(stations within the ROM domain in th NWS and Canadian meteorological ne works. The ROM preprocessors assir ilate raw meteorological data on atmo pheric pressure, temperature, moistur winds, and clouds from these location In addition to the surface measurement the North American upper air soundir network contains 24 stations within ar near the boundaries of the ROM doma where twice daily upper air sounding provided measurements of pressur temperature, moisture, and winds. Nir of these stations are located within th ROM domain boundaries, and the fr quency of soundings was increased four times per day at these statior during "intensive" NEROS field stuc periods.

During the 1980 summer season EPA Office of Research and Developme sponsored two major field studies in th Northeast U.S. The first was the Nort east Regional Oxidant Study (NEROS The NEROS field measurements conce trated on sampling strategies to clari and parameterize essential processe simulated within the ROM, to provic input data for the model, and to provic data with which to evaluate the mode The Persistent Elevated Pollutant Ep sode (PEPE) study was performed conjunction with the NEROS in 1980. I focus was on a regional perspective co cerning the broad regions of haz associated with large stagnant a masses.

The Northeast Corridor Region Modeling Program (NECRMP), spo sored by EPA's OAQPS, included urb field studies during the summer of 19 designed to collect the necessary a quality and meteorological data nece sary to apply the Urban Airshed Model

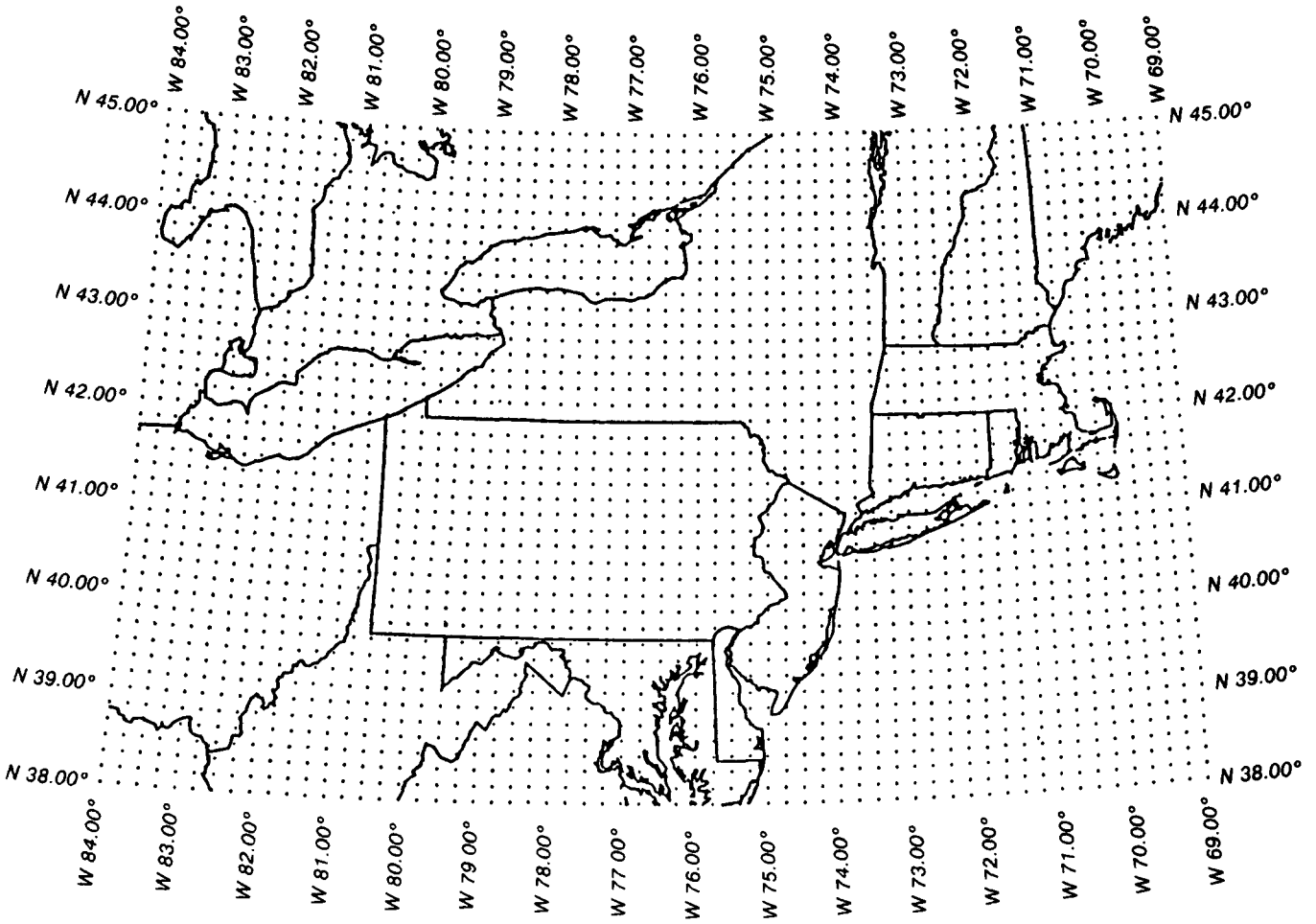

Figure 1. *Northeast U.S. ROM domain. Each dot represents a corner of a grid cell.*

Washington, DC, Baltimore, New York, and Boston. The data requirements for applying and evaluating this model led to the development of a monitoring program including air quality measurements by aircraft upwind of the urban area in the morning, and over and downwind of the urban area in the afternoon.

There are two major components to the emissions inventory data base needed by the ROM system: the anthropogenic emissions, and the biogenic emissions. Anthropogenic emissions of NO_x, CO, and various categories of non-methane hydrocarbons (NMHC) were obtained from the 1980 National Acid Precipitation Assessment Program (NAPAP) emissions inventory. The final inventory for use by ROM contains hourly emission rates on the ROM's 18.5-km grid resolution for NO, NO_2, CO, and the NMHC categories of ethylene, olefins, paraffins, formaldehyde, higher aldehydes, toluene, xylene, and non-reactives.

The biogenic portion of the emissions inventory, developed at EPA, consists of hourly, gridded values of natural hydrocarbon species. Three basic components are required to develop estimates of hourly, gridded biogenic NMHC emissions: (a) emission factors representative of vegetative species indigenous to the modeling region, (b) empirical relationships between emission factors and specific environmental parameters, and (c) quantitative estimates of vegetation density of the representative species in a designated area. The development of the biogenic inventory consisted of the compilation and assimilation of each of these factors.

To circumvent any model initialization problems, the ROM is initialized on a relatively clean day several days before any periods of high O_3 concentrations are found to exist in the domain. The full domain is assumed to contain spatially invariant values of clean tropospheric background concentrations of O_3, NO_x, and NMHC.[1] These species concentrations were then allowed to chemically equilibrate, using an independent chemical solution module. The set of initial values was spatially invariant in each horizontal model layer. To further isolate the model results from initial condition artifacts, the results for the first 24 hours of simulation are not used in the model evaluation analysis.

Boundary condition problems cannot be circumvented in the same manner as the initial condition problems. The goal, therefore, is to mitigate the effect of the

[1]Clean tropospheric background values used are O_3 = 35 ppb, NO_x = 2 ppb, NMHC = 15 ppb.

boundary conditions. This has been done to some extent by the specification of the size of the model domain. While the area of greatest interest in the simulations is in the Northeast Corridor from Washington, DC through Boston, the model domain extends in an upwind direction to the Ohio Valley and south to northern Virginia and West Virginia. In this way the model assimilates all of the significant upwind sources potentially making an impact on the Northeast Corridor and thereby reducing the influence of boundary conditions in that area. Obviously, the farther west and south in the domain one goes from the Corridor, the greater will be the boundary condition influence.

We have assumed the same tropospheric background conditions at the ROM boundaries as were described for the initial conditions, with one exception. We allow for the fact that O_3 may deviate from this background value at lateral boundaries. We set the boundary O_3 concentration from ambient monitoring data for each day of simulation. For a given day and time period the same O_3 value was used at all lateral boundaries.

Model Evaluation

Most methods of model evaluation used in previous studies have been largely deterministic. That is, the model concentration predictions for a specific location are compared to observations taken at that location on a given day. The ROM's developer (R. Lamb) maintains that there are inherent limitations on the predictability of air quality models, particularly on regional and larger spatial scales. He suggests that, even if the model is formulated perfectly, the data needed to drive the model are not sufficient to exactly determine the state of the atmosphere. This uncertainty in the atmospheric state gives rise to corresponding uncertainties in the concentration predictions from even a perfect model. The magnitude of the uncertainties in the concentration values is directly proportional to the extent to which concentrations at a receptor are affected by distant sources. Furthermore, the level of uncertainty increases with increasing distance from the nearest meteorological station and with increasing distance to significant sources. These ideas concerning the uncertainty in atmospheric state translate directly into uncertainty in the wind-driven transport component of the air quality model solution.

In the analyses used for this project we use a quasi-deterministic mode for the evaluation of ROM results when comparing predictions to surface-monitored observations of O_3, NO_2, and NO_x. In the quasi-deterministic mode we aggregate concentrations at groups of receptor locations and compare the aggregate frequency distribution of concentrations from the receptor group with that of the observed concentrations from the group. This is possible because there are sufficient numbers of these monitoring stations to form coherent groups for the aggregation step in the analysis. For surface monitoring of NMHC and also for all aircraft monitoring there are not sufficient stations, or, in the case of aircraft, the data are obtained too intermittently, to form groups for aggregation. It is therefore more difficult to implement the quasi-deterministic method of evaluation. In this case we form the most appropriate spatial and/or temporal averages of data to compare with observations. With this combination we attempt to maintain a balance of the use of all available data with the conscious desire to use the data *appropriately*. In this context, the rigorous comparison of data of dissimilar scales is considered *inappropriate*.

Data Preparation

The evaluation of surface-based O_3 concentrations follows along the lines of the quasi-deterministic analysis. We focus on the ability of the ROM to simulate O_3 concentration frequency distributions at groups of receptor locations during the simulation period. Each group of receptors shall have certain characteristics common to all members of the group. In our case we have chosen the observed frequency distribution of O_3 concentrations to be the common characteristic. The first step toward forming coherent groups of stations was to do a histogram-type analysis on the observed concentrations from the monitors in the surface network over the model simulation period. Only daytime (0800-1900 h, LST) hourly values were included in the histogram analysis. The values used in the histogram were normalized to represent the fraction of observations at each monitoring station that fell within a given concentration range for the daylight hours over the model simulation period. Similar data were compiled for the 214 monitoring sites and were then subjected to a cluster analysis to form coherent groups of sites based on frequencies o observed O_3 concentrations. Six suc groups were formed.

Air quality monitoring for NO_x and NO_2 is generally not as extensive as that fo O_3 in the U.S. As with O_3, nearly all o the sites were located within urban areas Using this network to verify the per formance of the ROM for predicting NO and NO_2 concentrations poses severa problems. The greatest problem is tha the nature of these chemical species ir the atmosphere is largely *primary*; that is the flux of emissions of the species is generally the largest contributor to the ambient concentrations. This being the case, the large spatial variations ir emissions patterns are expected to produce large spatial variations ir ambient concentrations, especially on the urban scale where the emissions hetero geneity is greatest. Urban-oriented moni tors, such as those available here for NO and NO_2, will capture concentratior patterns characteristic of the local area only. Monitors located in rural areas, o areas that might be more regionally representative, were generally not avail able in 1980. Similar scale problems exis for O_3 but are mitigated somewhat by the *secondary* nature of the pollutant and its smaller spatial concentration gradients For the quasi-deterministic analysis usec for the evaluation, data from NO_x and NO_2 monitors were aggregated ir individual urban areas with large NO source emissions. These were majo metropolitan areas that contained multi ple monitoring stations. The number o monitors in each group varied from 2 fo the Boston and Washington areas to 1(for the New York area.

Unlike the evaluation analyses fo surface concentrations of O_3, NO_2, and NO_x, the ROM evaluation for surface NMHC and all aircraft observations is more diagnostic than operational. There are a number of reasons for this. First the intermittency of the data poses some problems in setting up evaluative tests The NMHC and aircraft data are sparse ir both space and time. Group aggre gations, as performed for the othe species, are more difficult to perform Therefore, the comparisons of ROM pre dictions and observations are made in a deterministic manner for diagnostic pur poses to show whether there is a clea systematic bias in the way in which the ROM predictions of surface NMHC and concentrations predicted aloft compare with observations. Close statistical com parison of the data set is not warranted.

Summary of Results and Conclusions

The analysis using standard surface-monitored concentrations has shown that ROM2.0 predicts hourly O_3 concentrations above 80 ppb fairly well, and concentrations between 60 and 100 ppb particularly well. The average percentage of daylight hours (08-19 h, LST) over the simulation period showing concentrations above 80 ppb was 21.7% in the observed data set and 19.3% in the predicted data set for station groupings 1-3. These were the station groups which showed high to moderate values of observed O_3 concentrations. This statistic was compiled from data over the nearly 50 day simulation period spanning the summer of 1980. Cumulative concentration frequency distributions and histogram analyses of daytime hourly O_3 data have shown the model to typically predict a narrower overall concentration range than ambient observational data indicate. High hourly values are generally underpredicted and low values are overpredicted. This feature was evident in nearly all groups of monitoring stations. The cumulative frequency distributions treat the observed and predicted data sets independently for the simulation period. Table 1 compares the predicted and observed frequency distributions for daylight hourly O_3 concentrations for all station groups.

The model has shown good performance in predicting maximum daily O_3 concentrations averaged within station groups during the simulation period. The average daily maximum O_3 concentration over the simulation period for station group 1 was 88.1 ppb for the observed data set and 82.6 ppb for the predicted, a 6.2% underprediction. An examination of the time series of observed and predicted maximum O_3 concentrations, however, revealed that for this group there were often more significant differences between predictions and observations on any given day. At the 75th percentile level of daily maximum O_3 concentrations the ROM tended to underpredict by 30-50 ppb during episodic periods. For group 2, the average daily observed maximum was 76.9 ppb and the corresponding predicted value was 79.5 ppb, a 3.4% overprediction. For group 3, the observed maximum was 64.5 ppb and the predicted maximum was 70.1 ppb, an 8.7% overprediction. The time series analysis for group 3 showed that, on a daily basis, the median and 75th percentile levels of the maximum O_3 concentration showed better agreement between predictions and observations than did the group 1 results, with both over- and underpredictions occurring throughout the simulation period. The average performance over these three groups was a 2% overprediction of the daily maximum O_3 concentration.

A key indicator of model performance on the regional scale is the accuracy of simulating the spatial extent and location, as well as the magnitude of the pollutant concentrations within plumes from significant source areas on the regional scale, such as those emanating from major metropolitan areas within the model domain. ROM2.0 performance analyses in plume simulation were conducted for the Northeast Corridor subdomain of NEROS, including the major metropolitan areas from New York through Boston. During episodic conditions during the simulation period, the urban plumes from Washington, DC, Baltimore, New York and Boston could be clearly discerned in the model predictions. There was a systematic underprediction in the O_3 concentrations downwind of Philadelphia, often making it difficult to discern an integral urban plume. Figure 2 presents an example of the comparison of observed and predicted contours of maximum hourly O_3 concentrations over an episode.

Aggregate groups of monitoring stations for NO_x and NO_2 were formed for the major urban areas within the model domain. Results of model performance analyses for NO_x and NO_2 showed that the ROM2.0 significantly underpredicts NO_x and NO_2 among urban station groups at the 90th percentile of the cumulative concentration frequency distribution. Table 2 presents the average ratio of observed to predicted NO_x and NO_2 concentrations at the 50th and 90th percentile levels, averaged over all station groups. Spatial patterns of 06-09 h, LDT concentrations of NO_x and NO_2 were also analyzed for the Northeast Corridor sub-domain. Results of this analysis demonstrate that concentrations in this morning period show factors of 2-3 model underprediction. Exceptions occur in the Toronto and Philadelphia areas where the model predictions of NO_x and NO_2 appear to be significantly higher than than those in other areas. This may infer overestimates of NO_x emissions for these areas. A confounding factor in the performance analysis for NO_x and NO_2 was the scarcity of significant areas of concentrations above 5 ppb, either in the model predictions or in the observations, for large areas of the domain. The accuracy of standard NO_x and NO_2 monitors degrades greatly below this level, and observed data cannot be trusted to be accurate when concentrations are less than 5 ppb.

Model performance on surface-derived NMHC concentrations was performed in a diagnostic manner because of the inter-

Table 1. *Frequency Distribution of Daylight Observed and Predicted O_3 Concentrations Within Receptor Groups for the Period 14 July - 31 August 1980*

Group	Number of Sites	Percent of Daylight (08-19 h, LST) O_3 Concentrations Between: 5-20 ppb Obs.	5-20 ppb Pred.	21-40 ppb Obs.	21-40 ppb Pred.	41-80 ppb Obs.	41-80 ppb Pred.	>80 ppb Obs.	>80 ppb Pred.
1	35	8	1	19	7	39	66	34	26
2	39	15	1	23	9	41	69	21	21
3	64	16	1	28	9	46	79	10	11
4	54	22	1	37	12	36	80	5	7
5	20	42	2	38	16	19	73	1	9
6	2	0	0	26	4	63	93	11	3

Table 2. *Average Ratio (Obs./Pred.) Over Station Groups at 50th and 90th Percentiles of Cumulative Frequency Distributions*

	Daytime 08-19h, LST Percentile 50th	Daytime 08-19h, LST Percentile 90th	All Hours Percentile 50th	All Hours Percentile 90th
NO_x	1.8	2.3	1.9	2.5
NO_2	2.2	2.2	2.1	1.9

mittency of the measurements. All observed concentrations were calculated from a sum of the concentrations of individual species determined from gas chromatographic analyses of the NMHC samples. The ratios of observed to predicted NMHC concentrations averaged over the 06-09 h, LDT period were generally in the range of 4-7 for large urban areas, and the ratio for afternoon hours outside of major urban areas was in the range of 1-3. Despite the large model underpredictions, an analysis of the carbon distribution among reactivity classes in the observed and predicted data sets showed quite good agreement in the relative distribution of carbon mass between higher and lesser reactive classes. The biogenic portion of the NMHCs (as judged by the explicitly modeled isoprene compound) did not show up consistently in the measured NMHC samples. It was therefore difficult to judge ROM2.0 performance on an absolute basis. The model predictions of isoprene concentrations, however, were in good agreement with surface values quoted in the literature from special field studies.

Model performance analyses using aircraft data were also performed in a diagnostic manner because of the intermittency of the flights both in time and space. The analysis for O_3 concentrations shows that, in general, the overall background values in the planetary boundary layer do not build up as high under episodic conditions in the predictions as they do in the observations. Typically, regionwide O_3 concentrations predicted by ROM2.0 were in the 40-60 ppb range, while corresponding concentrations measured by aircraft monitors were 40-90 ppb, with the higher observed values occurring under strong episodic conditions. Most instances of regionally elevated O_3 values were underestimated by 20-30 ppb. Ozone aloft in the morning, upwind of urban areas, was generally underpredicted on days when aircraft measurements showed values greater than ~70 ppb. The regional background O_3 underprediction by the ROM was, however, not a totally pervasive phenomenon. Examination of predicted spatial patterns reveals the model to predict large areas of >80 ppb concentrations which build up, especially during daytime hours, and are transported through the model domain. These areas often merge with each other as they grow larger and are moved by the transport fields. Typically they shrink in size after the sun sets and gradually dissipate over the nighttime hours. Ambient patterns, as judged by the intermittent aircraft observations appear to show more widespread areas of >80 ppb O_3 concentrations under episodic conditions. These areas also apparently do not shrink in size and magnitude over the nighttime period as much as the ROM predictions suggest as judged by the early morning aircraft flights. Background values of NO_x were also underpredicted, by factors of 2-3. Observed values of NO_x concentrations were often in the 5-10 ppb range o

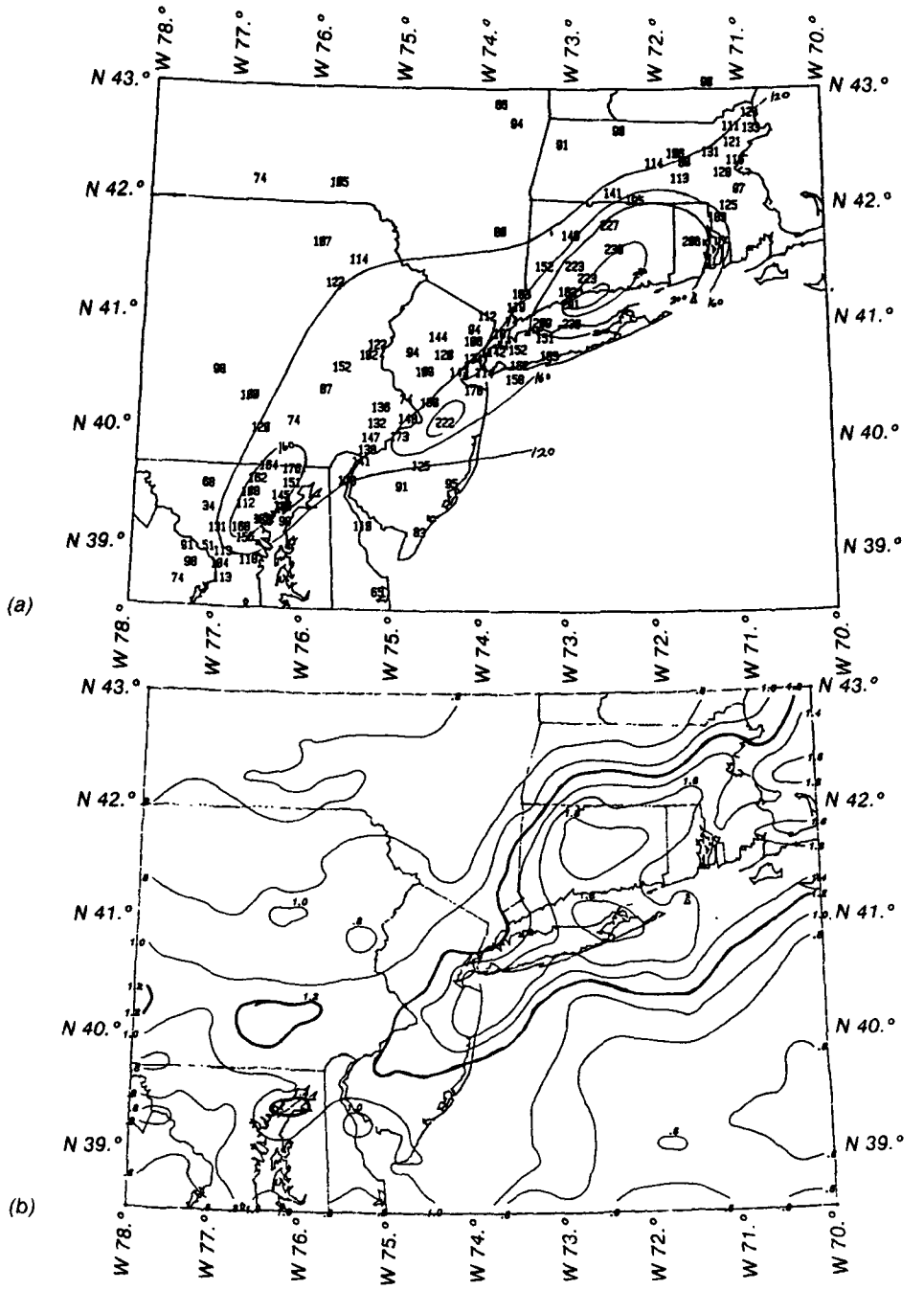

Figure 2. *Contours of maximum hourly O_3 concentrations over the period July 20-22, 1980 for (a) observed and (b) predicted data sets. Contours of observed data are in concentration units of ppb and contours of predicted data are in units of ppb/100.*

regional flights with predictions along the ame flight path typically less than 5 ppb.

Flights conducted in and downwind of urban areas showed ROM's O_3 concentration predictions to be very credible, especially when model wind fields provided transport along the Northeast Corridor from a southwesterly direction. The urban plumes simulated downwind of Boston, New York, and Washington especially showed good agreement with aircraft observations. The ROM tended to underpredict the maximum concentration areas within the plumes, although the degree of underprediction was not great for concentrations as high as 150-180 ppb. For observed concentrations over 200 ppb, as measured by aircraft monitors, the model underprediction was generally more substantial. Model performance or NO_x concentrations observed in the aircraft data shows ROM2.0 to generally underpredict, with observed to predicted ratios of +5 ppb NO_x in the range of 1-2. These ratios were generally somewhat less than those for the corresponding NO_x ratios in the aircraft background measurements and for the surface concentration analyses. Concentrations of NMHC aloft during the morning period were observed to be considerably lower (by up to a factor of 10 or more) than the morning surface NMHC concentrations. ROM2.0 still generally underpredicted the NMHC concentrations observed from aircraft samples, but by a smaller factor (1-3, as opposed to 4-7 in the surface data).

It is clear from our study that a critical factor in the utility of regional model results is the correct simulation of the location of regional concentration patterns, especially from urban area plumes. The important factor in the vicinity of major urban areas is the correct simulation of the mixing of upwind, over, and downwind air concentrations. Providing accurate wind fields to regional models will help ensure this. We are attempting to correct some transport biases discovered in ROM2.0 from this performance evaluation. The problem of localized circulation features that often exist on scales smaller than the existing meteorological monitoring network will still remain. For Northeast Corridor cities, these local circulations may greatly influence plume transport patterns. Ultimately, a nested scale predictive meteorological model may be needed to generate the transport fields used in ROM to most accurately model plume transport on regional scales in areas of complex flow.

While the overall performance of ROM2.0 for predicting O_3 concentrations has not shown sufficient cause for rejecting use of the model in regulatory analyses,[2] a number of issues have arisen that prevent recommending use of the model in a simple, unassisted manner for studies involving violations of the O_3 air quality standard. These include day-to-day variability of model bias results, systematic transport biases seen in the spatial patterns, and perhaps most importantly, a general underprediction of regional background O_3 concentrations under episodic conditions. For applications studies, such as the upcoming ROMNET (Regional Oxidant Modeling for Northeast Transport) project, the ROM results must be used with care. Episodes must be chosen for which the model is operating at its most consistent level with best fidelity. Careful analysis of the simulated transport patterns and background O_3 levels must be made for a given episode before model results can be used for subsequent analysis. Sensible guidance for use of ROM results will require strong interaction of the modelers, with their insights and ability to interpret model results, and the regulatory users of the information.

This model performance study has also raised some key questions concerning the data base with perhaps the greatest inherent uncertainty, the emissions inventory. The large and consistent underprediction of NMHCs seems to indicate deficiencies in the hydrocarbon portion of the inventory. These deficiencies may be across the board on all sources since the normalized carbon fractions seem to agree fairly well in the observations and predictions of NMHC reactive classes. Missing or incorrect temperature sensitivities of both biogenic and anthropogenic precursor emissions may be responsible, in part, for the failure of the model to build the background tropospheric O_3 concentration levels as high as they have been observed from field study results during episodic conditions. Current efforts toward the development of a 1985 base year emissions inventory by NAPAP will take steps to rectify some of the known deficiencies in the 1980 inventory.

To complete the picture of model performance evaluation, there must be a corresponding model sensitivity study. A systematic sensitivity study of a complex regional model such as ROM is a very resource-intensive task, and was outside the scope of the present effort. Results from a sensitivity study will, among other purposes, help guide the application of resources to narrow the uncertainty in those required data for which the model is most sensitive. Results from a model evaluation study, such as this one, will help to provide invaluable guidance toward establishing some of the critical sensitivity tests. We hope to begin such a sensitivity study soon. The combination of performance evaluation and sensitivity results is key to understanding ***why*** a model has performed in the manner it has under varying conditions.

Model performance is a continuing task. Once a model is used in production for multiple applications some type of performance study is required for each application. Even though the science embodied in the model may be up to contemporary standards, the stochastic nature of the atmosphere can cause variations in model performance from location to location and from time to time. Detailed data bases, including aircraft based measurements, such as those used in this study, are not commonly available. However, the methods of evaluation demonstrated here may also be applied to routinely available data bases for less rigorous evaluation exercises.

[2]Model evaluation studies, such as this one, will not indicate model acceptance, but rather will only indicate whether model rejection is warranted. Model acceptance comes only after non-rejection in many evaluation studies ovier tme.

The EPA author, ***Kenneth L. Schere*** *(also the EPA Project Officer, see below), is on assignment to the Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC 27711, from the National Oceanic and Atmospheric Administration, Richard A. Wayland is with Computer Sciences Corporation, Research Triangle Park, NC 27709.*

The complete report, entitled "EPA Regional Oxidant Model (ROM2.0): Evaluation on 1980 NEROS Data Bases," (Order No. PB 89-200 828/AS; Cost: $36.95, subject to change) will be available only from:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650

The EPA Project Officer can be contacted at:

Atmospheric Research and Exposure Assessment Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300

EPA/600/S3-89/057