EPA/600/3-85/067

AUGUST 1985

EPA REGIONAL OXIDANT MODEL: DESCRIPTION AND EVALUATION PLAN

ATMOSPHERIC SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

EPA REGIONAL OXIDANT MODEL: DESCRIPTION AND EVALUATION PLAN

by

Kenneth L. Schere
Meteorology and Assessment Division
Atmospheric Sciences Research Laboratory
Research Triangle Park, North Carolina 27711
and

Allan J. Fabrick
MEF Environmental, Inc.
9705 Burnet Rd., Suite 311
Austin, Texas 78758

ATMOSPHERIC SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

DISCLAIMER

This report has been reviewed by the Atmospheric Sciences Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

AFFILIATION

Mr. Schere is on assignment to the Meteorology and Assessment Division, Atmospheric Sciences Research Laboratory, from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

ABSTRACT

The U.S. EPA Regional Oxidant Model and NECRMP/NEROS data base are described. The model incorporates a comprehensive description of the physical and chemical processes thought to be important to tropospheric O_3 production on 1000 km scales. The data base employed for the first application of the ROM was collected during the summers of 1979 and 1980 in the Northeast U.S. It contains meteorological and air quality data both from regular monitoring networks as well as enhanced networks or special field project measurements made during that period.

The evaluation procedure that will be used to determine the ROM performance on this data base is outlined. A number of episodes will be simulated from the period July 23 through August 16, 1980 for which performance statistics will be developed. The evaluation of any given day within an episode will proceed in 2 distinct stages. The first stage will focus on model performance for an individual model realization, irrespective of all other realizations. Model realizations for a given day are functions of the possible flow fields that existed for the day. The second stage will attempt to evaluate model performance using the full probabilistic abilities of the ROM that consider all realizations concurrently. The focus of the evaluation will be on O_3. The exact pathway through the evaluation study will be determined by the resources available at the time.

CONTENTS

Abstract iii
Figures vi
Tables vii
Acknowledgements viii
1. Introduction 1
2. The ROM System 7
Overview 7
Preprocessor descriptions 10
Core model and postprocessors 19
3. The NECRMP/NEROS Data Base 23
4. Model Evaluation 36
Overview 36
Data base analysis 37
Evaluation by individual realization 41
Evaluation over all realizations 46
Further testing 66
5. Summary and Conclusions 69

References 73
Appendices
A. Surface Monitoring Sites 75

FIGURES

Number		Page
1	ROM modeling domain for the Northeast U.S. application . . .	3
2	ROM vertical layer structure	4
3	Data flow for the ROM system	8
4	Network of ROM processors	11
5	P11 processor solution schemes	17
6	Typical Lagrangian sampling flight track	26
7	Locations of upper air sounding stations	28
8	Locations of surface meteorological stations	29
9	Locations of NECRMP O_3 monitoring sites	31
10	Locations of SAROAD O_3 monitoring sites	32
11	Surface O_3 frequency distribution from SAROAD sites	34
12	Structure diagram of the ROM evaluation by individual realization	47
13	Hypothetical concentration distributions from the ROM . . .	57
14	Box-plot of predicted O_3 concentrations from urban grid cells	59
15	Source-receptor combinations in ROM evaluation over all realizations	61
16	Structure diagram of the ROM evaluation over all realizations	63
17	Emission reduction analysis using the ROM and urban scale models	70
18	Emission reduction analysis using the ROM alone	72

TABLES

Number		Page
1	NEROS Aircraft Instrumentation	21
2	Sample Data Channel in ROM System	24

ACKNOWLEDGEMENTS

Discussions with Dr. Robert Lamb and Dr. Robin Dennis concerning model evaluation techniques and the Kriging analysis for observed data were extremely helpful in the preparation of this document.

SECTION 1

INTRODUCTION

The perception of air quality problems has increased in recent years from an urban scale to a larger regional scale as the effects of multi-day and long-range transport of air pollutants have been studied and understood. During the past five years the U.S. Environmental Protection Agency (EPA) has undertaken a model development effort to produce an air quality simulation model capable of treating all of the chemical and physical processes that are thought to affect the concentrations of air pollutants over several day/1000 km scale domains. The first generation version of the EPA Regional Oxidant Model (ROM) is now operational. Among the processes it treats are horizontal transport, atmospheric chemistry, nighttime wind shear and turbulence episodes associated with the nocturnal jet, cumulus cloud effects on vertical mass transport, diffusion and deposition, subgrid scale chemistry processes, and emissions of natural and anthropogenic precursors. These processes are simulated in a three-dimensional (3-D) Eulerian framework with 3 1/2 vertical layers extending through the boundary layer and the capping inversion or cloud layer. In the present configuration of the ROM domain, horizontal resolution is approximately 20 km.

Lamb (1983, 1984) has described in detail the theoretical basis and design of the ROM. The modeling system is actually a network of interconnected preprocessors whose output data are reformatted for input to the simulation model. The following section summarizes the functions of the components in the modeling system.

The chemical kinetics mechanism, cumulus cloud vertical flux formulation, and nocturnal jet transport solution are all preliminary versions in the 1st generation ROM. Within one year, the production version (2nd generation) of the ROM will be completed and operational with final versions of the components cited above. It is this 2nd generation ROM that will be used in the model evaluation study outlined in this report.

The data base that will be used in the evaluation study is that of the Northeast Corridor Regional Modeling Project (NECRMP). The NECRMP ambient data base consists of measurements made during 4 EPA field measurement programs in 1979 and 1980 in the northeastern U.S. These include the 1979 Northeast Regional Oxidant Study (NEROS I), the 1980 NEROS II, the Persistent Elevated Pollution Episodes Study (PEPE), and the 1980 Urban Field Studies. Together, these studies have provided a variety of air quality and meteorological measurements on regional, urban, and site-specific scales (Freas, 1983). Also, a parallel effort within the NECRMP program assembled a complete source emissions inventory specifically addressing the ROM requirements (EPA, 1982). This emissions inventory will be supplemented by updated emissions for the U.S. and southern Canada contained in the National Acid Precipitation Assessment Program (NAPAP) data base and the revised inventory will be used in the ROM evaluation study.

Figure 1 shows the extent of the model domain for the NECRMP application. The individual grid cells are 15 minutes of longitude wide in the E-W direction and 10 minutes of latitude in the N-S direction, or about 18.5 km square cells. There are 60 cells E-W and 42 cells N-S giving a total horizontal extent of approximately 1100 km by 780 km. The vertical structure of the ROM consists of 3 1/2 layers as shown in Figure 2. The bottom (1/2) layer is actually a diagnostic surface layer less than 100 m deep where surface deposition and subgrid scale chemical effects are modeled in a diagnostic manner. Layers 1 and 2 are prognostic model layers and extend through the depth of the well mixed layer during the day and the surface inversion and old mixed layer at night. The top prognostic model layer extends up to 1 km above the top of the mixed layer and includes any convective cloud elements. Some of the physical and chemical processes being simulated in each layer are shown in Figure 2 for the daytime situation.

This report summarizes the ROM system and the NECRMP data base and presents a plan for an evaluation study of the ROM. The ROM model evaluation must account for the stochastic nature of the predicted concentration field. This is a result of the multiple wind field realizations (interpo-

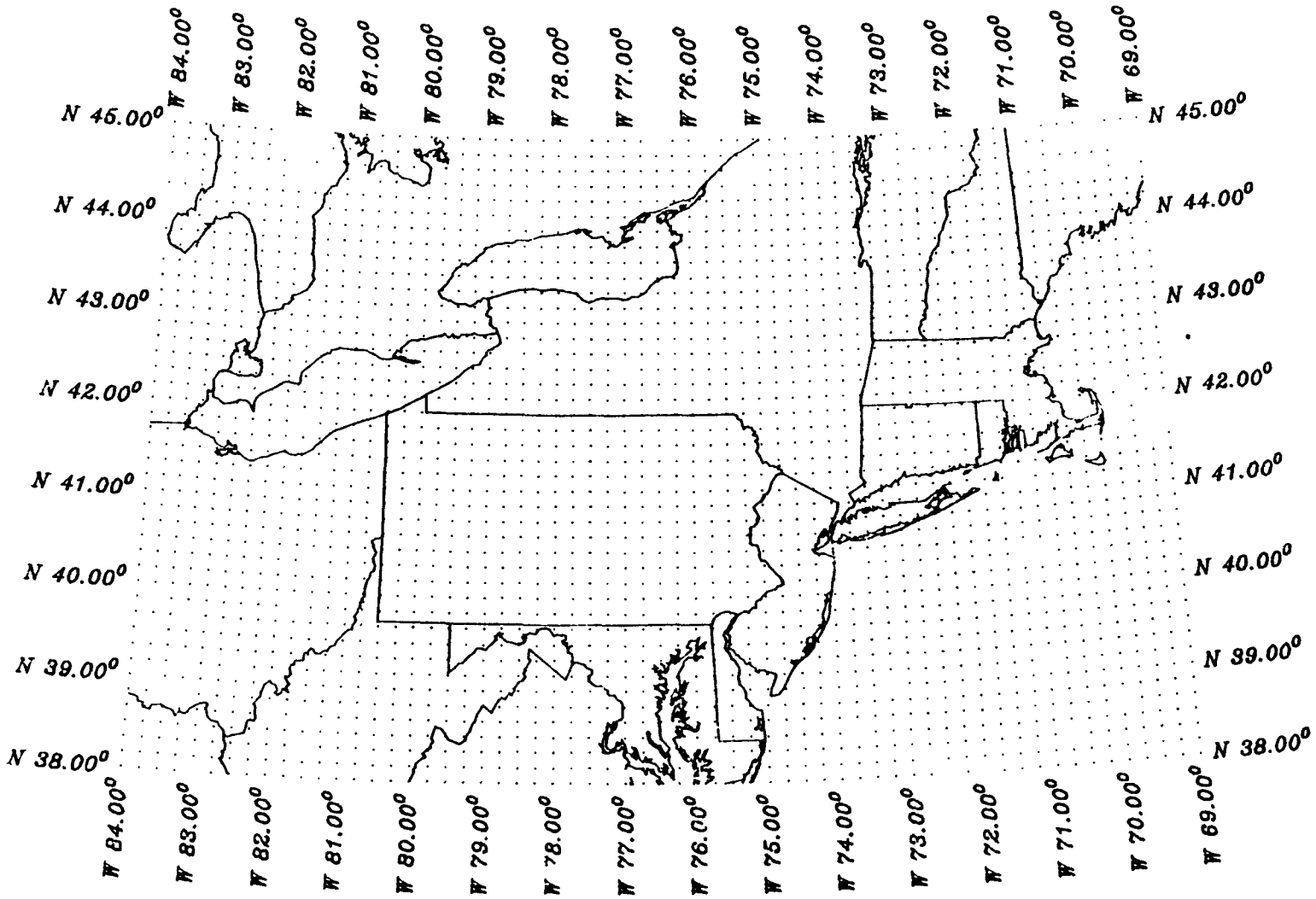


Figure 1. ROM modeling domain and grid resolution for the Northeast U.S. application.

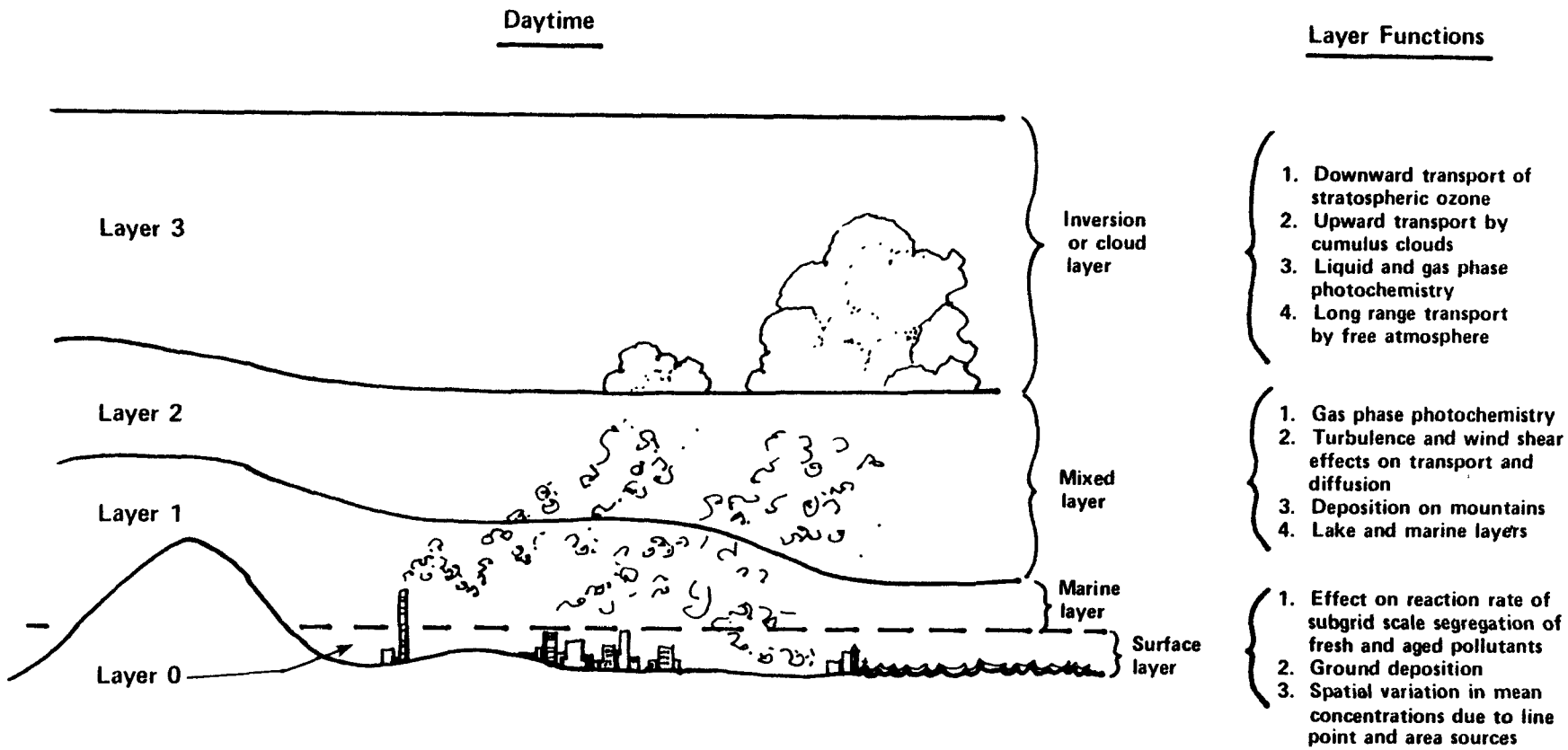


Figure 2. ROM vertical layer structure during the daytime.

lations) that the ROM wind field processor can generate from a given set of hourly wind observations. These multiple wind fields are each consistent with both the observations and physical laws governing atmospheric flow and are assigned probabilities of occurrence based on the inherent kinetic energy contained within the field. From the resulting set of wind field realizations used in a corresponding set of ROM simulations the model generates a concentration probability distribution for a given receptor site within the domain instead of a single concentration value. When all members in the set of realizations are considered concurrently, this aspect of the model prediction makes the ROM different than most other air quality simulation models and presents more of a challenge to the evaluation effort.

The first and second moments of the predicted concentration distribution can be used to determine the expected frequency with which the observed concentrations should fall within a given interval. The utility of the model for regulatory use is measured partly by the width of this interval. The second moment of the distribution (the concentration variance) is the parameter that defines this width. It is a measure of the inherent transport uncertainty of the model. If the width is very large, the model may provide no more information than one would gather by guessing the expected concentration. This is true even if the model is shown to be accurate in other respects. Therefore part of the ROM evaluation should consist of an analysis of the predicted concentration variances. For regulatory use, the predicted concentration for maximum ozone at a given receptor location for a single realization is the primary parameter of interest. This value may be compared to the corresponding observed value. Also, if groups of receptors can be identified having similar predicted concentration distributions, the observed distribution composed of the measurements from each receptor within the group can be compared to the predicted distribution from the group. Such groups might be stratified by receptor location with respect to downwind distance from major source emissions areas. Section 4 discusses these and other potential evaluation methods for analyzing the ROM predictions in the NECRMP application.

Base case evaluation studies with the ROM will be conducted for 4 or 5 regional smog episodes from the data base. The individual episodes range from 2 to 6 days in duration, and are mostly from the 1980 summer period. The exact choice of modeling episodes is tentative as of yet because the analysis of the air quality data base has only recently begun. A full 2 week simulation will also be performed.

SECTION 2

THE ROM SYSTEM

OVERVIEW

The ROM described here is the 2nd generation version of the model. It is designed to simulate air pollution chemistry and dispersion over a 1000-kilometer-by-1000-kilometer area for multiple day periods. The simulation of photochemical air pollution over such scales is a complex problem. The ROM reflects this complexity, consisting of some 25 programs accessing almost 100 files. The overall structure of the data flow for the ROM modeling system is illustrated in Figure 3.

The input files contain the "raw" data accessed by the modeling system. Actually, these data files are produced by a combination of manual and computer manipulation of raw data. However, the input data files constitute a boundary between the ROM and the rest of the world, and therefore define the input data for the model. The initial preprocessors take the "raw" input data and transform it for use by other preprocessors. These data are transferred between preprocessor programs through the processor input files (PIF). The preprocessor programs are used to develop intermediate parameters that are then transformed into core model inputs. Examples of these intermediate parameters are the wind fields, emissions, and turbulent fluxes. The preprocessors transfer the processed data through the PIF files and the model input files (MIF).

The data contained in the MIF files are converted into the form required by the CORE model through the execution of the b-matrix compiler (BMC) program. The BMC translates the parameter fields in the MIF files (layer thicknesses, horizontal winds, interfacial fluxes, deposition velocities, etc.) into the matrix and vector elements necessary to operate the CORE model. The BMC transfers the information needed by the CORE model via the core input files.

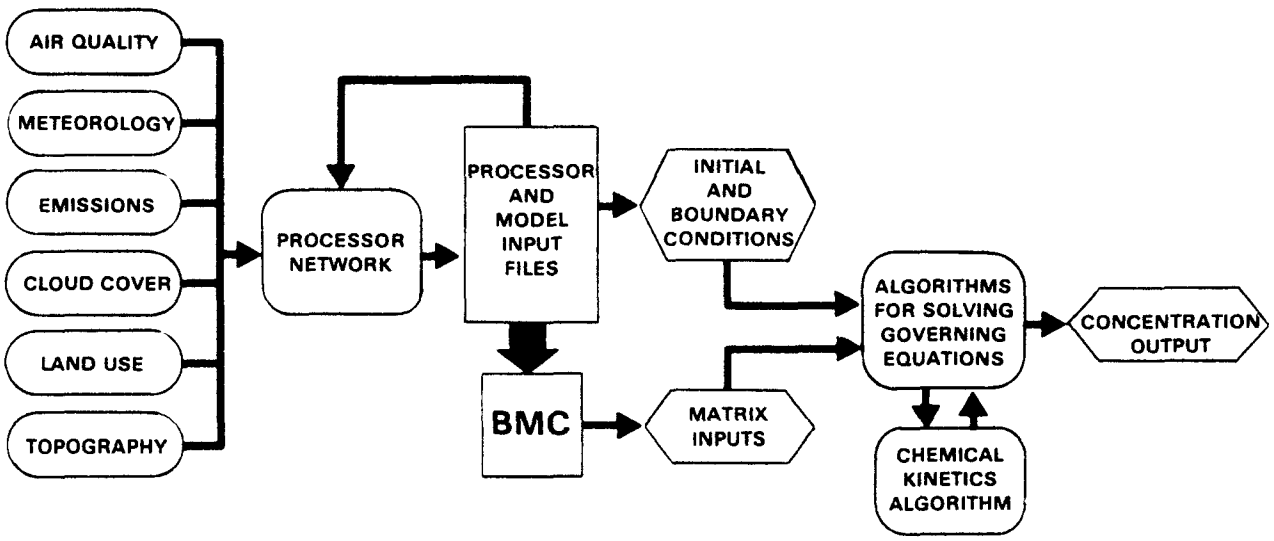


Figure 3. Data flow for the ROM system.

The CORE model is the computer language analogue of the differential equations that describe the processes involved in the chemistry and dispersion of regional photochemical air pollution. The CORE model is expressed in fundamental mathematical form. All inputs to the model are matrices and vectors whose elements are composites of meteorological parameters, chemical rate constants, etc.

The output of the CORE model consists of the layer-averaged pollutant concentrations for every grid cell every 30 minutes. However, these data are not the final model results. The ROM incorporates two new concepts in air pollution modeling. The first is a method for simulating the physical and chemical processes that occur within about the first 100 meters of the ground. The result of this scheme is that both the ground-level concentration and the root-mean-square concentration variation within each selected cell are produced. These data can be extracted using the results of the CORE model output (contained in the model output files) and parameters contained in the model input files for a time period of 30 minutes or longer.

The other concept is the incorporation of the uncertainty of the wind field representation. It is known that many different wind fields can be constructed that match the observed wind data and empirical and theoretical constraints. In the ROM, a family of possible wind fields is used in the model and leads to a result of an ensemble of concentration fields. Therefore the model results will consist of a distribution of concentration values and associated probabilities of occurrence for each grid cell for each time step. The implementation of the multiple wind field concept requires that the model be run many (~10) times for each simulation. ROM postprocessor programs will be used to transform the results of the individual CORE outputs into ensemble concentration distributions. Other model postprocessors will transform the layer 1 concentration values into ground level (layer 0) concentration distributions. These data constitute the model results. Measures of model accuracy will be made by comparing the model predictions of concentration probability distributions with observed

measurements as well as examining individual model runs (realizations) on a single wind field.

PREPROCESSOR DESCRIPTIONS

The model preprocessors transform the "raw" input data into the matrix and vector elements required by the CORE model. Some of the preprocessors are quite simple; others such as the wind field processors are complex computer models in themselves. All of the preprocessors communicate with each other along data flow channels through the use of the PIF and MIF files. The preprocessor system is designed to proceed in stages (0-6) with data developed in one stage used as input for the next stage. Figure 4 is a schematic diagram of the network of ROM processors and data channels. In this section an overview of the model processors proceeding from stage 0 through stage 7 is presented.

Stage - 0

Processor P03 - This processor performs standard operations on the surface meteorological data to reformat them for the higher level processors in the network. The processor reads the surface meteorological data that are usually available every three hours and interpolates the data into hourly values. Some stations report every hour, in which case temporal interpolation is not necessary. The variables produced by this preprocessor include surface wind velocity components, surface mixing ratio, relative humidity and dew point, surface virtual temperature, geopotential of the 1000mb surface, time rate of change of the 1000mb geopotential, sea level pressure, surface potential temperature, and the surface air density.

Processor P25 - This processor converts land use percents into land use fraction for each grid cell in the model region.

Processor P05 - This processor transforms the raw cumulus cloud digitized satellite data into hourly gridded sky coverage of this type of cloud.

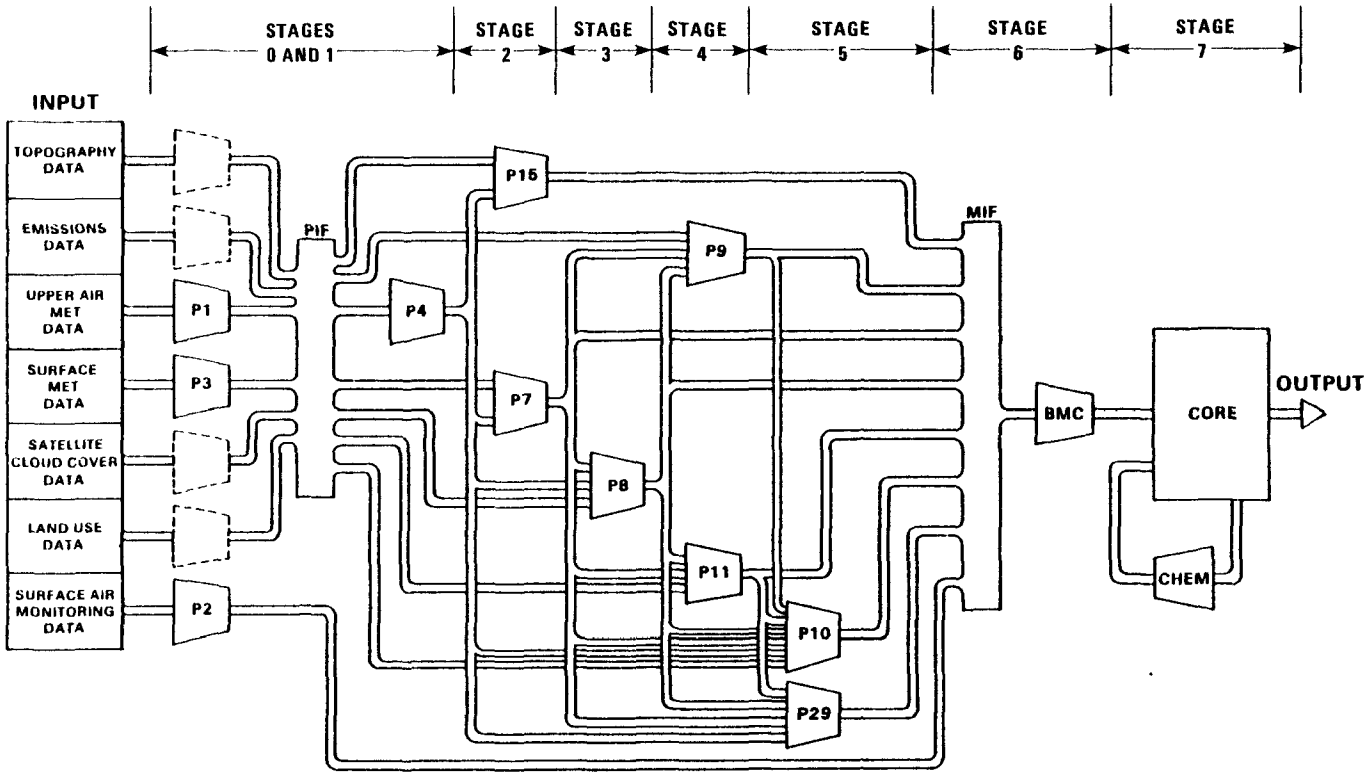


Figure 4. Schematic illustration of the principal components in the ROM processor network.

Processor P06 - This processor reads the input terrain elevation data and computes the average elevation for each cell in the modeling region. The preprocessor also computes smoothed elevations for use by the P08 processor.

Processor P13 - This processor computes the gridded line source length totals for each cell.

Processor P14 - This processor prepares the files containing stack descriptions and emissions for all major point sources, and hourly gridded emissions for all other sources.

Processor P16 - This processor interpolates the rawind upper air data into hourly observations.

Processor P19 - This processor reads the hourly surface meteorological observations and computes hourly gridded fraction of cloud coverage for all cloud types combined.

Processor P20 - This processor uses the ROM grid definition to produce required matrix elements for P11 when it tranforms gridded divergence data to Fourier series expansions.

Processor P21 - This processor contains indices that classify each grid cell in each layer to one of 5 possible sets of initial concentration profiles based on the magnitudes of measured pollutant loadings of O_3 and precursor species.

Stage - 1

Processor P01 - This processor performs a number of standard operations on the interpolated hourly rawind data for use by other, higher level, processors. Wind vectors, mixing ratio, relative humidity, dew point, temperature, potential temperature, static stability, air density, and pressure are computed

for every 50 meters from the surface to 5000 meters. The geopotential of the 850, 700, 500, and 1000mb surfaces are also computed.

Processor P02 - This processor determines the initial concentrations throughout the grid and the upper and lateral concentrations at the grid boundaries for all pollutant species simulated by the CORE model. The first-generation model assumes typical "clean atmosphere" concentrations for species entering the grid boundaries. The initial concentrations are specified using five sets of preequilibrated initial concentrations. Clean atmosphere concentrations are used to define rural initial conditions. For grid cells in or near urban areas, observed data are used to guide the manual selection of one of the five sets of initial concentrations that best represents the initial conditions in that grid cell. The second-generation model will likely employ the same procedure. The matrix defined in P21 classifies each cell according to 1 of the 5 sets of initial concentrations.

Processor P04 - This processor computes the surface roughness, Obukhov length, surface heat flux, and the friction velocity for each grid cell for each hour.

Stage - 2

Processor P15 - This processor computes the effective deposition resistance and deposition velocity for each pollutant species every hour for each grid cell. The processor uses the surface roughess, friction velocity and Obukhov length computed by processor P04 in stage 1.

Processor P07 - This processor computes terrain penetration factors (Lamb, 1983), the height of the surface based inversion layer, and the wind field in the surface inversion. Within the processor is a complex numerical model that simulates the motion of a viscous, hydrostatic fluid of constant density over irregular terrain.

The processor first calculates $\Delta\rho_1$. This parameter is used as an indicator of the presence of a surface based inversion (1, if surface based inversion is present, and 0, otherwise). The value of $\Delta\rho_1$ is computed hourly for use by the wind field processor P11. The model assumes that an inversion layer forms over the entire model domain at the hour that the surface heat flux averaged over the whole region becomes negative, and it disappears everywhere at the hour that the mean "temperature deficit" (defined by equation 7-11, Lamb, 1983) first becomes negative.

For time steps occurring when $\Delta\rho_1$ equals 1, preprocessor P07 computes the growth of the surface inversion and the layer averaged velocity field within the inversion. Since the surface wind observations during these conditions either are erratic or reflect purely local conditions, the inversion layer flow model uses observed pressure gradients to initialize the flow field and then drives the flow model using observed temperatures. No wind observations are used in the calculation of the inversion layer wind field.

For all time steps, processor P07 computes the elevation of the top of layer 1, and sets the depth of layer 0 to one-tenth of the value of layer 1. The processor also converts the height of layer 1 and the height of the virtual surface into pressure coordinates. The height of the virtual surface is the terrain elevation plus the depth of the surface inversion when an inversion is present, and is equal to the terrain elevation otherwise.

A final computation performed by this processor is the estimate of the fractions of the top surfaces of layers 0 and 1 that are penetrated by terrain. Currently the value of the terrain penetration fraction for layer 0 is set to the value calculated for level 1.

Processor P18 - This processor computes matrix elements needed for Fourier series expansions of wind fields evaluated at all rawin and surface meteorology stations.

Stage - 3

Processor P08 - This processor determines the elevations of the top of layers 2 and 3, computes the mean vertical velocities on these surfaces, and estimates cloud updraft velocities and other parameters (listed below) needed to specify pollutant fluxes across the surfaces of layers 2 and 3.

The following parameters will be calculated by processor P08 every time step for each grid cell in the model domain: the plume entrainment velocity, the cumulus updraft speed, the vertical air speed on the interface between layers 2 and 3, the elevation of the interface of layers 2 and 3, the elevation of the top of the grid (the top of layer 3), the volume flux through the top of the grid, the local time derivative of the elevation of the interface between layers 2 and 3, the thickness of layer 3, the thickness of layer 2, the pressure at the top of layer 3, the pressure at the top of layer 2, and an estimate of the cumulus flux partition function (see equation 5-44, Lamb, 1983).

Stage - 4

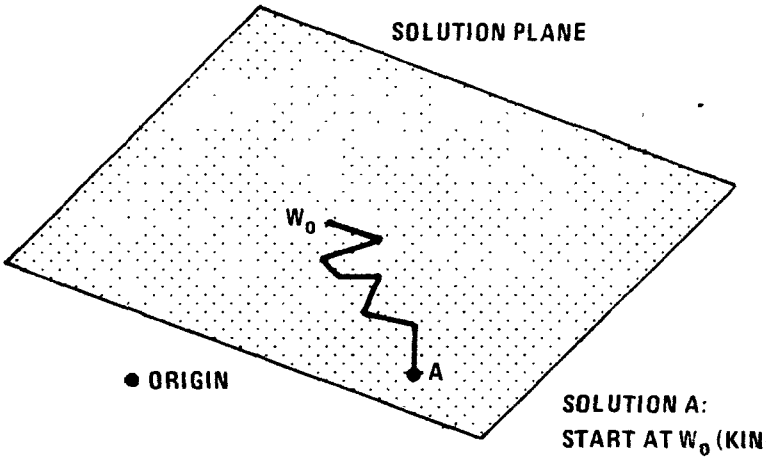
Processor P09 - This processor prepares the information necessary to correct the chemical reaction rate constants for variations in atmospheric density, temperature, cloud cover, and solar zenith angle. Layer-averaged density correction factors and layer-averaged virtual temperatures are used by the CHEM portion of the CORE model to correct the chemical reaction rate constants. The solar zenith angle and the cloud cover correction factor are used to modify the photolytic rate constants used in CHEM.

Processor P11 - This processor generates a family of vertically integrated horizontal winds for each of the model's layers. During the daytime hours (i.e., when $\Delta\rho_1 = 0$ this processor produces wind fields for each of the model's three layers. At night (i.e., $\Delta\rho_1 = 1$) the processor produces wind fields for layers 2 and 3 only. The flow field generated by the P07 processor for layer 1 is passed to P11 for combination with the wind fields derived there.

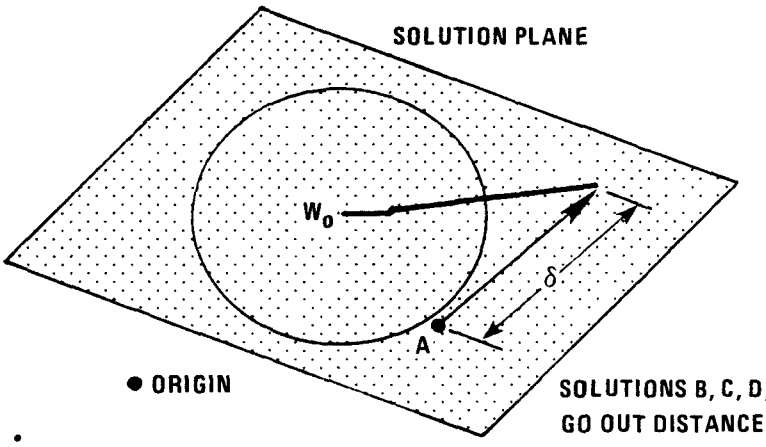
This processor uses a novel numerical scheme to produce layer-averaged wind fields that are: (1) consistent with observed wind data, (2) consistent with the mass-continuity equation and derived gridded wind field divergences, and (3) consistent with the theory and observations that, in two dimensional fluid flow, kinetic energy is partitioned among the spatial fluctuations in the flow proportional to the inverse cube of the wave number of the fluctuations (minus-3 law).

P11 calculates a family of wind fields consistent with the observations and theory. The base wind field, W_0, is obtained by calculating a wind field whose average kinetic energy is the same as the average kinetic energy of the observations, whose components match the measured winds, and whose divergence is consistent with specified gridded wind field divergences. An iterative convergence procedure is then used to search for a wind field that obeys the minus-3 law and is close to the initial wind field solution, W_0. This wind field is designated as field A.

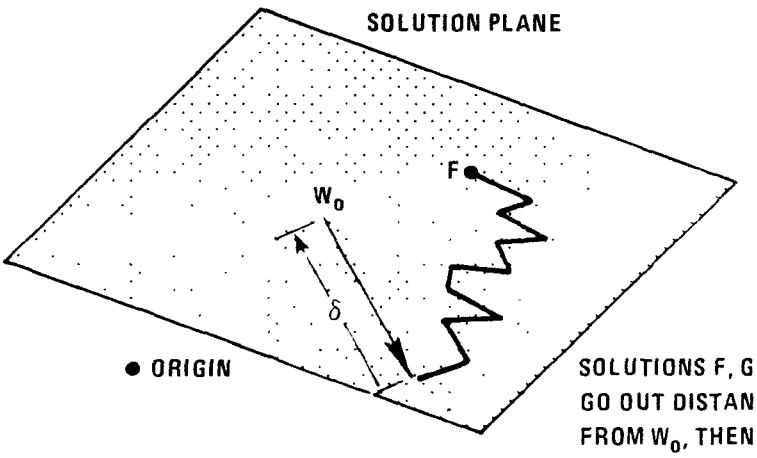
The next four wind fields, B, C, D, and E, are computed by moving a specified distance in a random direction from the first wind field, A, (in phase space) and then projecting back to the energy sphere whose intersection with the solution plane contains the first wind field solution A. This procedure results in wind fields whose average kinetic energy is identical to the first wind field solution but whose kinetic energy partition does not exactly obey the minus-3 law. Wind fields F, G, H, I, and J are obtained by starting the iterative convergence procedure at locations that are selected by moving a specified distance on the solution plane (in phase space) in a random direction from the initial solution point, W_0. The random displacement is the same as for solutions B, C, D and E and was determined by experimentation. Wind fields F, G, H, I, and J have average kinetic energies that differ from observations but the kinetic energy is partitioned in a manner consistent with the minus-3 law. A detailed description of the methodology in P11 can be found in Lamb (1984). A graphical illustration of the wind field calculation procedures are presented in Figure 5.



SOLUTION A:
START AT W_0 (KINETIC ENERGY = OBSERVED)
APPLY POWER LAW CONVERGENCE
TO GET SOLUTION A.



SOLUTIONS B, C, D, AND E:
GO OUT DISTANCE δ IN A RANDOM DIRECTION
FROM A; THEN PROJECT BACK TO 'ENERGY
SPHERE' CONTAINING A.



SOLUTIONS F, G, H, I, AND J:
GO OUT DISTANCE δ IN A RANDOM DIRECTION
FROM W_0, THEN APPLY POWER LAW
CONVERGENCE TO GET F TO J.

Figure 5. Illustration of the wind field processor (P11) solution schemes.

The ROM will be run ten times, with wind fields A through J, for each scenario simulation. The results of the individual runs and their associated probabilites will be used to construct the concentration probability function at each time step and grid cell.

During daylight hours the surface wind velocities are used as input to the calculation of the layer 1 averaged wind fields. The upper air wind velocities are used as input to the calculation of the layer 3 averaged wind fields. Both surface and upper air observations are used for layer 2. Because of the numerical method employed in this processor no pair of surface wind observations can be closer than about 25 kilometers. Since there are a large number of surface meteorological stations in the modeling region some surface stations will be excluded based on the 25 km criteria. In practice, this means that the availability of surface wind observations is not likely to be an important criteria in the selection of model validation scenarios. All upper air data within the model will be used for the wind field calculations.

Processor P12 - This processor calculates the parameters required to describe the material flux between layer 0 and layer 1 and between layer 1 and layer 2. It also calculates the horizontal eddy diffusivities in all layers.

Processor P17 - This processor calculates the local time derivatives of the cell heights in layers 1 and 2.

Stage - 5

Processor P10 - This processor transforms the emission data into the source strength functions required by the b-matrix compiler. The source strength functions partition the point source and area emissions into layers 0, 1 and 2. Processor P10 also computes the plume volume fraction which is used in the layer 0 equations to parameterize subgrid scale chemistry effects.

Processor P29 - This processor transforms the velocity and diffusivity fields into parameters required by the CORE model. The velocity, diffusivity and cell volumes are used to compute an effective velocity field that is then used to compute "back track" trajectories. The coordinates of the "back track" trajectory origin associated with a given cell for a given time step are used by the CORE model to solve the advection-diffusion equation.

Stage - 6

Processor P28 - This processor, the b-matrix compiler (BMC), translates the variables residing on the MIF files into the elements of the "b-matrix" defined in Chapter 9 (Lamb, 1983). The variables prepared by processor P29 are passed through the BMC and saved on the file MFBMAT. The MFBMAT file and the MFIBC file written by the initial and boundary condition processor (P02) constitute the input data for the CORE model.

CORE MODEL (Stage - 7) AND POSTPROCESSORS

The CORE model is the main program of the ROM. It reads the MFIBC and MFBAT data files and computes the concentration of each pollutant specie in layers 1, 2 and 3 in each grid cell for each time step (30 minutes). Since the CORE model operates on one wind field realization at a time the model will be run many times (~10) for each simulation scenario.

The concentration distribution functions, taking into consideration results from all realizations, for selected pollutant species in layers 1, 2 and 3 in each grid cell at each time step will be calculated by a postprocessor. This program will read the concentration fields produced by the individual CORE model runs and the probability weighting functions assigned to the various realizations (Lamb, 1984). Pollutant concentrations in each cell for each time step are described by a concentration value from any given realization or a concentration distribution when values from all realizations are considered together. Concentrations in layer 0 are computed

using the current concentrations for layer 1 and parameters contained in the MIF files. The equations used for the evaluation of the layer 0 concentrations and their root-mean-square variation are given in Chapter 5 (Lamb, 1983).

As an example of how input data flows through the preprocessor chain, a section of a data stream, or channel, in the ROM system is given in Table 1. The section that appears there follows the surface and upper air meteorological processing from the raw input data through the first 2 stages of processing. Data processing must occur in 6 stages of processors before the simulation model can be executed. The processed parameters shown in the table would flow through principal and branch channels for all stages. Similar data channels mark the flow of air quality, emissions terrain and land use, and miscellaneous parameters. Lamb (1984) discusses the data channels and the details of the data processing in detail.

TABLE 1. SAMPLE DATA CHANNEL IN ROM SYSTEM

SELECTED STAGE 0 PROCESSORS

Raw Input	Description	Processor	Output Parameter	Description
SID	surface station id and location	P3	u,v	wind components
		P19	q	mixing ratio
CLD	surface observations of cloud cover, amount, heights	(Surface) Meteorological Processors	RH	relative humidity
			T_d, T_v, T_p	dew point, virtual, and potential temperature
P,Ps	station pressure and sea level pressure			
WS	surface wind speed		ϕ_o, $\dot{\phi}_o$	height and rate of change of 1000 mb surface
WD	surface wind direction			
T,T_d	surface temperature and dew point		P_s	sea-level pressure
			ρ	air density
			σ_c	fractional coverage of all cloud types
UID	upper air station id and location			
P_a	pressure level aloft	P16 (Upper Air Meteorological Processor)	u_a,v_a	wind components aloft
z_a	height aloft		T_a,T_{da}	
T_a,T_{da}	temperature and dew point aloft		Pa	
RH_a	relative humidity aloft			parameters above are interpolated regularly in time
WS_a	wind speed aloft			
WD_a	wind direction aloft			

SELECTED STAGE 1 PROCESSORS

Input Parameter	Processor	Output Parameter	Description
u,v	P4 (Surface Meteorological Processor)	L	Obukhov length (gridded values)
q		z_o	surface roughness (gridded values
RH		u*	friction velocity (gridded values)
T_d, T_v, T_p		Q	surface heat flux (gridded values)
ϕ_o, $\dot{\phi}_o$			
Ps			
ρ			
σ_c			

TABLE 1. (continued)

SELECTED STAGE 1 PROCESSORS

Input Parameter	Processor	Output Parameter	Description
u_a, v_a T_a, T_{da} P_a	P1 (Upper Air Meteorological Processor)	z_a u_a, v_a T_{da}, T_{va}, T_{pa} q_a RH_a P_a $\Phi 850$, $\Phi 700$, $\Phi 500$ ρ_a σ_a	height aloft wind components aloft dew point, virtual, and potential temperature aloft mixing ratio aloft relative humidity aloft pressure level aloft height of 850, 700, 500 mb surfaces air density aloft static stability aloft parameters above interpolated regularly in vertical space and time

SECTION 3

THE NECRMP/NEROS DATA BASE

The NECRMP and related field studies (NEROS I, NEROS II, and PEPE) conducted during the summer periods of 1979 and 1980 provided a wealth of meteorological and air quality data over the Northeast U.S. These data include measurements made specifically for the field studies as well as routine monitoring of the atmosphere from existing networks. Measurements above the surface were obtained from instrumented aircraft, balloon-borne ascending and tethered instrument platforms, and satellite imagery. Surface measurements were made from fixed monitoring sites as well as special mobile vehicles.

Table 2 shows the instrumentation and parameters measured from the principal aircraft used in the NEROS experiments. During NEROS I (1979) the flight plans were designed to provide Lagrangian sampling of pollutants within specific air parcels over the regional (1000 km) domain. Measurements were made at several altitudes within the daytime boundary layer and aloft above the surface-based inversion at night. Aircraft flights during NEROS II and the NECRMP urban field studies (1980) concentrated on sampling in and near the urban plumes of Columbus, OH, Baltimore, MD, Washington, DC, New York, NY, and Boston, MA (Possiel and Freas, 1982) as well as Lagrangian regional scale sampling. Figure 6 shows a typical Lagrangian sampling flight track. Air parcel movement was monitored with tetroons. Over 200 separate flights were made during the 2 NEROS projects over the July-August periods of 1979 and 1980. Although it is not indicated in Table 2 hydrocarbon samples were also obtained from the aircraft. These were analyzed by gas chromatography for individual organic species concentrations. An attempt was made to obtain samples over rural as well as urban areas. The hydrocarbon samples obtained aloft are a supplement to those obtained from surface sites across the regional domain.

TABLE 2. NEROS AIRCRAFT INSTRUMENTATION

NEROS I Aircraft Operator	Element[1]	Equipment
Research Triangle Institute	O_3	Bendix 8002
	NO/NO_x	Monitor Labs 8440
	SO_2	Meloy 285
	CCN	Environment One Rich 100
	B-scat	MRI 1550 B
	AT	Rosemount 102
	DPT	EG & G 880
Brookhaven National Laboratory	O_3	AID 560
	NO/NO_x	TECO 14d
	SO_2	Meloy 165
	B-scat	MRI 1550
	AT	Yellow Springs 705
	DPT	EG & G 137-3C
	RH	Weather Measure HM 111
	TSR	Eppley Pyranometer 8-48 A
	UVR	Eppley UV Radiometer
Washington State University	O_3	Bendix 8002
	NO/NO_x	Monitor Labs 8440 HP
	SO_2	TECO 43
	CCN	Environment One Rich 100
	B-scat	MRI 1550
	AT	Metrodata M-8
	RH	Metrodata M-8

[1] O_3 Ozone
NO/NO_x Nitric Oxide/Oxides of Nitrogen
SO_2 Sulfur Dioxide
B-scat Light Scattering Coefficient
AT Air Temperature
DPT Dew Point Temperature
TSR Total Solar Radiation
UVR Ultraviolet Radiation
CCN Cloud Condensation Nuclei

[2] Information Not Available

TABLE 2. (continued)

NEROS II Aircraft Operator	Element[1]	Equipment
Environmental Monitoring Incorporated	O_3	Dasibi
	NO/NO_x	Monitor Labs 8440
	SO_2	Meloy 285
	SO_4	Meloy SA285
	B-scat	NA[2]
	AT	NA
	DPT	NA
AeroVironment	O_3	Bendix 8002
	NO/NO_x	Monitor Labs 8440
	SO_2	TECO 43
	SO_4	(not measured)
	B-scat	NA
	AT	NA
	DBT	NA
Stanford Research Institute	O_3	Dasibi 1003 AAS
	NO/NO_x	(not measured)
	SO_2	Meloy SA285E
	SO_4	(not measured)
	B-scat	NA
	AT	NA
	DBT	NA

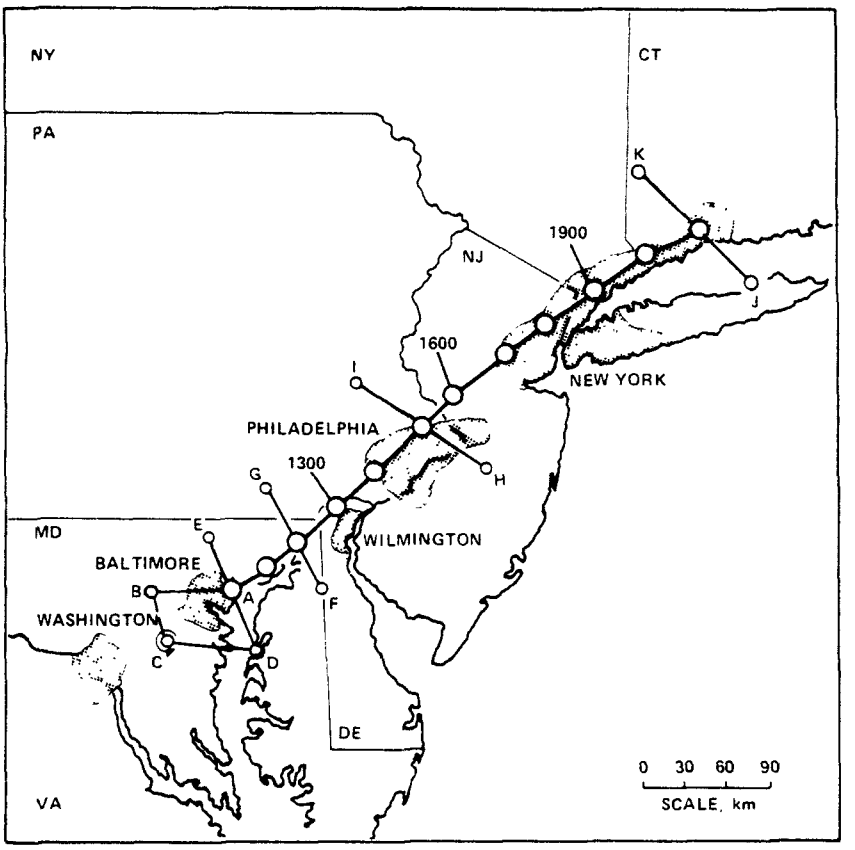


Figure 6. Typical Lagrangian sampling flight track along the East Coast corridor. Lettered sections indicate cross-trajectory traverses. Numbers are sampling times.

Meteorological data aloft were obtained from radiosonde releases at regular National Weather Service (NWS) upper air stations within and outside the boundaries of the ROM domain. Releases are normally made at 0000 GMT (1700 EST) and 1200 GMT (0700 EST) on a routine daily basis from these stations. Additional releases were made at 0600 GMT (0100 EST) and 1800 GMT (1300 EST) for 7 stations within the ROM domain during the NECRMP study. Figure 7 shows the locations of these upper air sounding stations that are used in the ROM data processing. Additional soundings are also available from the NECRMP urban field studies program (Possiel and Freas, 1982). Photographs of cloud cover over the model domain made by satellite observation are another form of non-surface based meteorological measurements used in the data processing for the model. These provide measures of the total cumulus cloud cover and the maximum depth of these cloud elements which are required model input parameters.

Surface based meteorological data were obtained primarily from the more than 100 NWS stations in the U.S. and southern Canada within and close to the boundaries of the ROM domain. Figure 8 shows the distribution of the surface stations providing data for one particular hour. Hourly observations of temperature, dew point, station pressure, wind speed and direction, and precipitation amount are available. Other observations, some of which are subjectively determined, include the amount of sky cover, cloud type and altitude, and prevailing visibility. Most of the meteorological stations report on a 24-hour basis although a few send reports only during daylight hours. The large number and high density of surface meteorological stations within the ROM domain provides more than adequate resolution for interpolating the required variables over the grid network.

Air quality measurements from surface-based monitoring locations were derived from three major data source categories. In addition to the routine measurements available from the SAROAD network described below, two special monitoring networks were established during the NECRMP study. The special networks included several monitoring sites for O_3, NO/NO_x, and hydrocarbons in the vicinity of the Columbus, OH and Baltimore, MD urban areas as part

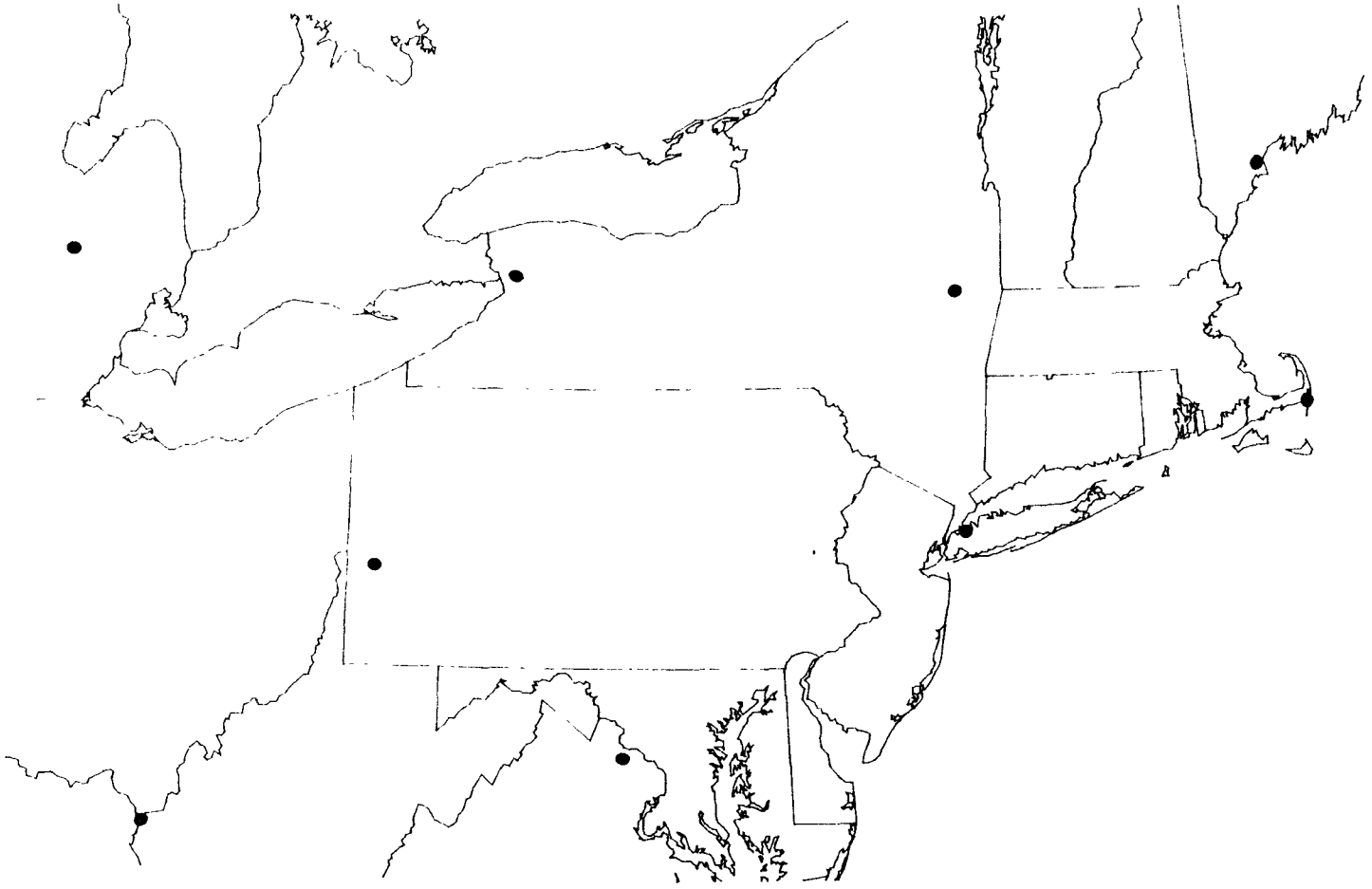


Figure 7. Locations of upper air sounding stations within the ROM domain.

Figure 8. Locations of surface meteorological monitoring stations within the ROM domain reporting at 1 particular hour.

of the 1980 NEROS II project to study urban plume transport. Also, during the 1980 NECRMP urban field studies several monitoring stations were established in the vicinity of major Northeast urban areas including Washington, DC, Baltimore, MD, New York, NY, and Boston, MA. These sites supplemented the existing networks with O_3, nitrogen oxides, and hydrocarbon measurements (Possiel and Freas, 1982). The quality assurance activities conducted for the surface monitoring programs of the NEROS and NECRMP studies were quite extensive. For all instruments these activities included frequent systems audits of monitoring sites, verification of calibration procedures, instrument calibrations and zero/span and precision checks, and instrument performance audits. The 1980 field study period will be the focus for obtaining data for the ROM evaluation because of the greater data availability, the coincident NECRMP urban field studies on the east coast, and the establishment of a base year emissions inventory for 1980.

The majority of available surface air quality monitoring stations are part of the established network of the Storage and Retrieval of Aerometric Data (SAROAD) system. This system assimilates air quality data from the networks of the National Air Monitoring Stations (NAMS) and the State and Local Air Monitoring Stations (SLAMS) and includes approximately 5,000 individual sites in the U.S. Hourly average O_3 values are available from the majority of these sites, while nitrogen oxides and hydrocarbon measurements are obtained at fewer locations than O_3. Quality assurance procedures are part of the SAROAD system. This minimizes the amount of erroneous data entering the system, although the number and frequency of audits on such a large network of regular monitoring stations are considerably fewer than those performed on the special NEROS and NECRMP sites specifically set up for those projects.

Figure 9 shows the distribution of surface O_3 monitoring sites included in the NECRMP field project and Figure 10 shows the regular sites for O_3 monitoring in the SAROAD network that are included within the ROM domain. A complete list of the SAROAD monitoring stations that measured O_3 appears in Appendix A.

Figure 9. Locations of monitoring sites for O_3 operating as part of the 1980 NECRMP Urban Field Studies.

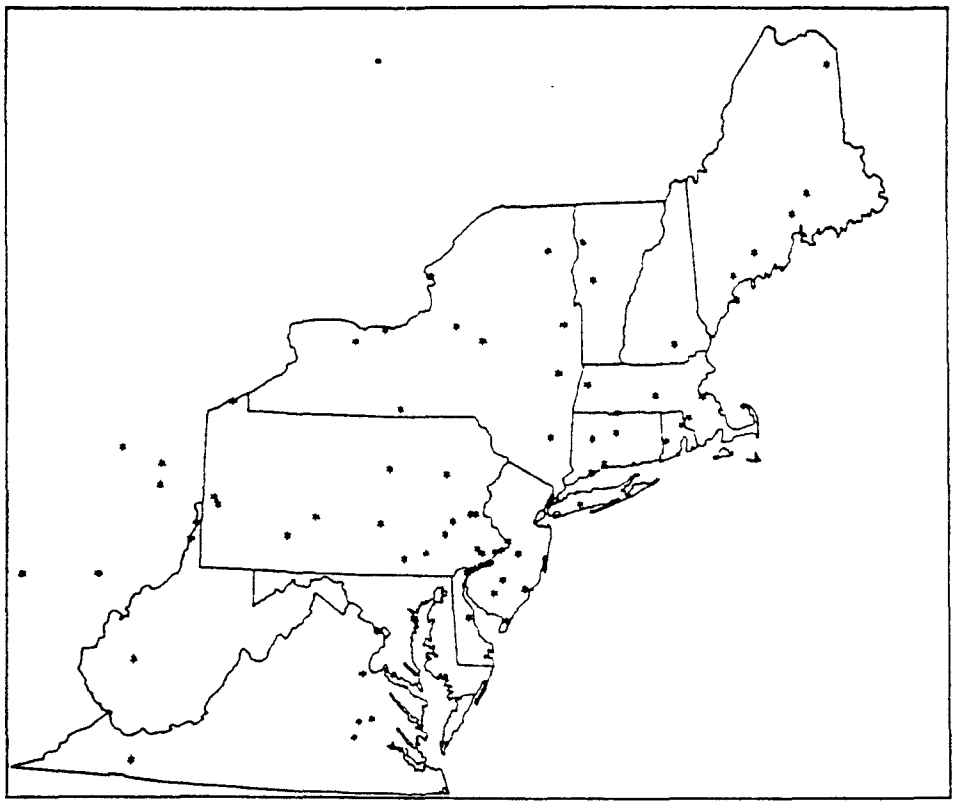


Figure 10. Locations of regular SAROAD monitoring sites for O_3 located in the northeast U.S. operating during 1980.

The final component of the data base is the source emissions inventory. Detailed data are required by the ROM to describe hour by hour emissions of gaseous species from anthropogenic and natural sources within the modeling domain. The primary species of interest are CO, NO_x, and speciated categories of natural and anthropogenic hydrocarbons. As part of the NECRMP project a complete emissions inventory for 1979-1980 for all point and area sources was prepared with the cooperation of the U.S. EPA and all of the states within the ROM domain (EPA, 1982). These data are being incorporated into the newer emissions inventory of the interagency National Acid Precipitation Assessment Program (NAPAP). The base year for these emissions data is 1980. The major data components in the inventory are U.S. and Canadian anthropogenic sources and U.S. and Canadian natural (biogenic) sources. The spatial extent of the NAPAP inventory is sufficient to provide complete coverage of the ROM grid. An EPA-developed processing routine transforms the annual emissions of the inventory into the hourly, speciated emissions used by the ROM emissions preprocessor. Temporal and spatial allocation factors for each source type and emitted species are used in the transformation (Sellars et al., 1985).

The availability of required input data as well as the availability of observed air quality data with which to compare model predictions will guide the selection of specific scenarios for modeling. Further, since the primary interest in the ROM is in the prediction of photochemically produced pollutants, the highest priority for modeling scenarios will be for regional smog episodes as determined from a distribution of observed hourly O_3 concentrations. Figure 11 shows a frequency distribution of surface O_3 concentrations over the SAROAD and NECRMP monitoring stations within the ROM domain during the 1980 NECRMP/NEROS study period. A number of high-O_3 episodes are evident within the period from July 23 through August 16. This period also lies within the intensive field efforts of the NECRMP and NEROS studies providing an enhanced data network of meteorological and air quality observations. Selected episodic periods of 2 days or longer within the above time "window" will be chosen for model simulation for the ROM

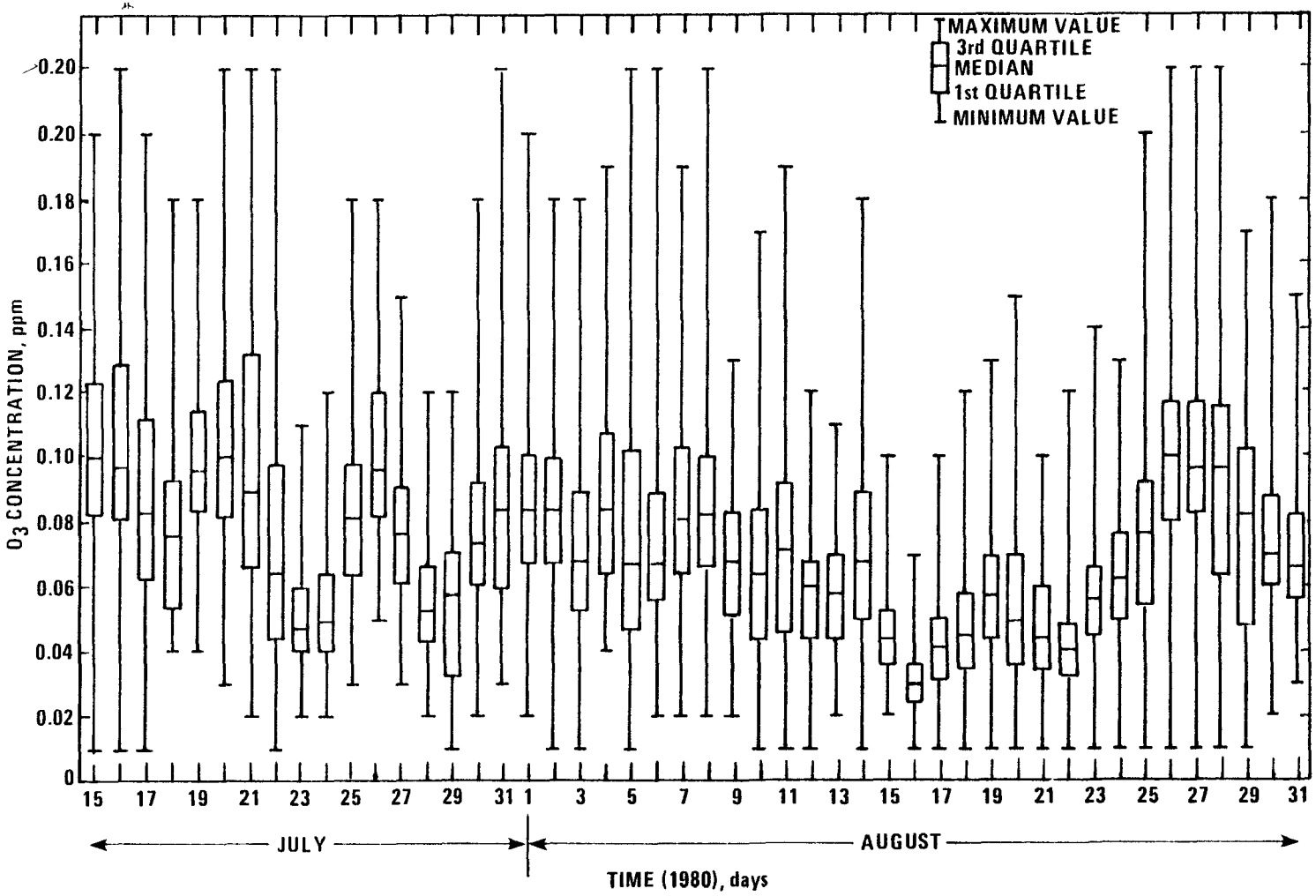


Figure 11. Box-plot frequency distribution of hourly O_3 concentrations from the SAROAD monitoring network during the 1980 NECRMP/NEROS study period.

evaluation study. Also, a longer-term period of approximately 2 weeks from within the window may also be simulated without reinitializing the model at any point after the start of simulation.

The emphasis in the model evaluation will be on O_3 predictions, although there are other pollutants also of interest. Some study may be given to the ROM's ability to predict CO concentrations within the domain. The nearly inert character of CO makes it a quasi-tracer species with which to evaluate the transport portion of the model without the interference of non-linear chemistry effects. Also, the primary importance of NO_2 in the photochemical cycle suggests that some evaluation statistics may be developed for this pollutant as well.

SECTION 4

MODEL EVALUATION

OVERVIEW

The comprehensive evaluation of a 3-D gridded air quality simulation model is a complex task. At a basic level it is necessary to test whether the mathematical representations of the individual physical and chemical processes are correct. This is done for each component process in the ROM. The validity of the chemical kinetic mechanism, for example, may be tested with data from controlled smog chamber experiments. The vertical cloud flux algorithm can be validated with aircraft measurements of material below, within, and above a convective-type cloud from the NEROS field program. The wind field algorithm cannot be precisely tested in an independent fashion, although individual trajectories can be evaluated from tetroon data from the NEROS study and the individual wind realizations can be analyzed for their consistency with measurements from fixed monitoring sites. Lamb (1983) describes some of these independent tests on component parts of the ROM during the developmental phase. The evaluation outlined here considers the model as a whole and attempts to compare its predictions with ambient observations.

This view would suggest that the ROM could then be viewed as a "black box" model producing results that could be statistically analyzed against observations. Although it is tempting to take this view, it must be avoided. The blind application of statistical tests will not provide sufficient understanding of model performance to draw any meaningful conclusions from the study. The goal of a model evaluation study, such as the one proposed here, is to gain insight into the model predictions and the observed data. One wants to know if the model is producing the right answer for the right reason and whether the model prediction is good enough for the user's purposes. Dennis and Irwin (1985) argue that model evaluation is as much art as it is science. They advocate that the careful

analysis of observational data will provide the direction for a meaningful path through the course of a model evaluation. The proper choice of statistical comparisons between observations and predictions can be established from an understanding of the phenomena shown by the data sets. This is the approach that we take in the ROM evaluation.

The approach will consist of 3 levels of analysis. The first level will explore the observed data set with diagnostic tests to bring out the important features, both on temporal and spatial scales, and to order the data according to the features that are found. Because the principal pollutant of interest included in the ROM simulations is O_3 the analysis of observed data will focus on it. The second level of analysis will be an evaluation of the ROM results for an individual realization of an episode. This approach is a deterministic analysis of each simulated realization without regard to the other members in the family of realizations. Finally, the third analysis level will consider concurrently the entire family of simulated realizations and the resulting probability distributions of concentration. This analysis structure is actually a hierarchy with the results obtained at each lower level guiding the steps taken at the next higher level.

DATA BASE ANALYSIS

A thorough analysis of the existing ambient data base is a key step in any model evaluation study. This analysis should elicit the major phenomenological features on the spatial and temporal scales of the model domain. In turn, the results obtained in this step will provide guidance in the model evaluation steps. The major thrust in the data base analysis will be on air quality data from surface monitoring sites. Pollutants of particular interest include O_3, CO, NO, NO_2, and hydrocarbon species. Air quality data obtained aloft will provide some idea of the model's predictive abilities in the upper mixed layer and cloud layer, but these measurements were obtained at irregular intervals and only at selected locations over the model domain. They cannot provide the regular hour-average type data

needed to conduct a complete evaluation. Moreover, the National Ambient Air Ouality Standards (NAAOS) are written for ambient surface level concentrations. The ROM evaluation will emphasize the model's ability to predict these concentrations.

There are many diagnostic tests that could be applied to elucidate the phenomena inherent in the ambient data. At a basic level we are interested in observing whether the air quality data, particularly O_3, define a plume downwind of major source areas. The ROM evaluation in an episodic mode is predicated upon the impacts of such plumes on downwind locations. A careful spatial analysis of the available data should reveal the extent of these plumes and whether they are broad enough to be resolved on the 18.5 km grid scales of the ROM. For receptors outside the influence of strong hydrocarbon sources a simple chemical check of the photostationary equilibrium between NO, NO_2, and O_3 can be made under daylight conditions. This test is particularly useful for flagging monitors (either NO_x or O_3) that may be malfunctioning for some reason. The expected interrelationship among these species outside of hydrocarbon influence provides an independent gauge on the adequacy of the monitored values. This test is of limited use however for low NO concentrations when instrument sensitivity becomes a factor. Another diagnostic procedure is a check on the NMHC/NO_x ratio in the air quality data. For a given time of day, this ratio can be studied over all monitoring stations. Values in the vicinity of urban source areas can be indicators of the potential smog reactivity of the pollutant mixture.

A seemingly recent data technique of spatial (and temporal) analysis, called Kriging (after the mining engineer D. H. Krige), should prove to be especially useful in the analysis of the network data from fixed monitors and in transforming that data into a spatially smoothed form most relevant for comparison with ROM predictions. Kriging is based on an application of the statistical properties of random functions. It has come from geostatistics (Delhomme, 1978; David, 1977), but for spatial analysis it is the same as structure function analysis (Gandin, 1970) and also has application to time-series analysis.

Kriging was developed as a method for optimal interpolation, accounting not only for the geometry of data values (distance), but also for distance-dependent structure in the data related to spatial covariability (spatial correlation). It is a method for interpolating fields of data that minimizes the error variance of the interpolated point. Thus Kriging provides a means to estimate the interpolation error of points and/or regions in a data field, under certain assumptions of random behavior of the data. In developing such estimates, Kriging provides a means to estimate inherent sampling error together with variability due to subregional effects that are evident from the data. Thus, Kriging provides a means to estimate a more regionally representative pollutant field for comparison with grid-averaged pollutant predictions. It also provides a means of estimating the extent of influence of spatial covariation and the degree to which such a range of influence is anisotropic. Hence, Kriging provides an estimate of the accuracy or uncertainty of both the monitoring network values and the interpolated values.

Kriging should be a useful tool in the comparison of model output data sets with a validation data set that contains inherent sampling (and measurement) error. As an interpolation method, it can be used to develop point or spatial estimates from network values, together with an estimate of the associated uncertainty, corresponding to ROM grids. Using these uncertainty estimates and the theory of standard normal distributions, the statistical significance of differences between individual ROM realizations and interpolated validation data can be assessed. Areas of significant model error can be identified and contour maps of regions of significant error can be produced to investigate the spatial characteristics of model errors. The uncertainty estimates can also be used in more complex characterizations of spatial and temporal model errors.

Through the use of Kriging, together with co-located monitoring data, it may also be possible to establish the expected variance in the air quality measurements at monitoring stations due to measurement error. The variance estimated for a particular monitoring site for a given time period,

such as 1 hour, is a combination of instrument variability and small scale chemical effects plus other subregional effects. ROM layer 0 O_3 predictions include an estimate of these subgrid scale chemical effects. Thus some comparisons should be possible between the expected variance as determined from selective Kriging of the network data and that from the ROM predictions.

In addition to analyzing the air quality data, the emissions and meteorological data bases can be used in diagnostic testing. The source emissions distribution across the ROM domain in the Northeast U.S. application is not at all homogeneous. Major urban areas emit large amounts of nitrogen oxides and hydrocarbons from a distribution of sources. Large single point sources, such as fossil-fuel electric power plants, are also scattered throughout the ROM domain in urban and nonurban areas. Emissions from these point and area sources are transported downwind according to the existing flow regime. Determining the trajectories along which these emissions travel can be an interesting and useful exercise. These trajectories should provide guidance on the location of actual and simulated plumes from these sources. A subset of major point and urban area sources can be chosen from all sources within the ROM domain. For each gridded wind flow realization during an episode a forward trajectory is tracked for the duration of the episode. This can be done for each realization in the family of possible realizations. Comparisons can be made among the computed trajectories for a given source area. If the trajectories show good agreement in location among all realizations the model should provide a clearly defined predicted plume centerline in the air quality simulation with which to compare against the observed plume data. Trajectories showing spatial disparities among the realizations may indicate a less precise predicted plume location. In this case some trajectories may intercept other major downwind source areas and other trajectories may not, leading to possible large differences in the computed air quality values in the predicted plume. In any case, the envelope of trajectories defined by the family of realizations should provide the boundaries of the territory where the actual source plume is located. When these boundaries are very

far apart, the utility of the air quality predictions for short temporal periods may not be great, although the result shows the spatial area that is "at risk" over a longer period of time for the same set of synoptic conditions. This type of information is important for the model evaluation because it quantifies some of the limits of predictability for the ROM as used on a specific data base.

EVALUATION BY INDIVIDUAL REALIZATION

The simulation of any given episode with the ROM will involve a finite number (~ 10) of realizations based on possible flow fields. We can regard any one of these realizations as determining the "real" outcome, while excluding all the other realizations. This is a deterministic perspective where we search for the best predictive capability of the model by evaluating each realization independently. In this case ROM predictions can be directly compared with monitoring station observations. Because of the strong influence of initial conditions on the modeled results, the first 6 to 12 hours of simulation should probably be excluded from the evaluation. This will enable the ROM to begin producing concentration fields based on the source emissions input at the point the evaluation begins. Since model initialization occurs at 1200h (noon) the first results subject to evaluation should be those of 0000h the next day, with the evaluation proceeding in increments of 24 single hour periods.

The focus of this evaluation will be on the accuracy of the ROM predictions of hourly O_3 concentrations, including the daily spatial and temporal patterns of the predictions, and the accuracy of the magnitude of the peak prediction for each day. The first point is important in interpreting the results, establishing confidence intervals, and diagnosing errors, while the analysis of O_3 peaks is at the heart of the regulatory use of the ROM.

Dennis et al. (1983) give a thorough discussion of various measures of model performance. The choice of particular measures depends upon the attributes of performance deemed important for a particular model and data base. For the ROM evaluation by individual realization, the performance measures appropriate for determining the accuracy of the hourly predictions are the bias, gross error, noise, and variance, complemented by graphical analysis of temporal and spatial behavior. For evaluating the peak predictions, all of these measures as well as a linear regression of predicted and observed values are appropriate. Estimated model bias is the mean of the residual or difference (d) between observed and predicted values, or,

$$\overline{d} = \overline{(C_o - C_p)} = \overline{C}_o - \overline{C}_p.$$

If the bias is significantly different from zero, it indicates a systematic tendency in the model's predictions with respect to timing, spatial location, or magnitude of the peaks. Gross error, a measure of model inaccuracy, can be estimated by the average absolute deviation in the residual concentrations,

$$\overline{|d|} = \overline{(|C_o - C_p|)}.$$

The estimated noise is the variance in the residuals, or,

$$s_d^2 = \frac{\Sigma(d - \overline{d})^2}{n - 1}$$

Both the noise and bias are components of the gross error in the model. The noise level is generally high for model results that are strictly paired with observations by time and location because a prediction that is off by a few

hours or grid cells can greatly affect the noise level. The noise may be a more meaningful performance measure when it is based on less strict pairings, such as the observed and predicted daily maxima.

Graphical analysis is also a key element in model evaluation. Informative plots may be constructed from the following elements:

(1) Frequency histograms of C_o, C_p, and d,
(2) Time series plots of C_o and C_p,
(3) Scatterplots of d vs. time or C_o,
(4) Boxplots of d within natural categories, such as monitoring site, hour or day of observation,
(5) Spatial contours of C_o, C_p, and d fields at time of maximum C_o, and
(6) Plots of wind flow trajectories.

For each ROM realization we can apply the relevant performance measures to hourly values of O_3 for all receptor locations in the domain, strictly pairing the observed value with the predicted. This should be done both for predictions in layer 0 (shallow surface layer) and layer 1 (lower portion of mixed layer) in order to determine the effect of simulating the near-surface physical and chemical processes on the model's predictive performance. Model predictions of CO might also be evaluated in this manner to test the performance on a nearly inert pollutant species, although regionally representative CO monitoring data are not commonly available. Also, for all receptor locations the estimated variance in the hourly O_3 observations from the Kriging data analysis can be compared with the layer 0 predicted variance in the ROM O_3 values for the same time and location. Receptor sites for the comparison of variances might be stratified according to urban and non-urban locations to separate the microscale chemical effects expected only in the urban areas from the total variance. Thus comparisons can be made for the hourly average values of O_3 and its estimated variance between the predictions and observations.

The step outlined above in the ROM evaluation should provide estimates of the model bias for each realization as well as possible signs of any

systematic errors in the model results. The analysis of these results can be done for each individual day of the simulated episodes (up to 24 strictly matched pairs of C_o and C_p at each receptor location) as well as cumulatively for all days in an episode, and for all episodes. Furthermore it is sometimes useful to form subgroups of the hourly values on which statistical evaluation can be performed. For example, model performance can be judged for O_3 above a certain threshold level, such as 0.12 ppm which is the current NAAQS value. Another subgrouping might stratify the hourly values as to time of day such as 06 - 10h, 10 - 14h, and 14 - 18h. There may be other natural dividing lines for subgroups as well. The evaluation by subgroup can often detect performance characteristics unique to that subgroup which might be masked or averaged out in the full set.

The next step in the evaluation procedure is the analysis of the peak or maximum values; a step of great interest to the regulatory community since the peak values are those that dictate emissions control steps. We concentrate on peak hourly O_3 values in this discussion. There are 2 sublevels of analysis for peak values: local and global. At the local perspective we choose the daily maximum observed hourly value at each receptor site. This can be paired with the prediction at the same site and time as well as the daily maximum hourly prediction at the site for any time. Both pairings are useful comparisons. If there is a systematic bias of the predicted peak value in time of day it should appear in the comparison of results from these local maxima analysis methods.

From a global perspective the observed daily maximum hourly concentration over all sites within a specified area of the ROM domain is determined and is paired with the corresponding predicted value in one of three ways: (1) with the predicted peak at the same site as the observed, (2) with the predicted peak among all sites within the specified area, and (3) with the predicted peak from any grid cell within the specified area of the domain. As these pairings become less restrictive (from 1 to 3) we should expect a tendency

away from underprediction in the model performance. With all other factors being equal the third pairing should lead to overprediction because the observations are constrained to the monitoring sites only, a smaller sample of points than all the grid cells in the area. However, the simulation of high O_3 episodes sometimes leads to a tendency toward underprediction of O_3 in all 3 pairings because studies of model performance are often done for days where high O_3 concentrations have been observed and days with lower observed concentrations have been excluded (Dennis et al., 1983).

The specified areas of the ROM domain to be used in the global analysis of O_3 maxima should be chosen based on some common characteristics that tie the area together. Examples might include political jurisdictions such as individual states, airsheds for air quality regulation based on population, source emissions, terrain distribution, or large downwind sectors from major source areas.

Another aspect of the ROM evaluation by individual realization that might be more difficult to quantify is the comparison of the spatial patterns in the predicted and observed concentration fields. Graphical contouring of these fields as well as the residual concentration field can subjectively provide an estimate of the model performance in this area. An objective graphical technique described by Siegel (1982) compares the shapes of two data sets and demonstrates the motions necessary to overlay one on the other with the best possible fit without altering the intrinsic shape of either field. There is a potential utility for this type of analysis in the ROM evaluation. Each individual simulated realization is a possible analog of the actual concentration field. If coherent plumes of higher concentration are seen emanating from large source areas in both the simulated and predicted fields, we can apply the objective graphical technique to compare these plume shapes. The predicted plume can then be translated and/or rotated as necessary to provide the best fit to the observed plume. This will, in effect, normalize the predicted field for errors caused solely by the specified wind field. After this plume rotation has been performed on all possible

areas within the model domain the statistical analysis of predicted and observed O_3 concentrations at receptor sites is again made to see how the plume reorientations improved the effective model performance. This step will be most useful in comparing the fields surrounding the maximum hourly observed O_3 concentration to that of the predicted O_3 within the same plume area. This comparison can be made for each realization for each day of simulation.

Figure 12 shows a structural diagram of the possible steps in the ROM evaluation by individual realization. The structure diagram is a hierarchy of operations including sequences (shown by horizontally colinear boxes), iterations (denoted by an asterisk, *), and selections (denoted by an open circle, O). The exact path we choose to take during the evaluation will be guided by the results obtained at intermediate steps along the way (feedback mechanism) as well as the available resources at hand. The diagram encompasses the complete set of evaluation exercises of which all or a subset of those shown will actually be performed.

EVALUATION OVER ALL REALIZATIONS

The distinguishing aspect of the ROM setting it apart from most existing air quality models is its ability to determine an expected range within which the actual measured concentration should fall. This ability is achieved through the assimilation of prediction data from all ROM realizations for a given episode. From this perspective all the realizations must be considered concurrently. At any given receptor location the ROM results generate a frequency, or probability distribution of expected concentrations.

This section explores potential avenues of evaluating this probabilistic type of model prediction. It is clear that in this vein there is no longer a one to one correspondence between prediction and observation. At any receptor location and time a single observation value (and possible variance, based on

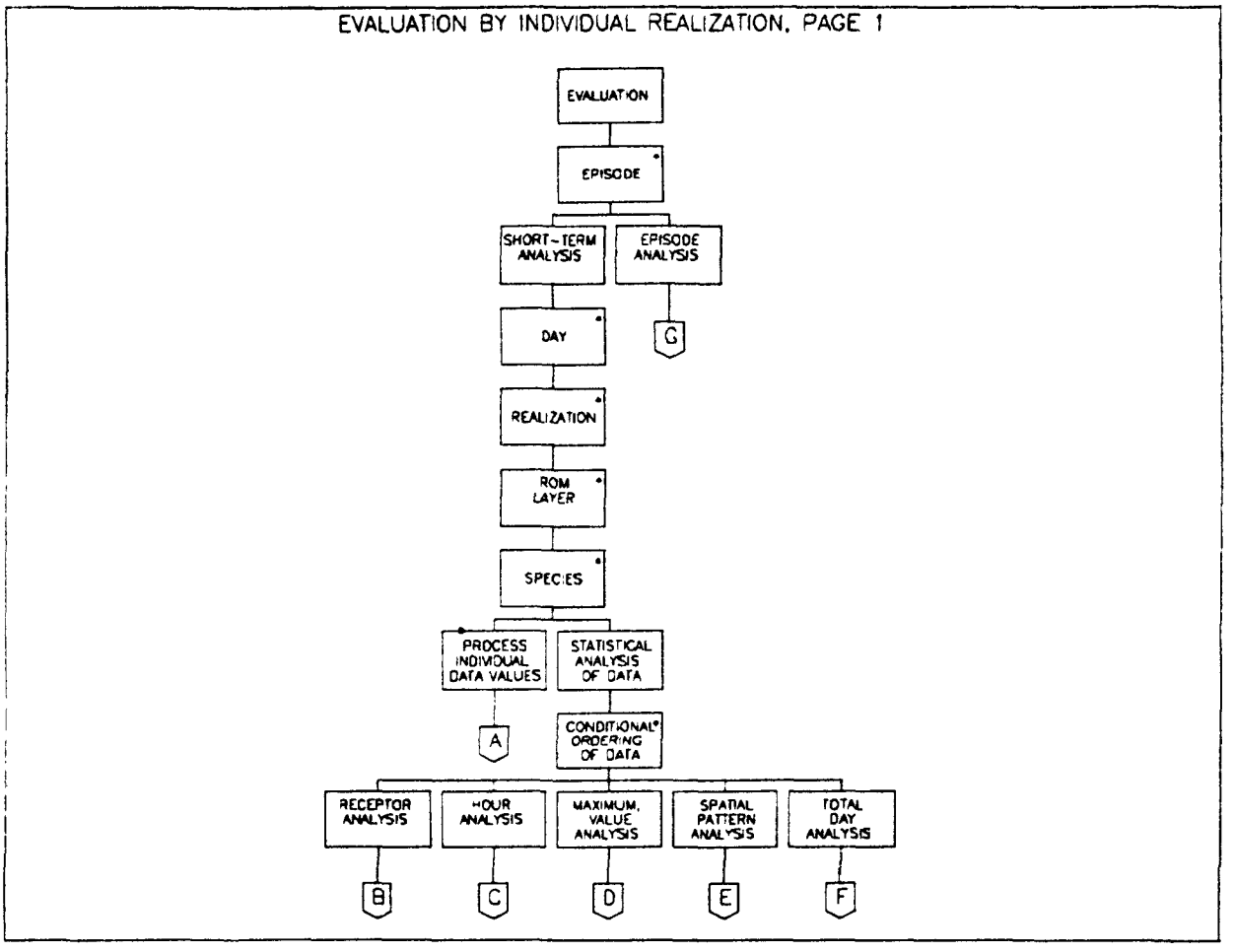


Figure 12. Structure diagram of the ROM evaluation by individual realization.

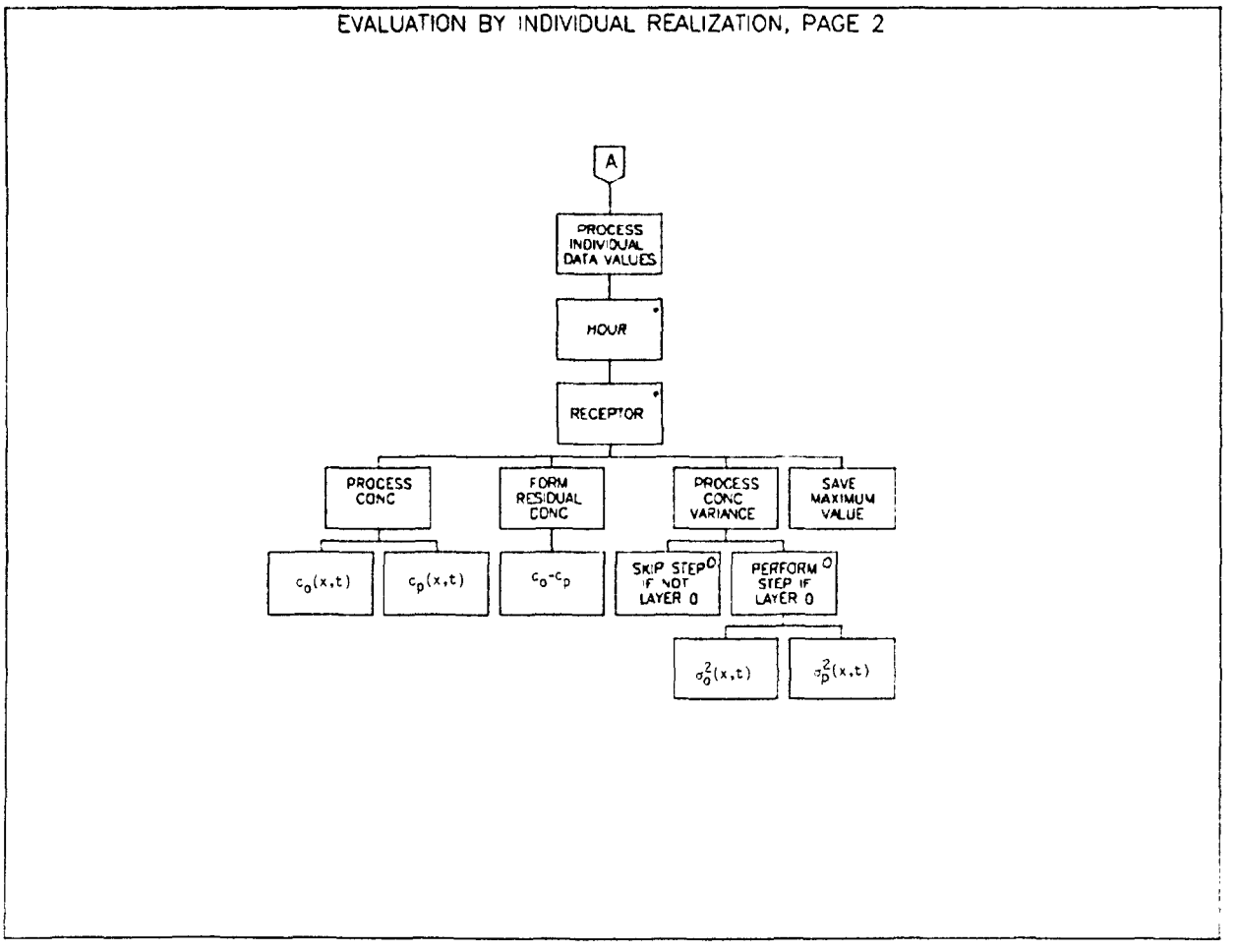


Figure 12. (continued)

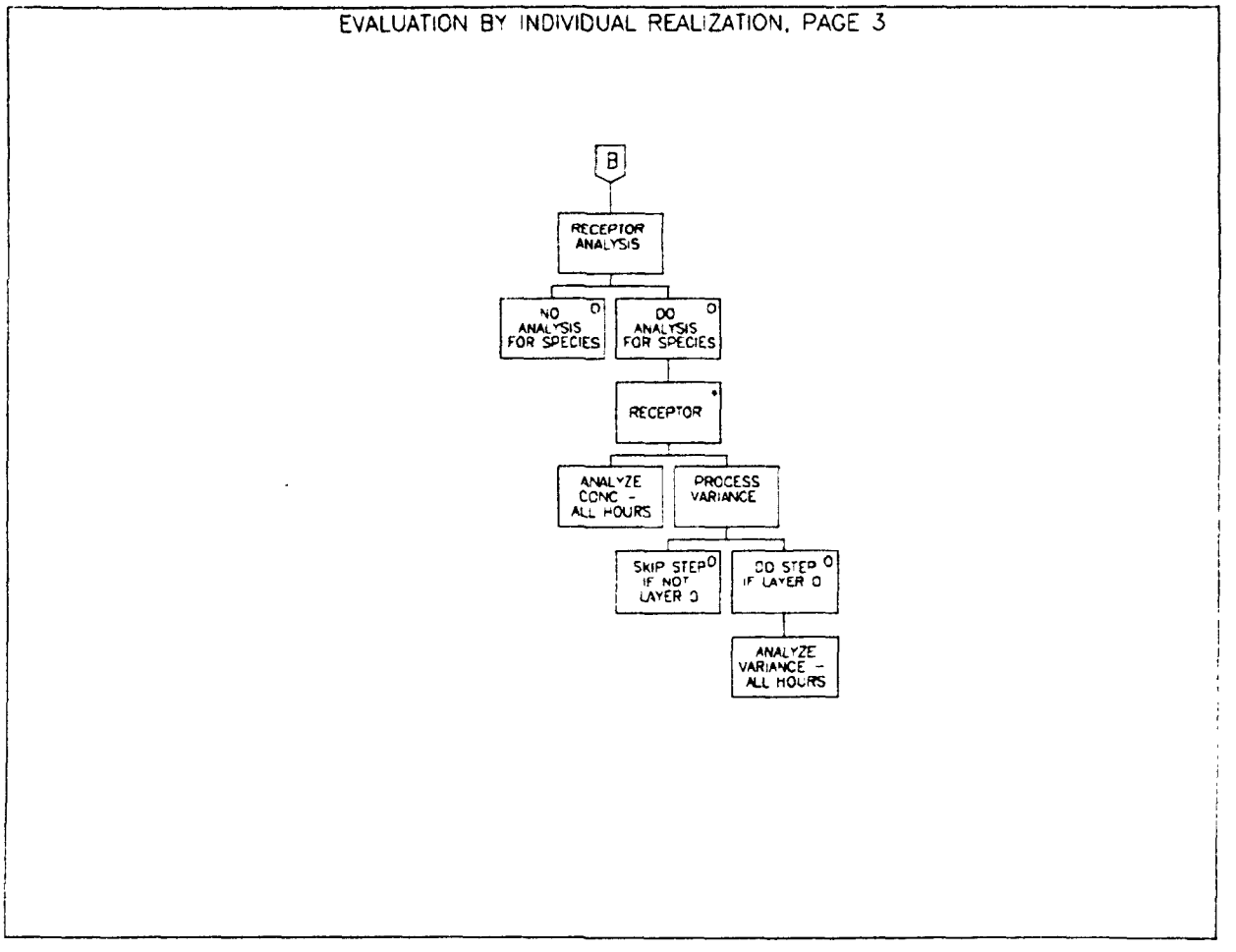


Figure 12. (continued)

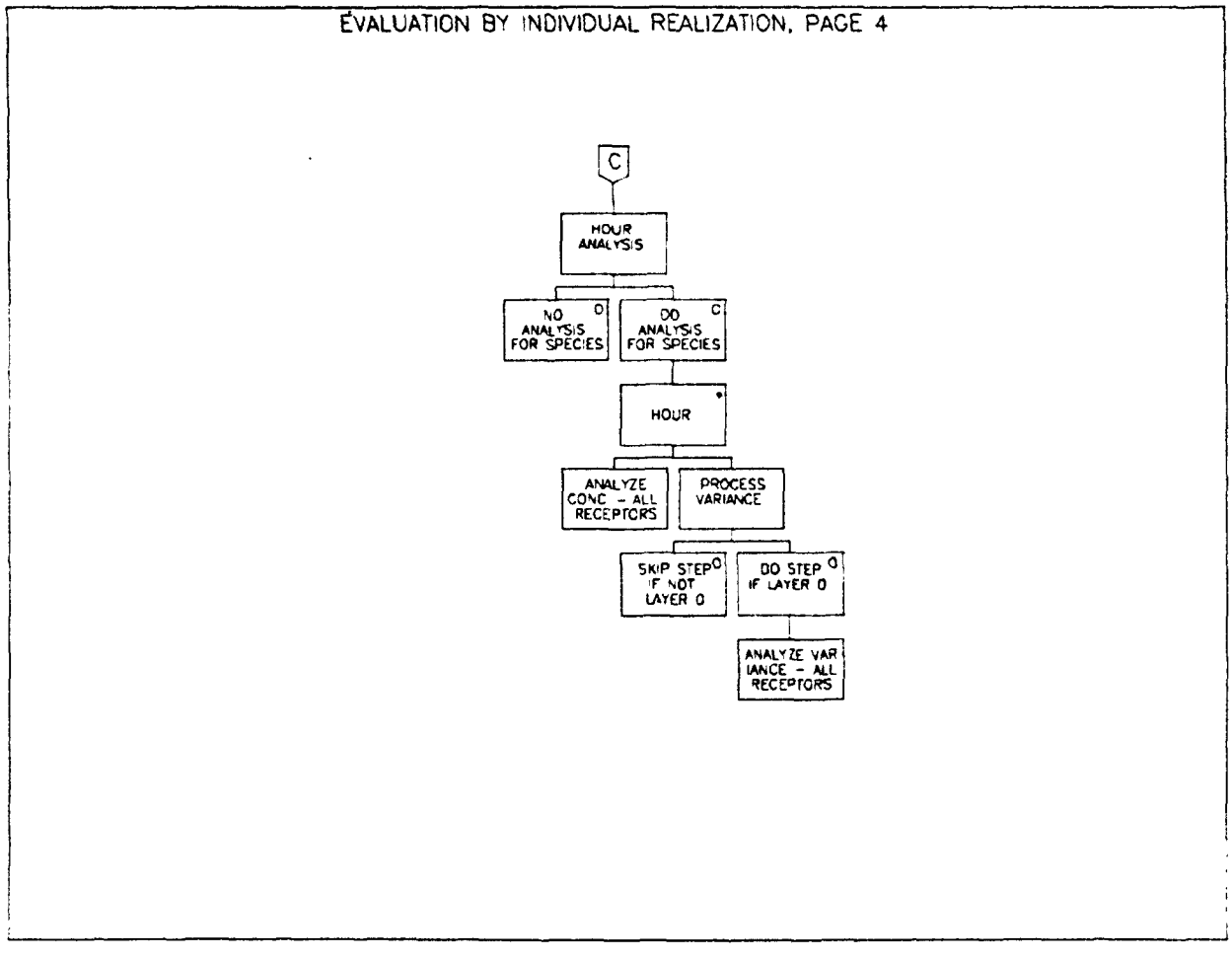


Figure 12. (continued)

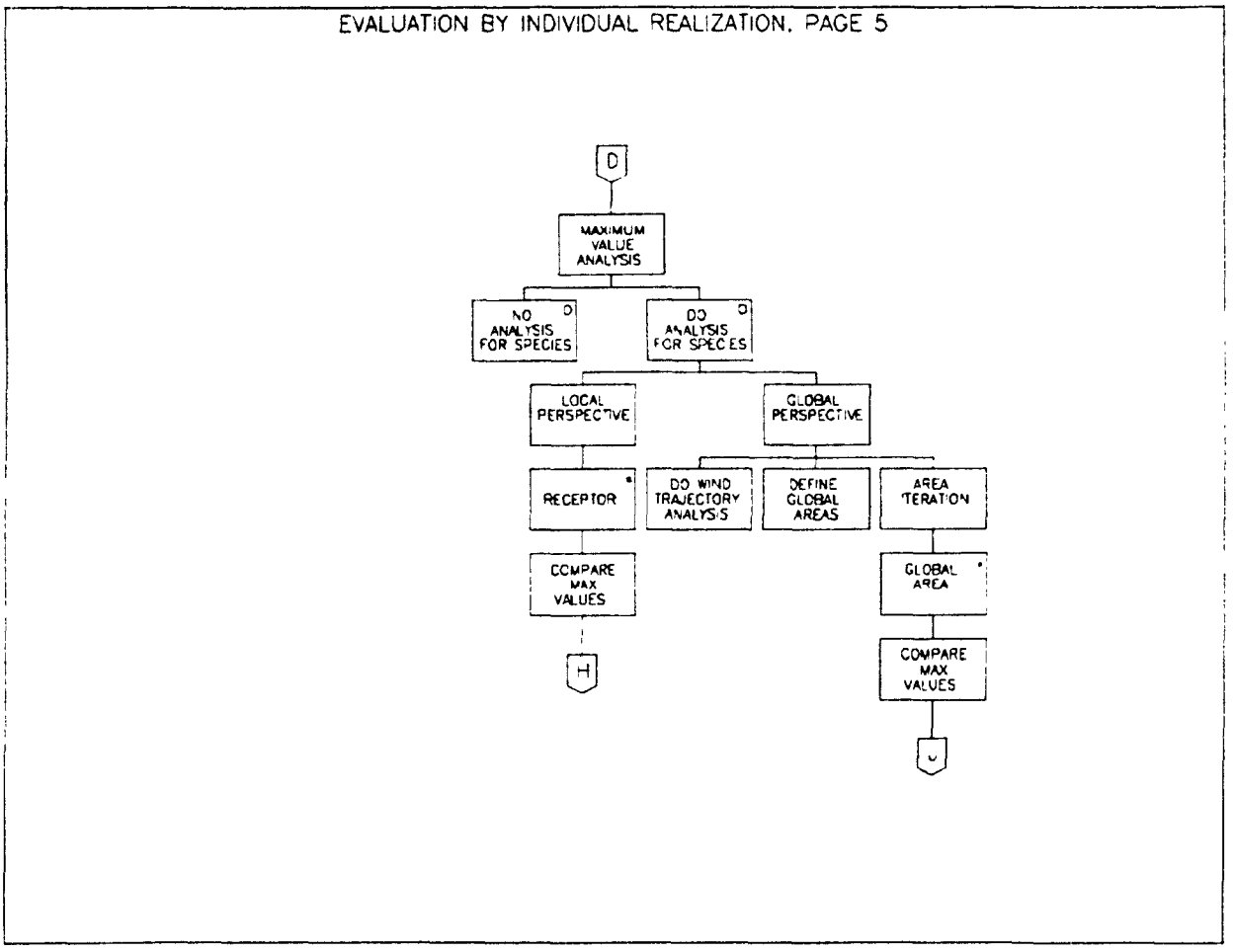


Figure 12. (continued)

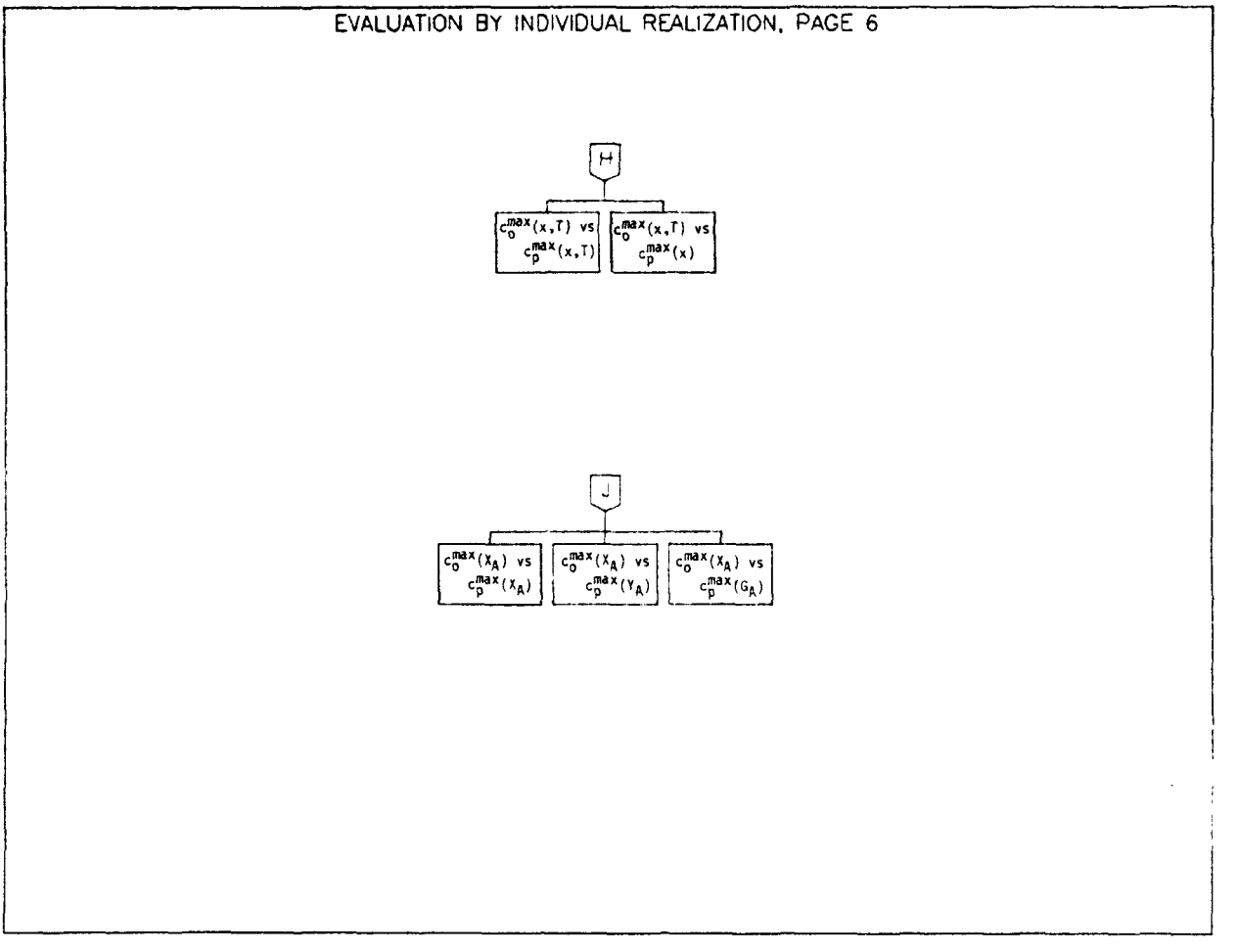


Figure 12. (continued)

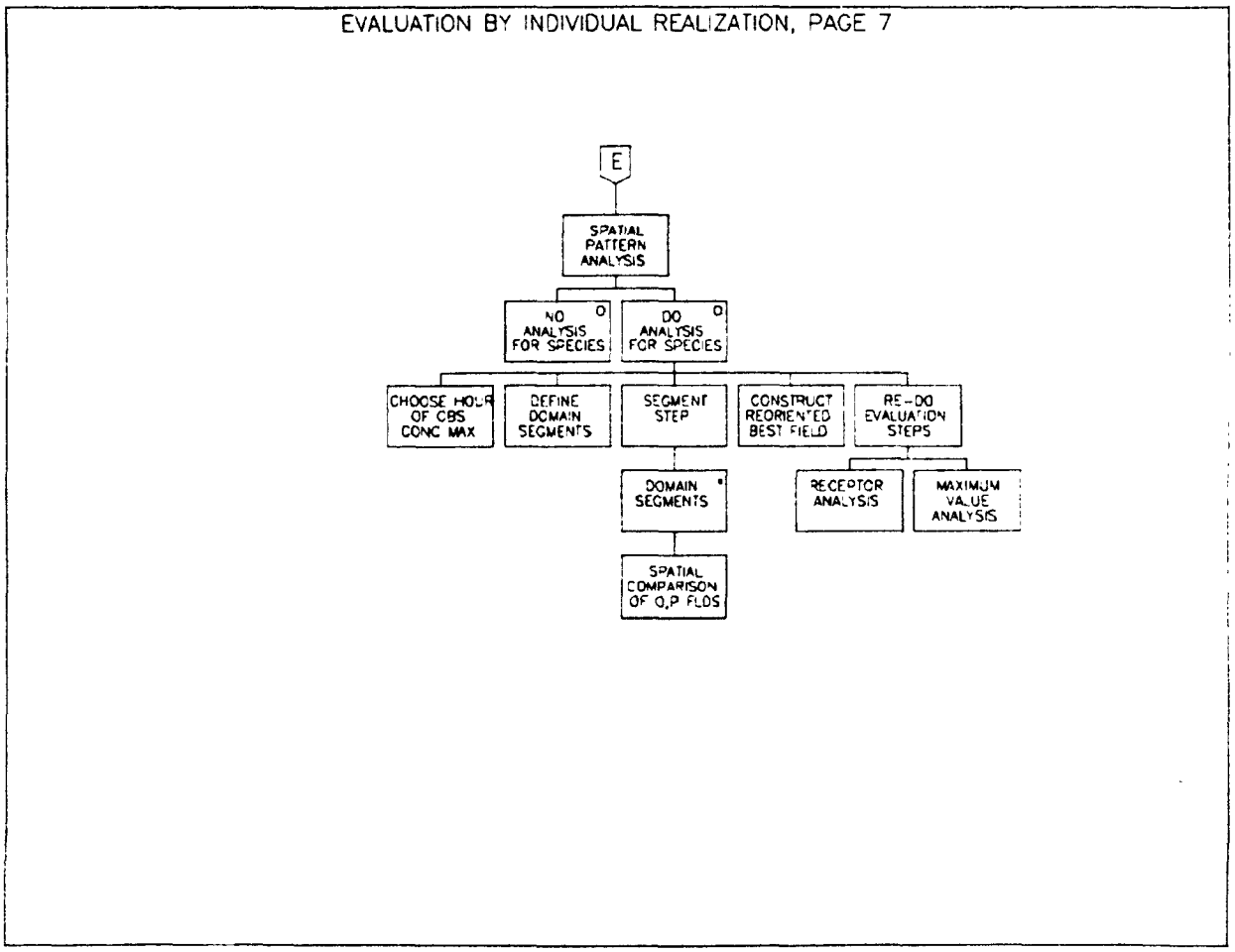


Figure 12. (continued)

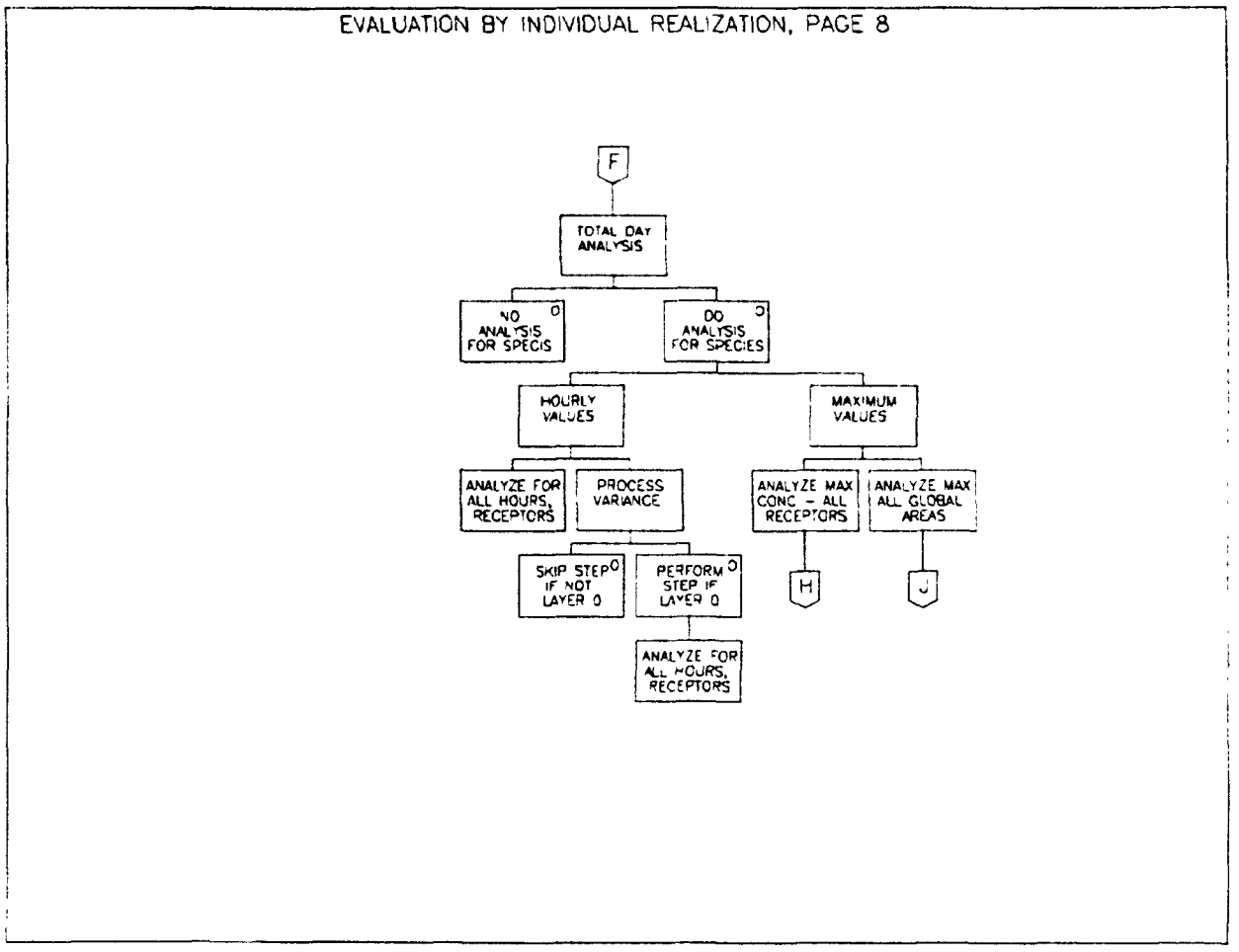


Figure 12. (continued)

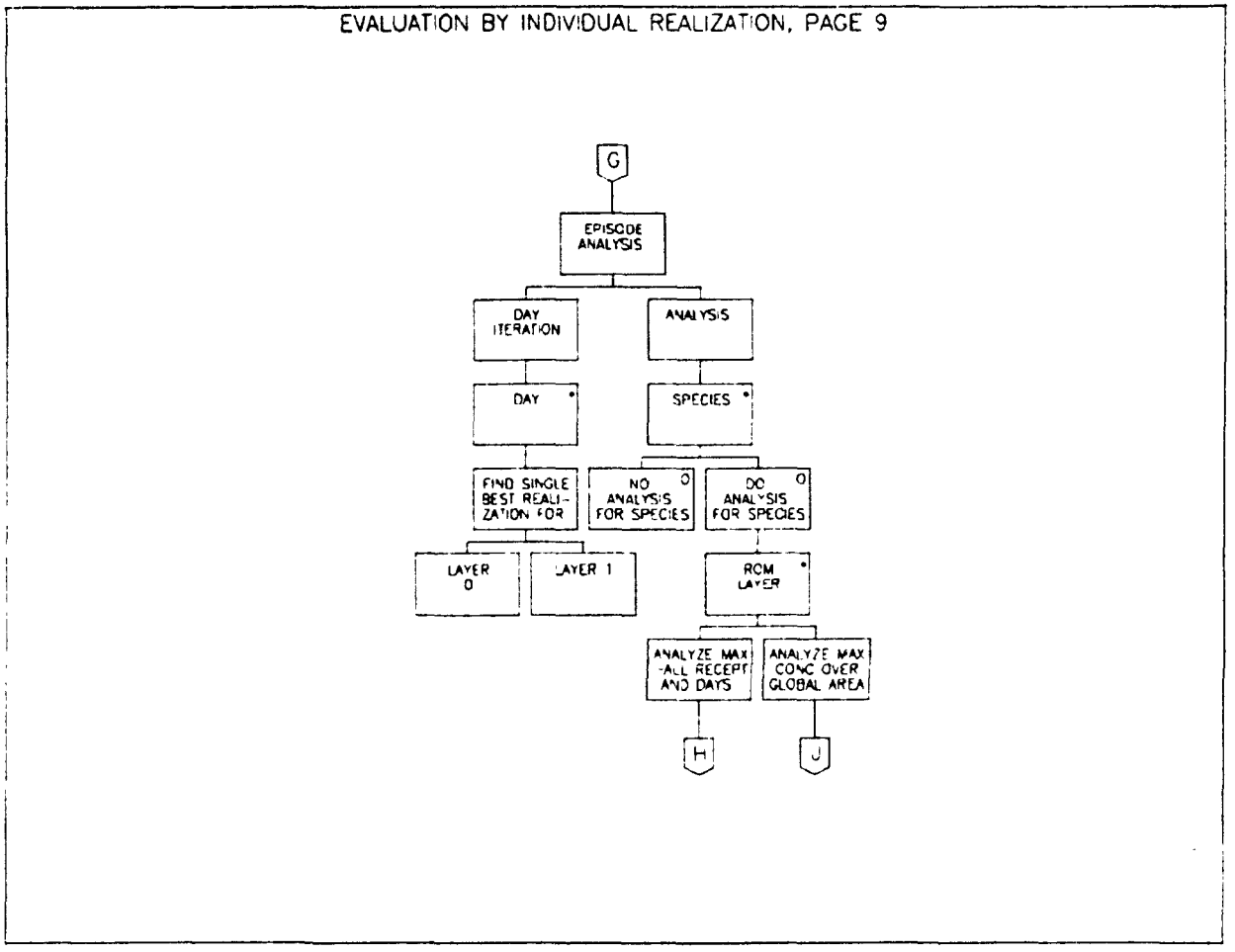


Figure 12. (concluded)

a Kriging analysis) is available. At the coarsest level of evaluation we can examine the resultant concentration distribution for a particular receptor and time, and judge its utility by its shape. For example, Figure 13 shows two hypothetical predicted concentration distributions at a receptor for a 1-hour period. The predicted concentration values for this receptor from the 10 realizations, along with their assigned probabilities of occurrence based on the wind field realization probabilities, define a concentration distribution. The distribution shown on the top is rather broad and flat, indicating that concentration values from C_1 to C_2 are nearly equally likely to occur. If, for example, the species is O_3 and the range of C_1 to C_2 is from 0.05 to 0.20 ppm, the model result is not very useful in an absolute sense. However, if the distribution changes uniformly with changes in emissions, then the prediction may prove useful to the regulator in a relative sense. One contributing factor to the possible prediction of a broad distribution such as this one is the location of the surface and upper air wind monitoring stations in relation to the receptor location. For example, consider the case of a receptor located in the same grid cell as a wind monitoring site. The method of wind field interpolation for the ROM forces the calculated fields to conform to the measured values at the monitoring site locations. Therefore the closer a receptor is to a wind monitoring site, the smaller the differences will be among the various wind estimates for the grid cell containing the receptor in the multiple wind realizations, and hence the smaller the differences will be among the predicted concentrations from the realizations. The prediction of a broad distribution, such as the exaggerated one at the top of Figure 13, is possible for receptor sites that are well removed from wind monitoring stations where the constraints of conformity to the observations are not so powerful.

The distribution shown at the bottom of Figure 13 shows a narrower peaked distribution. The expected range here should be more useful in an absolute sense to the regulator and to the evaluation procedure. At a minimum we can determine whether the observed concentration falls within the predicted distribution and whether it is within 1 standard deviation (σ) of the mean.

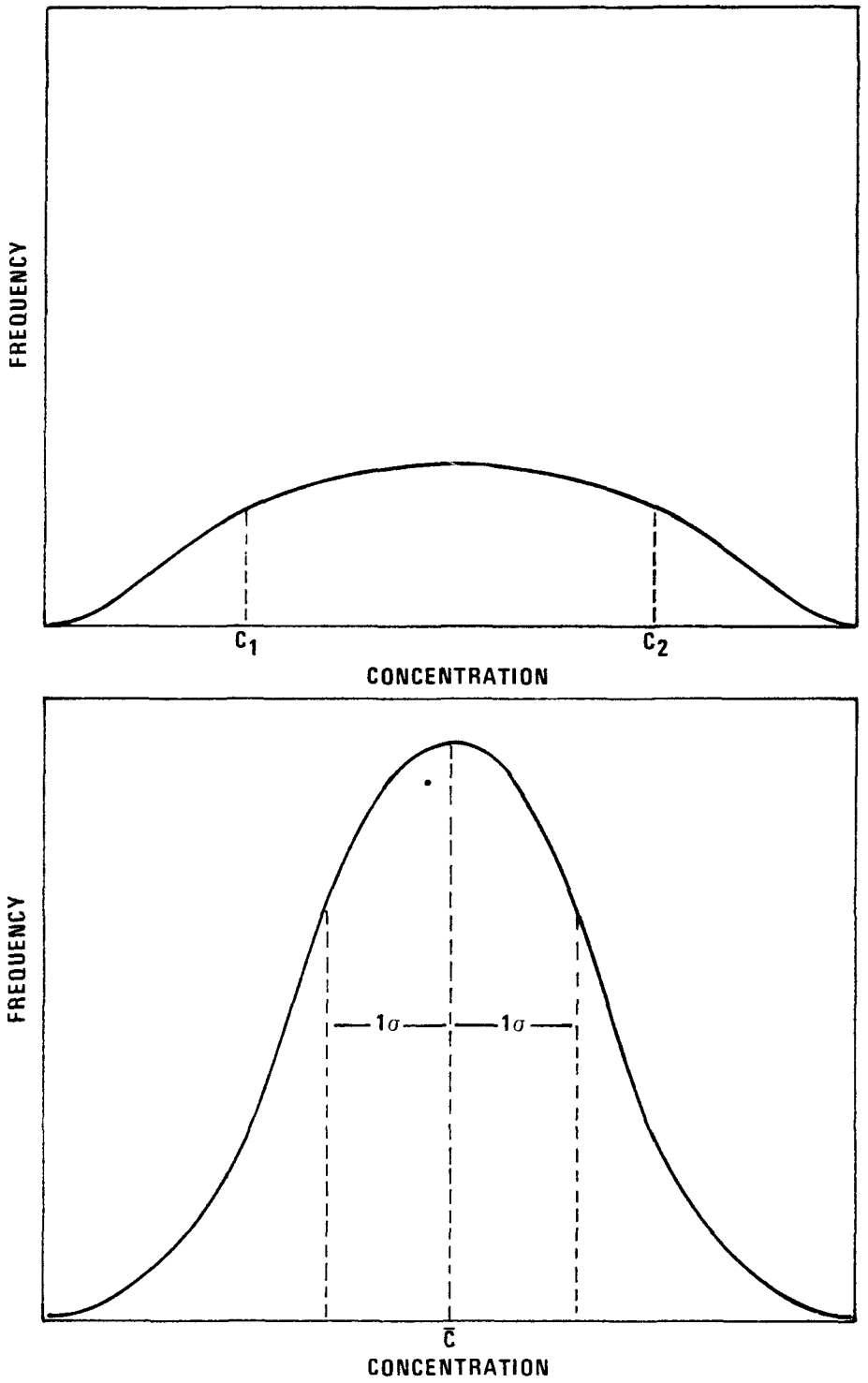


Figure 13. Hypothetical predicted concentration distributions from the ROM.

If the observation falls outside of or in the tail of the distribution, the model prediction is clearly in error. The concentration variance calculated from a Kriging analysis of the observed concentrations is not exactly comparable to the variance of the predicted distribution because the former is based on instrument variability and subgrid scale variability while the latter is based on expected variability from uncertainty in the flow field. New techniques, such as combined temporal and spatial Kriging, might produce more comparable estimates of the distribution.

From an evaluation standpoint we need a distribution of observed concentrations in order to perform objective statistical comparisons. There must be an orderly grouping of receptor stations to fulfill this requirement. It is not immediately obvious how such a grouping is achieved. One criterion for the grouping is that there should be no significant difference in the predicted concentration distribution among the receptor members of the group. We need a single predicted distribution with which to compare the observed distribution. Therefore the receptors in the group must have essentially the same prediction. This criterion can be used to decide if we have chosen a proper grouping of receptors. There is also no assurance that a proper grouping can be found from the receptor network set we will be working with.

Figure 14 shows an example from an early run of the 1st generation ROM of a box plot of the predicted O_3 concentration distribution from 69 layer 1 grid cells containing mostly urban land-use types. This plot was made from a single realization. If the shape and magnitude of this distribution remained constant over all realizations the grouping would be valid. This plot is shown by way of example only, for it is now thought that a grouping based on land-use type alone will not highlight the physical phenomena occurring in the regional scale model.

Intuitively one can imagine plumes of O_3 streaming downwind from major source areas in the model domain. The paths of these plumes will intersect receptor locations. One can hypothesize a grouping of receptors based on

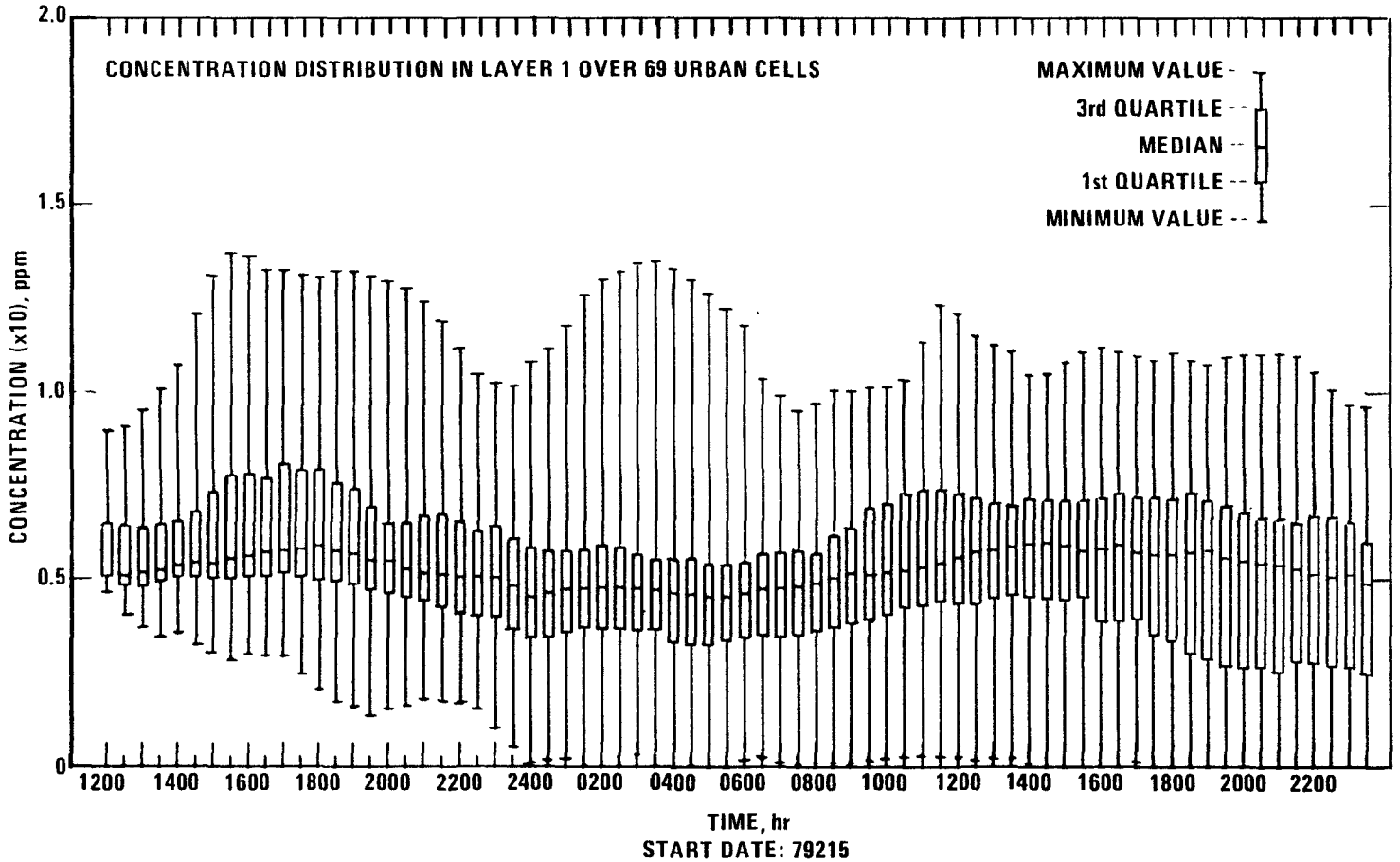


Figure 14. Box-plot of predicted O_3 concentrations from 69 urban grid cells in ROM layer 1 for August 3, 1979.

considerations of plume transport such that the members of the group lie within the downwind extent of such plumes as close as possible to the centerline and at comparable distances downwind of the respective source areas from which the plumes originate. The receptors should be chosen within plumes that have maintained their integrity and have not yet comingled with other major plumes. Figure 15 shows a conceptualized idea of how these receptors might lie in relation to several major urban source areas and their plumes. Any given plume should meander about a downwind receptor over the set of realizations. However the predicted concentration distribution formed from all realizations will have this meandering effect built in and should be common to all receptors in the grouping. This grouping can then be tested against the criterion described above.

Other considerations are necessary for this grouping also. For example, this discussion assumes that the major source emission areas have similar characteristics. In reality this is not true, so the predicted O_3 concentrations must be normalized in some manner by the source strengths of NO_x and hydrocarbons and the ratio between them (THC/NO_x). Also, the time period of interest must be specified for a receptor. This might correspond to the hour during which the maximum O_3 concentration was observed at a receptor during the day, or else an hour by hour O_3 comparison could be attempted between predicted and observed distributions.

If it is then determined that the predicted concentration distribution formed from all realizations is invariant over the set of receptors in the grouping, this distribution can then be compared to the distribution formed from the observed concentrations at the receptors. Should the predicted and observed concentration distributions show significant differences it would imply that some aspect of the chemistry or physics in the model is incomplete or incorrect. If, however, the distributions are similar the model results would be verified for this particular set of source-receptor relationships under the given set of meteorological conditions, but the general validity of the physics would not necessarily be verified.

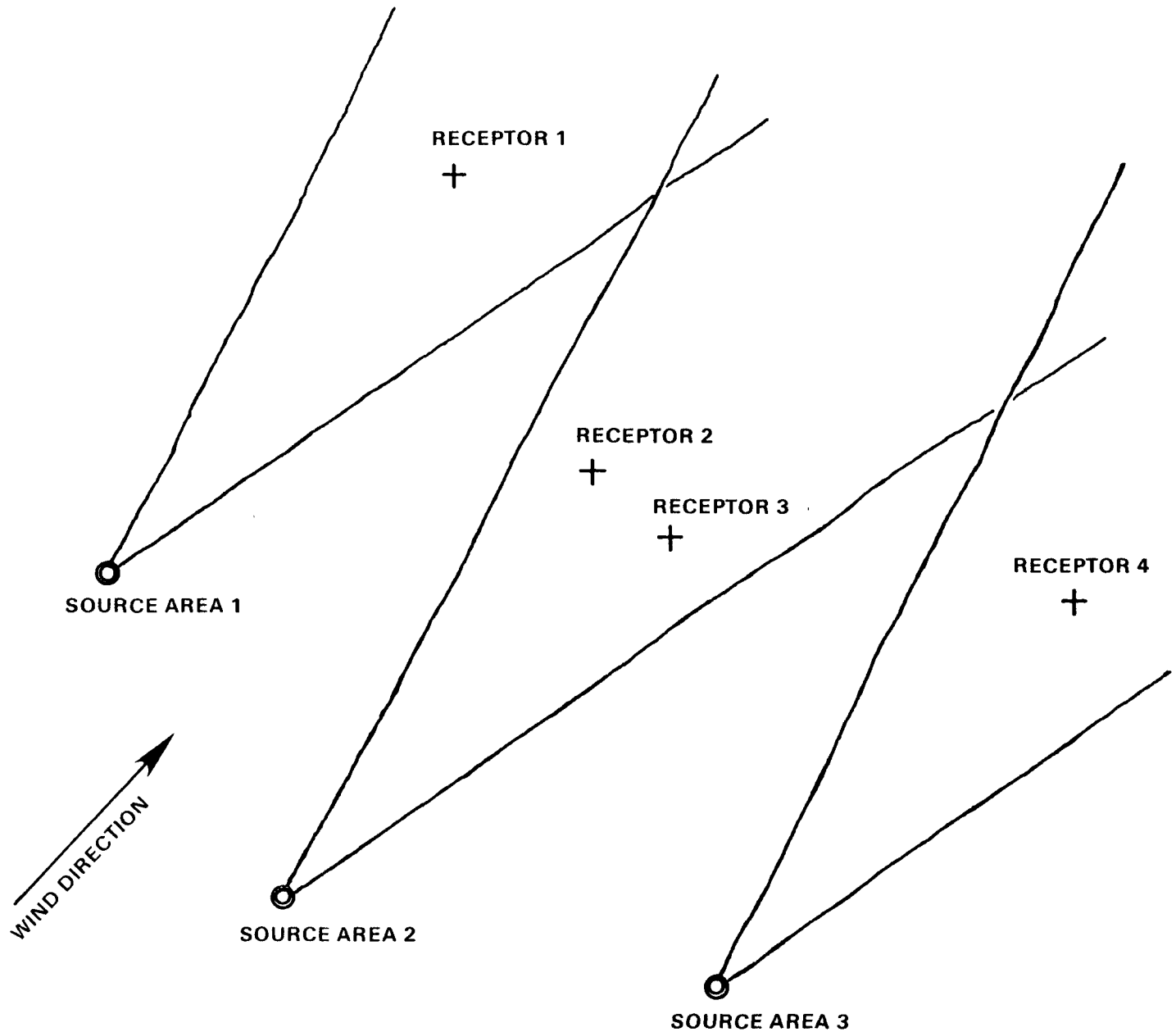


Figure 15. Source-receptor combinations in ROM evaluation over all realizations.

The test outlined above for the evaluation of the ROM over all realizations involves searching for a set of receptors that meet the criterion previously described. Dimensionally we are holding time fixed (within a single day period) and choosing a set of spatially varied receptors. There is an alternative path for evaluating the ROM over all realizations that involves focusing on only a single receptor at a time and allowing the time dimension to vary. This procedure would involve choosing a receptor within the downwind extent of an urban plume and examining the model results on a set of days that have essentially the same meteorological conditions and emissions functions. This last point implies a criterion that would assure invariance in the predicted concentration distribution from one selected day to the next at the chosen receptor. The observed distribution is then formed from the measured concentrations at the receptor on the selected days. A favorable comparison between the predicted and observed concentration distributions would imply that the ROM is mimicking the stochastic variability seen in the atmosphere. The problem with this evaluation method is that a very large data base spanning many weeks is required in order to select days that have virtually identical meteorology. In theory the approach is sound, but practically our application data base for the Northeast U.S. has only a limited temporal time span that is not sufficient to carry out this method of evaluation. For now this path of evaluation will be abandoned, although in the future it may be pursued given sufficient resources and a more tractable method of approach.

Thus the recommended method for the ROM evaluation over all realizations is the original one described here. The set of receptors used from day to day would vary depending on the wind direction. Preliminary trajectory analysis is necessary to determine whether there are potential source-receptor combinations available to carry out this analysis. Should it prove to be feasible the analysis should be made for ROM results from layers 0 and 1 for O_3 for the hour of the observed maximum and also possibly for all hours where the average concentration within the class of receptors is above a certain threshold value. Figure 16 is a structure diagram summarizing the steps

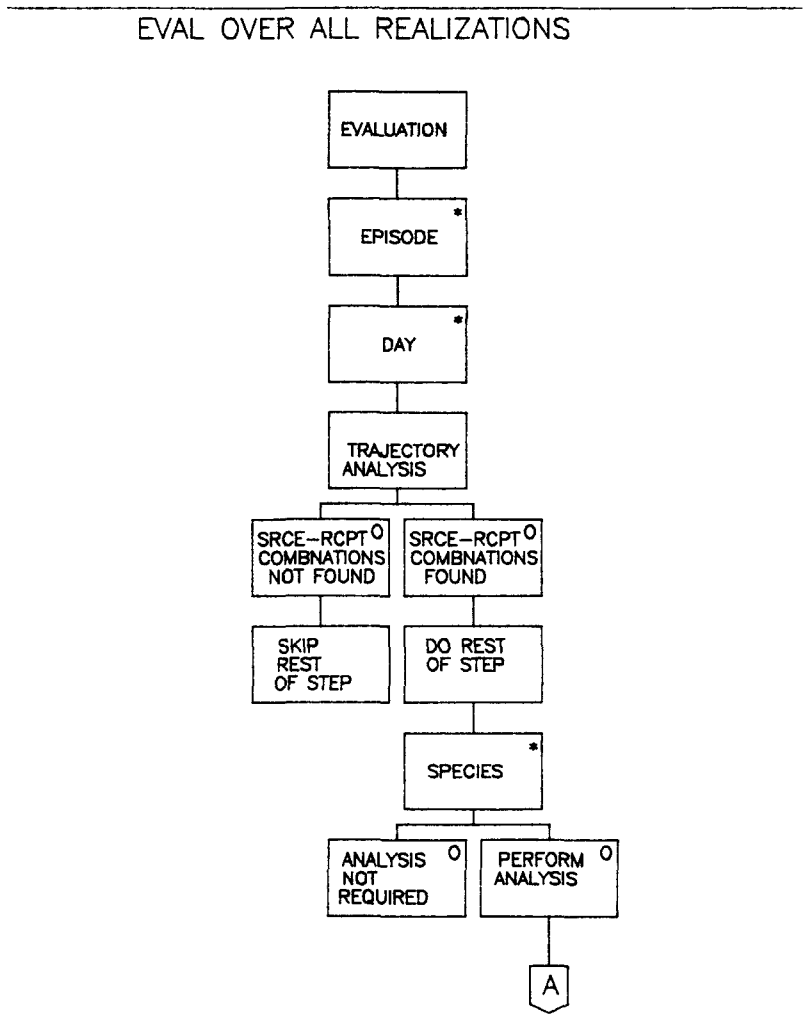


Figure 16. Structure diagram of the ROM evaluation over all realizations.

EVALUATION OVER ALL REALIZATIONS, PAGE 2

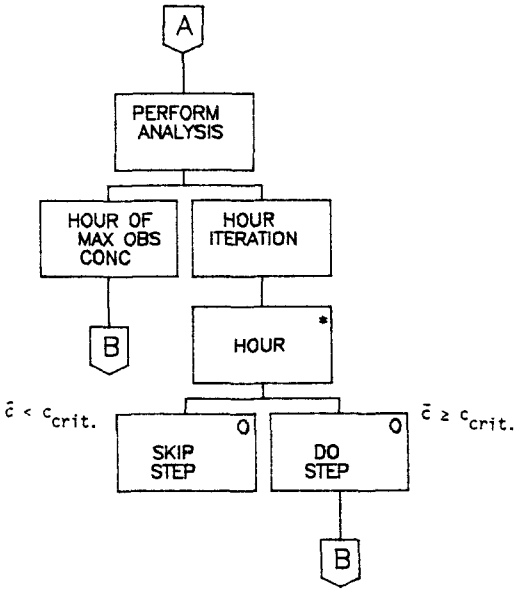


Figure 16. (continued)

EVALUATION OVER ALL REALIZATIONS, PAGE 3

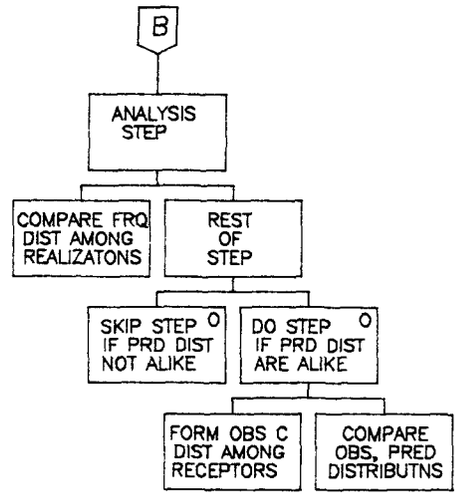


Figure 16. (concluded)

that are possible in the ROM evaluation over all realizations.

Various objective statistical tests may be used to compare the characteristics of the observed and predicted distributions. The standard t-test for comparison of means and the F-test for comparison of variances can be made. In addition, the Kolmogorov-Smirnov (K-S) test, a powerful "goodness-of-fit" test between two cumulative distributions can be made. The K-S test is valid for small sample size and it is conservative in that it minimizes the chances of making a Type II error (i.e., accepting the hypothesis that two samples have been drawn from the same population when, in fact, they are not).

FURTHER TESTING

An ideal experiment for evaluating the transport aspects of the ROM including the predictive uncertainty quantified in the multiple realization concept would involve the release of a controlled amount of passive tracer from a single source that is subsequently sampled in a dense receptor network on regional scales. A recent field study, the Cross-Appalachian Tracer Experiment (CAPTEX) conducted in 1983 was such a study (Clark et al., 1984). An inert trace gas was released on 5 separate occasions from a point source in Dayton, OH and on 2 other occasions from a Sudbury, Ontario point source. A dense sampling network was established in arcs at downwind distances of 300 to 1100 km from the Ohio release point in approximately 100 km intervals with additional sites to supplement the sampling from the Canadian release. In addition to the surface measurements, upper air wind data and aircraft sampling data are available as part of the CAPTEX data base.

An experiment could be designed using the ROM to simulate one or more of the tracer releases and subsequent plume evolution over the eastern U.S. and Canada. Some of the questions that might be investigated in this type of experiment include:

(1) Do the model's various realizations define the envelope in which the real plume is observed?

(2) What is the temporal and spatial variability in the real plume?

These questions, and others, might be answered more definitively in a controlled emission experiment such as this one. The current configuration of the ROM domain for the Northeast U.S. application would have to be enlarged on the west and north sides for this experiment to bring in the source locations. Alternatively, the sources could be specified as flux boundary conditions on the current domain, but this would involve making assumptions on the behavior of the gas from the source point to the model boundary, another potential source or error. Current resource levels preclude the undertaking of this modeling experiment with a new data base, although future ROM applications, when feasible, should include this one.

A model as complex as the ROM contains dozens of input parameters for which the model might be sensitive to slight errors in their specified values. A thorough sensitivity study of the ROM is beyond the scope of the project outlined here. However there are certain aspects that are relevant to the evaluation study and the present use of the model.

The choice of the horizontal grid cell size for the model domain is a significant decision for it determines the resolution on which regulators must base decisions. It also determines the resource level needed to operate the ROM in terms of computational time and memory requirements and the resolution and quantity of input data. For the ROM the grid-cell size will also affect the degree of subgrid scale chemistry that is accounted for in layer 0. One interesting sensitivity test that is recommended is to test the effect of horizontal grid size on model results. Since the current resolution of approximately 18.5 km was the resolution used to construct the source emissions inventory a smaller grid size would be difficult to use. However larger cells of double size (~37 km) or triple size (~55.5 km) might

be used in this test. For instance, would the predicted concentration in one large 37 km cell be equal to the average of the four 18.5 km cells contained within it? Discrepancies between these results might be attributed to subgrid scale processes.

Another test might be a ROM run to determine the contribution of individual large source areas on downwind receptors. This would involve an advection/dispersion run with the chemical kinetics turned off. Hypothetical tracers corresponding to the N species in the chemistry algorithm would be released from the N largest emission source areas, one species to an area. Then the concentration at downwind receptor sites could be analyzed for the component contributions from each source area. Results would be valid only for inert species, since the effects from chemical reactions could change the relative contributions based on advection/dispersion alone. Other tests might also be possible, depending upon the available resources.

SECTION 5

SUMMARY AND CONCLUSIONS

The U.S. EPA Regional Oxidant Model and NECRMP data base have been described here. The model incorporates a comprehensive description of the physical and chemical processes thought to be important to tropospheric O_3 production on 1000 km scales. The data base employed for the first application of the ROM was collected during the summers of 1979 and 1980 in the Northeast U.S. It contains meteorological and air quality data from regular monitoring networks as well as enhanced networks or special field project measurements made during that period.

The evaluation procedure that will be used to determine the ROM performance on this data base has been outlined. A number of episodes will be simulated from the period July 23 through August 16, 1980 for which performance statistics will be developed. The evaluation of any given day within an episode will proceed in 2 distinct stages. The first stage will focus on model performance for an individual model realization, irrespective of all other realizations. Model realizations for a given day are functions of the possible flow fields that existed for the day. The second stage will attempt to evaluate model performance using the full probabilistic abilities of the ROM that consider all realizations concurrently. The focus of the evaluation will be on O_3. Whether other species shall be considered and the exact pathway through the evaluation study will be determined by the resources available at the time.

In conclusion it should be worthwhile to review the intended uses of the ROM. The model was originally conceived as a way to provide the initial and boundary conditions required by urban scale models. The modeling system would be structured as shown in Figure 17. The ROM would first be run using projected future year emissions to establish initial and boundary conditions for the urban model. The ROM inputs would also include worst-case meteorology.

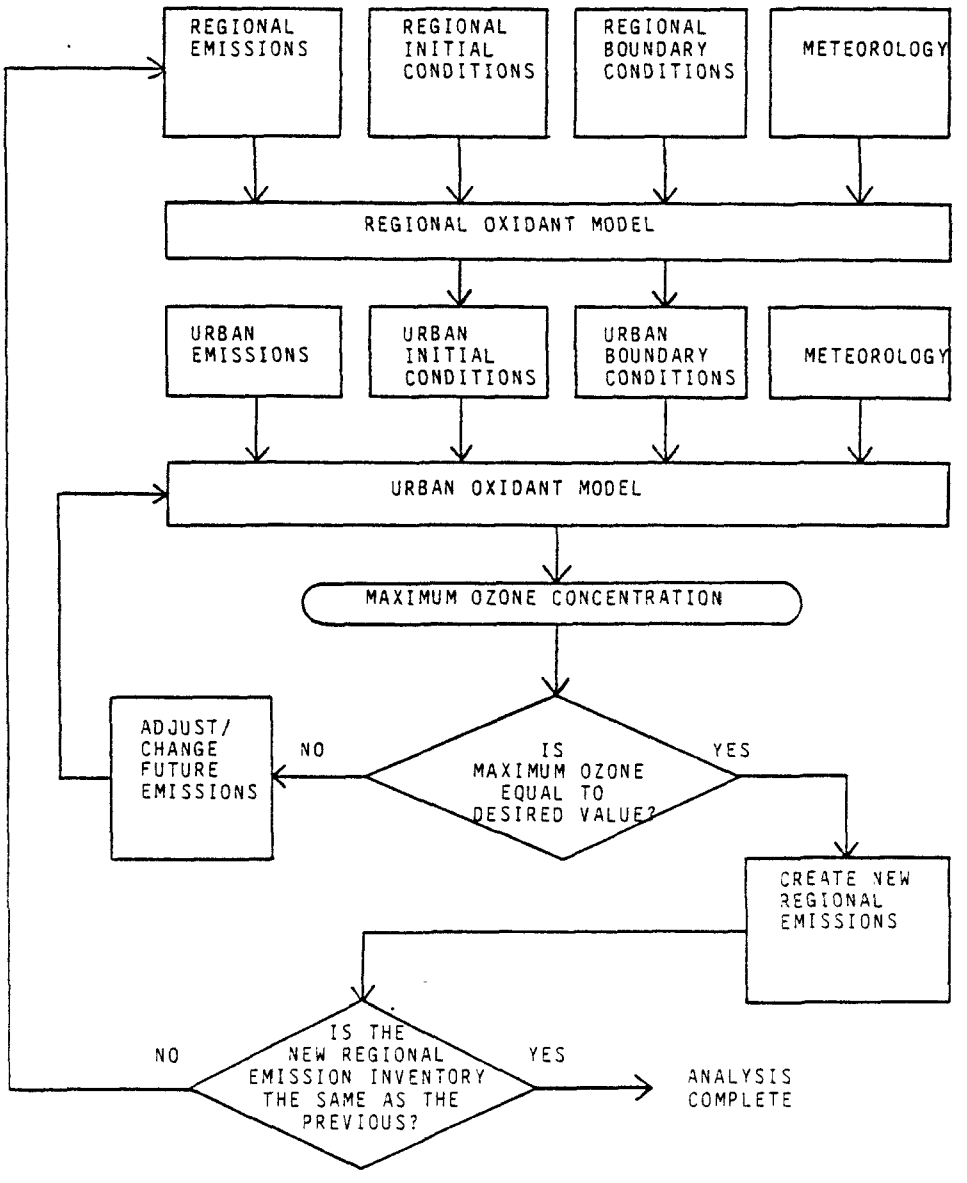


Figure 17. Emission reduction analysis using the ROM and urban scale models.

Model results would be used to specify the initial concentrations and boundary concentrations for each urban area. The urban worst-case meteorology would be a subset of the regional worst-case scenario. The emissions input into the urban model would be adjusted until the model predictions showed that the reduction in peak ozone concentration would be sufficient to insure that the oxidant standards are not violated.

After the emission reductions are determined for all urban areas a new regional emissions data base, describing the future year emissions assuming the implementation of control strategies, can be compared with the original regional emissions. If the emissions differ greatly, in either time or space, then the ROM would be run again to determine new urban initial and boundary concentrations. If the new initial and boundary conditions are substantially different, then the urban models would be run again to refine the emission reduction requirements. The procedure would be iterative.

Another use of the ROM has also been proposed. It would be used to predict ozone concentrations in urban and rural areas directly. The direct use of the ROM is illustrated in Figure 18. Emissions are adjusted and model predictions are made until the desired ozone reductions are achieved in all areas in the region. Also along these lines there is an interest in using the model to predict longer-term regional average concentrations. For example, a growing-season O_3 average is appropriate for determining potential crop losses in the agricultural belts of the midwest and central plains areas of the U.S. Logistical problems exist in a long-term ROM application, such as assembling the data base needed to make such a run or choosing a subset of days to represent an entire season.

Knowledge of how the ROM will be used is important in determining the objectives of the model evaluation. All of the potential uses of the model described here will be considered when the evaluation is performed and the results are analyzed. Explicit emissions scenario testing, however, is not a part of this study, but may be conducted in a future study.

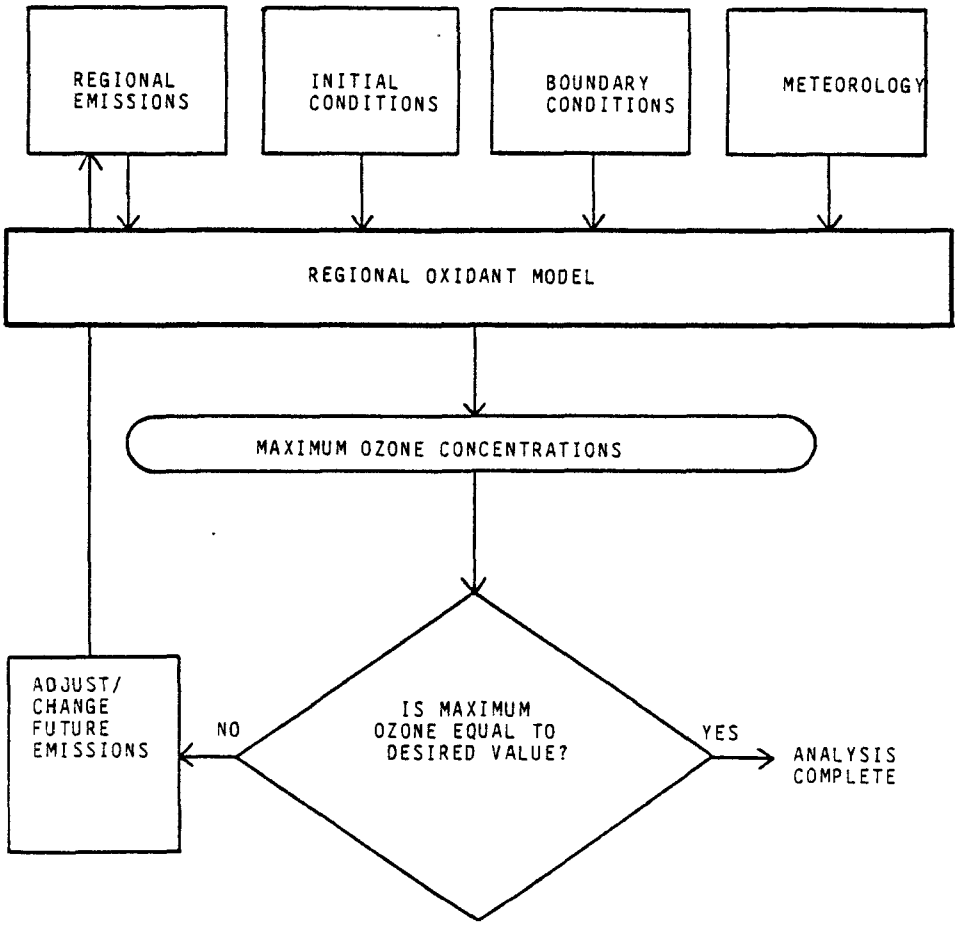


Figure 18. Emission reduction analysis using the ROM alone.

REFERENCES

Clark, T. L., G. J. Ferber, J. L. Heffter, and R. R. Draxler, 1984: Cross-Appalachian tracer experiment (CAPTEX '83). In Proceedings of the 4th Joint Conference on Applications of Air Pollution Meteorology, 16-19 October 1984, Portland, OR, American Meteorological Society, Boston, MA, 158-161.

David, M., 1977: Geostatistical Ore Reserve Estimation, Elsevier Scientific Publishing Co., Amsterdam, 364 pages.

Delhomme, J. P., 1978: Kriging in the hydrosciences. Advances in Water Resources, 1, 251-266.

Dennis, R. L., M. W. Downton, and R. S. Keil, 1983: Evaluation of performance measures for an urban photochemical model. EPA-450/4-83-021, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Dennis, R. L. and J. S. Irwin, 1985: Current views of model performance evaluation. In Proceedings of the DOE/AMS Model Evaluation Workshop (23-26 October 1984, Kiawah,SC), Vol. I: Participants and Invited Speakers Papers, A.H. Weber and A.J. Garret, eds, E.I. DuPont de Nemours and Co., Savannah River Laboratory, Aiken, SC 29808.

EPA, 1982: Northeast corridor regional modeling project - annual emission inventory compilation and formatting, Volume 1: Project approach. EPA-450/4-82-013a, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Freas, W. P., 1983: Northeast corridor regional modeling project-data base description. EPA-450/4-83-015a, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Gandin, L. S., 1970: The Planning of Meteorological Station Networks, WMO, No. 265. TP. 149, Geneva, Switzerland, 35 pp.

Lamb, R. G., 1983: A regional scale (1000 km) model of photochemical air pollution. Part 1 - Theoretical formulation. EPA-600/3-83-035, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Lamb, R. G., 1984: A regional scale (1000 km) model of photochemical air pollution. Part 2 - Input processor network design. EPA-600/3-84-085, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Possiel, N. C. and W. P. Freas, 1982: Northeast corridor regional modeling project - description of the 1980 urban field studies. EPA-450/4-82-018, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Sellars, F. M., T. E. Fitzgerald, J. M. Lennon, L. J. Maiocco, N. M. Monzione, and D. R. Neal, 1985: National Acid Precipitation Assessment Program emission inventory allocation factors. EPA Contract Report 68-02-3698, U. S. Environmental Protection Agency, Research Triangle Park, NC (March 1985).

Siegel, A. F., 1982: Geometric data analysis: an interactive graphics program for shape comparison. In Modern Data Analysis, R. L. Launer and A. F. Siegel, eds., Academic Press, New York, 103-122.

APPENDIX A

SURFACE MONITORING SITES

This appendix contains a listing of all surface monitoring sites in the special NECRMP field studies and the SAROAD networks in the Northeast U.S. taking 1-h air quality measurements of O_3 during 1980.

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
1	070060123F01	BRIDGEPORT	FAIRFIELD CO	CONNECTICUT	41 11 12 N	073 10 35 W	015	0020	215
2	070175123F01	DANBURY	FAIRFIELD CO	CONNECTICUT	41 24 05 N	073 26 41 W	015	0380	171
3	070190123F01	DERBY	NEW HAVEN CO	CONNECTICUT	41 18 59 N	073 05 16 W	015	0045	147
4	070330004F01	GREENWICH	FAIRFIELD CO	CONNECTICUT	41 04 46 N	073 42 19 W	006	0450	292
5	070420123F01	HARTFORD	HARTFORD CO	CONNECTICUT	41 45 39 N	072 40 44 W	015	0085	305
6	070478001F03	LITCHFIELD CO	LITCHFIELD CO	CONNECTICUT	41 40 21 N	073 08 35 W	006	0570	164
7	070570007F01	MIDDLETOWN	MIDDLESEX CO	CONNECTICUT	41 33 07 N	072 37 51 W	075	0190	167
8	070700123F01	NEW HAVEN	NEW HAVEN CO	CONNECTICUT	41 18 38 N	072 55 03 W	015	0025	235
9	071060001F01	STAFFORD	TOLLAND CO	CONNECTICUT	41 58 33 N	072 23 12 W	016	0870	146
10	071110007F01	STRATFORD	FAIRFIELD CO	CONNECTICUT	41 09 10 N	073 09 04 W	033	0010	172
11	080020001F01	DOVER	KENT CO	DELAWARE	39 08 55 N	075 30 00 W	014	0034	202
12	080180001F01	NEW CASTLE CO	NEW CASTLE CO	DELAWARE	39 48 30 N	075 27 00 W	014	0099	344
13	080180003F01	NEW CASTLE CO	NEW CASTLE CO	DELAWARE	39 46 30 N	075 31 00 W	014	0214	268
14	080180070A05	NEW CASTLE CO	NEW CASTLE CO	DELAWARE	39 32 00 N	075 41 11 W	003	0015	32
15	080180070F05	NEW CASTLE CO	NEW CASTLE CO	DELAWARE	39 32 00 N	075 41 11 W	003	0015	30
16	080240001A05	SUSSEX CO	SUSSEX CO	DELAWARE	38 35 03 N	075 16 25 W	020	0020	69
17	180080010F01	ASHLAND	BOYD CO	KENTUCKY	38 26 37 N	082 36 47 W	010	0570	298
18	180280003F03	BOONE CO	BOONE CO	KENTUCKY	38 55 05 N	084 51 10 W	012	0562	223
19	180500004F05	CALLOWAY CO	CALLOWAY CO	KENTUCKY	36 44 02 N	088 07 00 W	017	0420	163
20	180800007F01	COVINGTON	KENTON CO	KENTUCKY	39 04 15 N	084 31 31 W	014	0204	149
21	180900004F05	DANVILLE	BOYLE CO	KENTUCKY	37 40 28 N	084 46 03 W	010	0900	161
22	180960001F01	DAYTON	CAMPBELL CO	KENTUCKY	39 06 26 N	084 28 36 W	012	0552	352
23	181160001F01	FAYETTE CO	FAYETTE CO	KENTUCKY	38 07 35 N	084 28 00 W	012	0955	300
24	181580012F01	HANCOCK CO	HANCOCK CO	KENTUCKY	37 56 19 N	086 53 51 W	012	0400	69
25	181600004F05	HARDIN CO	HARDIN CO	KENTUCKY	37 45 27 N	085 43 14 W	010	0480	203
26	181680001F05	HARRODSBURG	MERCER CO	KENTUCKY	37 45 54 N	084 50 02 W	010	0850	153
27	181740008F01	HENDERSON	HENDERSON CO	KENTUCKY	37 51 21 N	087 34 31 W	015	0445	346
28	181840001F05	HOPKINS CO	HOPKINS CO	KENTUCKY	37 31 37 N	087 21 15 W	012	0450	157
29	181920027G03	JEFFERSON CO	JEFFERSON CO	KENTUCKY	38 08 14 N	085 34 42 W	024	0716	245
30	181920033G01	JEFFERSON CO	JEFFERSON CO	KENTUCKY	38 18 20 N	085 34 46 W			62
31	181920034G02	JEFFERSON CO	JEFFERSON CO	KENTUCKY	38 06 27 N	085 52 05 W	024	0524	42
32	181920041G05	JEFFERSON CO	JEFFERSON CO	KENTUCKY	38 09 56 N	085 52 57 W	015	0445	303
33	182300012F01	LEXINGTON-FAYETTE	FAYETTE CO	KENTUCKY	38 03 47 N	084 29 45 W	014	0099	321
34	182380015G01	LOUISVILLE	JEFFERSON CO	KENTUCKY	38 15 02 N	085 48 49 W	008	0458	36
35	182380021G01	LOUISVILLE	JEFFERSON CO	KENTUCKY	38 15 50 N	085 42 42 W	030	0530	260

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
36	182460027N02	MC CRACKEN CO	MC CRACKEN CO	KENTUCKY	37 09 21 N	088 47 14 W		0345	102
37	182960005N02	MUHLENBERG CO	MUHLENBERG CO	KENTUCKY	37 12 18 N	086 56 25 W	012	0470	317
38	182960009N02	MUHLENBERG CO	MUHLENBERG CO	KENTUCKY	37 18 41 N	087 00 04 W		0430	189
39	183080031F05	OHIO CO	OHIO CO	KENTUCKY	37 33 22 N	086 46 25 W	015	0435	83
40	183140005F01	OWENSBORO	DAVIESS CO	KENTUCKY	37 46 57 N	087 04 30 W	010	0400	338
41	183180020F01	PADUCAH	MC CRACKEN CO	KENTUCKY	37 04 59 N	088 36 16 W	020	0360	194
42	183180024F01	PADUCAH	MC CRACKEN CO	KENTUCKY	37 03 30 N	088 34 20 W	016	0035	82
43	183460002F05	PULASKI CO	PULASKI CO	KENTUCKY	37 09 31 N	084 28 59 W	012	0910	216
44	183710004F05	SHEPHERDSVILLE	BULLITT CO	KENTUCKY	37 59 51 N	085 42 14 W	012	0046	90
45	184100002F05	WINCHESTER	CLARK CO	MAINE	38 00 16 N	084 10 05 W	012	0970	91
46	200250001F05	CAPE ELIZABETH	CUMBERLAND CO	MAINE	43 33 57 N	070 12 00 W			142
47	200260001F05	CARIBOU	AROOSTOOK CO	MAINE	46 52 03 N	068 00 53 W	004	0189	138
48	200277002F05	CUMBERLAND CO	CUMBERLAND CO	MAINE	43 54 31 N	070 13 50 W	014	0184	131
49	200460001F05	GARDINER	KENNEBEC CO	MAINE	44 13 30 N	069 47 26 W	015	0180	139
50	200907004F05	PENOBSCOT CO	PENOBSCOT CO	MAINE	45 00 38 N	068 38 04 W	015	0126	155
51	200907005F05	PENOBSCOT CO	PENOBSCOT CO	MAINE	44 43 57 N	068 58 52 W	012	0213	132
52	200907006F05	PENOBSCOT CO	PENOBSCOT CO	MAINE	45 14 54 N	068 34 05 W	003	0055	96
53	210080001G01	ANNE ARUNDEL CO	ANNE ARUNDEL CO	MARYLAND	39 12 42 N	076 39 12 W	015	0075	56
54	210080003G01	ANNE ARUNDEL CO	ANNE ARUNDEL CO	MARYLAND	39 10 11 N	076 37 40 W	005	0045	228
55	210080014G01	ANNE ARUNDEL CO	ANNE ARUNDEL CO	MARYLAND	38 54 10 N	076 39 10 W	013	0145	117
56	210080019G05	ANNE ARUNDEL CO	ANNE ARUNDEL CO	MARYLAND	39 06 05 N	076 43 40 W	015	0151	151
57	210120004H01	BALTIMORE		MARYLAND	39 18 00 N	076 36 45 W	010	1100	186
58	210120011H01	BALTIMORE		MARYLAND	39 14 12 N	076 35 04 W	010	0035	244
59	210120018F01	BALTIMORE		MARYLAND	39 18 53 N	076 36 51 W	015	0145	263
60	210120019F01	BALTIMORE		MARYLAND	39 17 12 N	076 37 30 W	010	0040	37
61	210140003G01	BALTIMORE CO	BALTIMORE CO	MARYLAND	39 24 21 N	076 45 55 W	015	0540	308
62	210140010G01	BALTIMORE CO	BALTIMORE CO	MARYLAND	39 30 54 N	076 25 52 W	012	0340	30
63	210200005F01	BETHESDA	MONTGOMERY CO	MARYLAND	38 59 45 N	077 06 18 W	012	0332	121
64	210420080J05	CECIL CO	CECIL CO	MARYLAND	39 43 07 N	076 06 22 W	021	0042	65
65	210480004F01	CHEVERLY	PRINCE GEORGES	MARYLAND	38 55 26 N	076 54 16 W	010	0070	107
66	210500007G01	COCKEYSVILLE	BALTIMORE CO	MARYLAND	39 27 40 N	076 37 52 W	013	0413	318
67	210560006F01	CUMBERLAND	ALLEGANY CO	MARYLAND	39 38 42 N	078 45 47 W	015	0700	126
68	210565001F01	DAMASCUS	MONTGOMERY CO	MARYLAND	39 16 28 N	077 13 17 W	013	0740	130
69	210650001F01	EDGEWOOD	HARFORD CO	MARYLAND	39 24 35 N	076 17 50 W	014	0020	194
70	210680001G01	ESSEX	BALTIMORE CO	MARYLAND	39 18 38 N	076 28 29 W	015	0033	293

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
71	210740034G01	FREDERICK CO	FREDERICK CO	MARYLAND	39 28 47 N	077 13 55 W	015	0500	158
72	210840002F01	GREENBELT	PRINCE GEORGES	MARYLAND	39 01 12 N	076 49 40 W	010	0160	147
73	210860004F01	HAGERSTOWN	WASHINGTON CO	MARYLAND	39 38 44 N	077 43 12 W	010	0635	66
74	210920080J05	HARFORD CO	HARFORD CO	MARYLAND	39 37 37 N	076 19 46 W	022	0034	67
75	210960005G01	HOWARD CO	HOWARD CO	MARYLAND	39 10 10 N	076 53 55 W	010	0400	103
76	211360002G01	RIVIERA BEACH	ANNE ARUNDEL CO	MARYLAND	39 09 35 N	076 30 42 W	015	0045	152
77	211380001F01	ROCKVILLE	MONTGOMERY CO	MARYLAND	39 06 48 N	077 06 26 W	130	4150	124
78	211560001F01	SUITLAND-SILVER HILL	PRINCE GEORGES	MARYLAND	38 51 12 N	076 55 55 W	015	0040	300
79	220030002F01	AGAWAM	HAMPDEN CO	MASSACHUSETTS	42 02 08 N	072 38 09 W	018	0112	136
80	220120003F01	ATTLEBORO	BRISTOL CO	MASSACHUSETTS	41 55 48 N	071 15 54 W	015	0130	201
81	220240016F01	BOSTON	SUFFOLK CO	MASSACHUSETTS	42 22 06 N	071 02 25 W	015	0025	29
82	220240021F01	BOSTON	SUFFOLK CO	MASSACHUSETTS	42 22 41 N	071 01 42 W	014	0020	189
83	220372001F01	CHARLTON	WORCESTER CO	MASSACHUSETTS	00 00 00 0	000 00 00 0	030	1030	146
84	220480002F01	DANVERS	ESSEX CO	MASSACHUSETTS	42 35 31 N	070 58 37 W	040	0120	70
85	220535001F01	EASTON	BRISTOL CO	MASSACHUSETTS	42 04 02 N	071 05 26 W	014	0150	85
86	220730001F01	GEORGETOWN	ESSEX CO	MASSACHUSETTS	00 00 00 0	000 00 00 0	012	0100	346
87	220795001F01	HAMILTON	ESSEX CO	MASSACHUSETTS	42 37 50 N	070 48 55 W	018	0185	64
88	221210001F01	MEDFIELD	NORFOLK CO	MASSACHUSETTS	42 12 44 N	071 20 08 W	075	0275	200
89	221220003F01	MEDFORD	MIDDLESEX CO	MASSACHUSETTS	42 24 16 N	071 04 58 W	015	0050	57
90	221392001A05	MONTAGUE	FRANKLIN CO	MASSACHUSETTS	42 34 01 N	072 31 42 W	014	0329	73
91	221800005F01	PITTSFIELD	BERKSHIRE CO	MASSACHUSETTS	42 27 09 N	073 12 04 W	015	1015	205
92	221880002F01	QUINCY	NORFOLK CO	MASSACHUSETTS	42 14 40 N	070 58 19 W	015	0035	68
93	222100003F01	SOMERVILLE	MIDDLESEX CO	MASSACHUSETTS	42 24 13 N	071 06 30 W	014	0035	127
94	222196001F01	SUDBURY	MIDDLESEX CO	MASSACHUSETTS	42 23 06 N	071 23 20 W	014	0130	121
95	222252001F01	TEWKSBURY	MIDDLESEX CO	MASSACHUSETTS	42 36 29 N	071 13 14 W	012	0150	80
96	222380005F01	WATERTOWN	MIDDLESEX CO	MASSACHUSETTS	42 22 29 N	071 10 57 W	013	0063	132
97	222640012F01	WORCESTER	WORCESTER CO	MASSACHUSETTS	42 15 31 N	071 48 02 W	013	0493	166
98	222640015F01	WORCESTER	WORCESTER CO	MASSACHUSETTS	00 00 00 0	000 00 00 0	015	0900	77
99	230160001G01	ALLEN PARK	WAYNE CO	MICHIGAN	42 13 43 N	083 12 37 W	014	6000	311
100	231180016G01	DETROIT	WAYNE CO	MICHIGAN	42 21 30 N	083 05 51 W	012	0625	341
101	231180019G01	DETROIT	WAYNE CO	MICHIGAN	42 26 00 N	083 00 13 W	014	0630	343
102	231440006F01	ESSEXVILLE	BAY CO	MICHIGAN	43 36 56 N	083 50 43 W	012	0585	362
103	231580011F01	FLINT	GENESEE CO	MICHIGAN	43 01 29 N	083 41 09 W	027	0740	323
104	231700001F02	GENESEE CO	GENESEE CO	MICHIGAN	43 10 06 N	083 27 42 W	018	0840	208
105	231820002H01	GRAND RAPIDS	KENT CO	MICHIGAN	42 58 27 N	085 40 32 W	012	0602	131

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
106	232720001H02	KENT CO	KENT CO	MICHIGAN	43 02 31 N	085 24 45 W	022	0855	182
107	232840002F01	LANSING	INGHAM CO	MICHIGAN	42 44 21 N	084 32 47 W	010	0850	61
108	232840012F01	LANSING	INGHAM CO	MICHIGAN	42 44 15 N	084 32 10 W	011	0880	94
109	233040002G01	LIVONIA	WAYNE CO	MICHIGAN	42 22 55 N	083 24 14 W	012	0682	199
110	233140009G02	MACOMB CO	MACOMB CO	MICHIGAN	42 43 29 N	082 47 30 W	016	0620	152
111	233280003F01	MARQUETTE CO	MARQUETTE CO	MICHIGAN	46 48 52 N	087 43 41 W	035	0680	333
112	233740022G01	MUSKEGON	MUSKEGON CO	MICHIGAN	43 14 15 N	086 12 07 W	044	0625	31
113	233760006G01	MUSKEGON CO	MUSKEGON CO	MICHIGAN	43 17 03 N	086 13 20 W	042	0640	320
114	234340003F01	PORT HURON	ST CLAIR CO	MICHIGAN	42 56 52 N	082 26 30 W	015	0610	351
115	234790001F05	SALINE	WASHTENAW CO	MICHIGAN	42 09 20 N	083 46 41 W	021	0740	225
116	234880002F01	SOUTHFIELD	OAKLAND CO	MICHIGAN	42 27 14 N	083 13 10 W	014	0670	343
117	235260003G01	WARREN	MACOMB CO	MICHIGAN	42 30 48 N	083 00 22 W	011	0617	349
118	300420012F01	MANCHESTER	HILLSBOROUGH CO	NEW HAMPSHIRE	42 59 34 N	071 27 52 W	010	0264	228
119	300540009F01	PORTSMOUTH	ROCKINGHAM CO	NEW HAMPSHIRE	43 04 41 N	070 45 46 W	010		298
120	310080005F01	ATLANTIC CO	ATLANTIC CO	NEW JERSEY	39 31 29 N	074 27 35 W	013	0010	292
121	310180003F01	BAYONNE	HUDSON CO	NEW JERSEY	40 40 52 N	074 07 11 W	015	0005	268
122	310720003F01	CAMDEN	CAMDEN CO	NEW JERSEY	39 55 23 N	075 05 50 W	012	0020	300
123	310740001F01	CAMDEN CO	CAMDEN CO	NEW JERSEY	39 40 26 N	074 51 25 W	013	0115	306
124	310780003F01	CAPE MAY CO	CAPE MAY CO	NEW JERSEY	39 06 13 N	074 48 54 W	012	0010	225
125	311120001F01	DUMONT	BERGEN CO	NEW JERSEY	40 56 16 N	073 59 42 W	015	0085	223
126	311160003F01	EAST ORANGE	ESSEX CO	NEW JERSEY	40 45 23 N	074 12 02 W	013	0150'	277
127	311500001F01	FLEMINGTON	HUNTERDON CO	NEW JERSEY	40 31 00 N	074 48 36 W	015	0130	213
128	312580002J05	LINDEN	UNION CO	NEW JERSEY	40 36 15 N	074 15 12 W	000	0020	83
129	312750001F01	MC GUIRE AFB	BURLINGTON CO	NEW JERSEY	40 02 55 N	074 35 02 W	016	0130	275
130	313180004J05	MONMOUTH CO	MONMOUTH CO	NEW JERSEY	40 19 40 N	074 16 00 W	000	0150	78
131	313260001F01	MORRIS CO	MORRIS CO	NEW JERSEY	40 47 16 N	074 40 39 W	015	0900	128
132	313480008F01	NEWARK	ESSEX CO	NEW JERSEY	40 43 58 N	074 10 41 W	015	0030	143
133	313500002F01	NEW BRUNSWICK	MIDDLESEX CO	NEW JERSEY	40 28 56 N	074 26 27 W	012	0085	297
134	314300001F01	PLAINFIELD	SOMERSET CO	NEW JERSEY	40 36 04 N	074 26 31 W	013	0060	187
135	315060001F01	SOMERVILLE	SOMERSET CO	NEW JERSEY	40 34 10 N	074 36 49 W	016	0070	304
136	315400002F01	TRENTON	MERCER CO	NEW JERSEY	40 13 12 N	074 45 54 W	015	0050	303
137	315560002F01	VINELAND	CUMBERLAND CO	NEW JERSEY	39 29 32 N	075 01 15 W	012	0104	285
138	330130002F01	AMHERST	ERIE CO	NEW YORK	42 59 33 N	078 46 35 W	016	0600	230
139	330280002F01	BABYLON	SUFFOLK CO	NEW YORK	40 44 08 N	073 24 22 W	023	0085	335
140	330480007F01	BINGHAMTON	BROOME CO	NEW YORK	42 05 58 N	075 53 37 W	016	0837	327

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
141	331880003F01	ELMIRA	CHEMUNG CO	NEW YORK	42 06 34 N	076 48 08 W	013	0835	336
142	332020002F03	ESSEX CO	ESSEX CO	NEW YORK	44 21 54 N	073 54 06 W	050	4900	334
143	332480003F01	GLENS FALLS	WARREN CO	NEW YORK	43 19 00 N	073 37 26 W	016	0360	338
144	332900005F01	HEMPSTEAD	NASSAU CO	NEW YORK	40 44 43 N	073 35 13 W	015	0090	339
145	332900010J05	HEMPSTEAD	NASSAU CO	NEW YORK	40 53 30 N	073 35 55 W	000	0100	77
146	334100002F01	MAMARONECK	WESTCHESTER CO	NEW YORK	40 55 53 N	073 45 50 W	013	0030	342
147	334380005F01	MONROE CO	MONROE CO	NEW YORK	43 04 24 N	077 42 39 W	010	0535	304
148	334680004F01	NEW YORK CITY	QUEENS CO	NEW YORK	40 44 09 N	073 49 01 W	066	0100	341
149	334680007F01	NEW YORK CITY	KINGS CO	NEW YORK	40 35 33 N	073 56 15 W	057	0010	217
150	334680010F01	NEW YORK CITY	NEW YORK CO	NEW YORK	40 44 22 N	073 59 10 W	125	0125	226
151	334680067F01	NEW YORK CITY	RICHMOND CO	NEW YORK	40 35 50 N	074 07 35 W	066	0300	234
152	334680070F01	NEW YORK CITY	QUEENS CO	NEW YORK	40 46 39 N	073 54 31 W	020	0050	74
153	334680076F01	NEW YORK CITY	KINGS CO	NEW YORK	40 40 19 N	073 58 40 W	036	0090	78
154	334720006F01	NIAGARA CO	NIAGARA CO	NEW YORK	43 13 08 N	078 28 45 W	016	0461	150
155	334740006F01	NIAGARA FALLS	NIAGARA CO	NEW YORK	43 05 09 N	078 59 47 W	013	0570	320
156	335060004F01	ONEIDA CO	ONEIDA CO	NEW YORK	43 18 14 N	075 43 17 W	016	0512	311
157	335620003F01	POUGHKEEPSIE	DUTCHESS CO	NEW YORK	41 42 12 N	073 56 08 W	015	0100	327
158	335680001F01	RENSSELAER	RENSSELAER CO	NEW YORK	42 37 39 N	073 45 04 W	014	0015	341
159	335760004F01	ROCHESTER	MONROE CO	NEW YORK	43 10 01 N	077 31 01 W	013	0425	330
160	336020003F01	SCHENECTADY	SCHENECTADY CO	NEW YORK	42 47 47 N	073 56 22 W	016	0348	237
161	336520001J05	STONY BROOK	SUFFOLK CO	NEW YORK	40 54 45 N	073 07 45 W	000	0150	77
162	336620011F01	SYRACUSE	ONONDAGA CO	NEW YORK	43 02 43 N	076 08 52 W	000	0000	352
163	336620014F01	SYRACUSE	ONONDAGA CO	NEW YORK	43 03 47 N	076 10 33 W	015	0365	81
164	336620015F01	SYRACUSE	ONONDAGA CO	NEW YORK	43 03 36 N	076 07 36 W	015	0450	95
165	336880004F01	UTICA	ONEIDA CO	NEW YORK	43 05 59 N	075 11 48 W	016	0400	337
166	337240001F01	WAYNE CO	WAYNE CO	NEW YORK	43 13 49 N	077 10 22 W	016	0460	325
167	337480004F01	WHITE PLAINS	WESTCHESTER CO	NEW YORK	41 03 08 N	073 45 51 W	026	0195	337
168	360060018H01	AKRON	SUMMIT CO	OHIO	41 04 35 N	081 31 05 W	012	1022	119
169	360060020H01	AKRON	SUMMIT CO	OHIO	41 06 22 N	081 30 12 W	012	1002	236
170	360080001F01	ALLEN CO	ALLEN CO	OHIO	40 45 28 N	084 04 03 W	020	0900	31
171	360360008H01	BARBERTON	SUMMIT CO	OHIO	40 59 55 N	081 37 52 W	012	0106	113
172	360580002H01	BEREA	CUYAHOGA CO	OHIO	41 22 17 N	081 51 13 W	025	0790	295
173	361000001H02	CANTON	STARK CO	OHIO	40 47 50 N	081 22 30 W	025	1065	282
174	361220019H01	CINCINNATI	HAMILTON CO	OHIO	39 06 28 N	084 32 52 W	015	0525	151
175	361220020H01	CINCINNATI	HAMILTON CO	OHIO	39 12 39 N	084 28 56 W	015	0575	361

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
176	361220035H01	CINCINNATI	HAMILTON CO	OHIO	39 08 09 N	084 30 29 W	014	0790	43
177	361260001G01	CLARK CO	CLARK CO	OHIO	40 00 04 N	083 48 16 W	003	0953	351
178	361280002H01	CLERMONT CO	CLERMONT CO	OHIO	39 04 33 N	084 10 34 W	040	0600	149
179	361300034H01	CLEVELAND	CUYAHOGA CO	OHIO	41 33 17 N	081 34 28 W	025	0653	194
180	361340001F03	CLINTON CO	CLINTON CO	OHIO	39 25 38 N	083 47 14 W	012	1069	167
181	361460004F01	COLUMBUS	FRANKLIN CO	OHIO	40 00 00 N	083 02 30 W	017	0800	243
182	361460015F01	COLUMBUS	FRANKLIN CO	OHIO	39 57 44 N	082 59 33 W	020	0790	255
183	361460081F01	COLUMBUS	FRANKLIN CO	OHIO	40 05 13 N	082 57 37 W			270
184	361480001F01	CONNEAUT	ASHTABULA CO	OHIO	41 57 34 N	080 34 21 W	025	0660	352
185	361660019G01	DAYTON	MONTGOMERY CO	OHIO	39 48 50 N	084 11 40 W	008	0838	357
186	361660025G01	DAYTON	MONTGOMERY CO	OHIO	39 45 33 N	084 12 00 W	010	0770	331
187	361980002F01	ELYRIA	LORAIN CO	OHIO	41 22 01 N	082 06 20 W	038	0733	60
188	361980003F01	ELYRIA	LORAIN CO	OHIO	41 21 58 N	082 06 30 W	030	0765	244
189	362720006H01	HAMILTON CO	HAMILTON CO	OHIO	39 16 46 N	084 21 58 W	018	0878	151
190	362720010H02	HAMILTON CO	HAMILTON CO	OHIO	00 00 00 0	000 00 00 0	015	0519	360
191	363400002H01	LEBANON	WARREN CO	OHIO	39 25 37 N	084 12 09 W	035	0785	153
192	363560002F01	LOGAN	HOCKING CO	OHIO	39 31 38 N	082 22 33 W	020	0760	61
193	363920001I03	MAHONING CO	MAHONING CO	OHIO	41 06 53 N	080 58 58 W	010	0970	80
194	364160001F01	MEDINA CO	MEDINA CO	OHIO	41 09 17 N	081 54 10 W	020	1060	361
195	364340004H01	MIDDLETOWN	BUTLER CO	OHIO	39 31 53 N	084 23 29 W	012	0642	359
196	364500006G01	MONTGOMERY CO	MONTGOMERY CO	OHIO	39 47 47 N	084 07 33 W	016	0800	71
197	364500007G01	MONTGOMERY CO	MONTGOMERY CO	OHIO	39 47 47 N	084 07 29 W	016	0800	86
198	365320002G02	PAINESVILLE	LAKE CO	OHIO	41 43 18 N	081 14 34 W	035	0725	344
199	365580001H01	PORTAGE CO	PORTAGE CO	OHIO	41 10 56 N	081 19 49 W	016	1076	337
200	365640001G01	PREBLE CO	PREBLE CO	OHIO	39 50 21 N	084 43 29 W	013	1205	324
201	366400009H01	STARK CO	STARK CO	OHIO	40 52 15 N	081 19 51 W	025	1110	146
202	366420012I01	STEUBENVILLE	JEFFERSON CO	OHIO	40 21 36 N	080 37 20 W	030	0810	349
203	366600006H09	TOLEDO	LUCAS CO	OHIO	41 38 50 N	083 31 48 W	020	0580	233
204	366600081H09	TOLEDO	LUCAS CO	OHIO	41 43 05 N	083 28 33 W	023	0580	269
205	367280003H01	WESTLAKE	CUYAHOGA CO	OHIO	41 28 31 N	081 52 46 W	025	0660	308
206	367760007I01	YOUNGSTOWN	MAHONING CO	OHIO	41 05 59 N	080 39 01 W	049	0855	335
207	390100067G01	ALLEGHENY CO	ALLEGHENY CO	PENNSYLVANIA	40 22 55 N	080 11 08 W	030	1160	87
208	390100067J05	ALLEGHENY CO	ALLEGHENY CO	PENNSYLVANIA	40 27 25 N	079 44 57 W	013	0121	64
209	390100080J05	ALLEGHENY CO	ALLEGHENY CO	PENNSYLVANIA	40 26 12 N	079 57 16 W	033	0090	91
210	390100081J05	ALLEGHENY CO	ALLEGHENY CO	PENNSYLVANIA	40 26 34 N	079 59 17 W	033	0089	73

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
211	390100088G01	ALLEGHENY CO	ALLEGHENY CO	PENNSYLVANIA	40 28 20 N	079 49 12 W	025	1200	89
212	390120003F01	ALLENTOWN	LEHIGH CO	PENNSYLVANIA	40 37 07 N	075 26 56 W	007	0364	334
213	390140801F01	ALTOONA	BLAIR CO	PENNSYLVANIA	40 30 54 N	078 23 20 W	020	1240	66
214	390440004F01	BADEN	BEAVER CO	PENNSYLVANIA	40 38 08 N	080 13 51 W	012	0710	332
215	390580014F01	BEAVER FALLS	BEAVER CO	PENNSYLVANIA	40 44 52 N	080 19 00 W	012	0710	354
216	390780017F01	BETHLEHEM	NORTHAMPTON CO	PENNSYLVANIA	40 37 13 N	075 21 52 W	007	0300	337
217	390940001G01	BRACKENRIDGE	ALLEGHENY CO	PENNSYLVANIA	40 36 39 N	079 44 47 W	020	0965	101
218	391080012F01	BRISTOL (BOROUGH)	BUCKS CO	PENNSYLVANIA	40 14 51 N	074 59 50 W	007	0040	315
219	391260501F01	BUTLER CO	BUTLER CO	PENNSYLVANIA	41 00 19 N	079 43 02 W	011	1320	111
220	391400100F01	CARBONDALE	LACKAWANNA CO	PENNSYLVANIA	41 34 14 N	075 30 41 W	010	1050	62
221	391560005F01	CHARLEROI	WASHINGTON CO	PENNSYLVANIA	40 08 49 N	079 54 12 W	007	0760	325
222	391620002F01	CHESTER	DELAWARE CO	PENNSYLVANIA	39 50 13 N	075 22 29 W	007	0010	311
223	392020101F01	CONSHOHOCKEN	MONTGOMERY CO	PENNSYLVANIA	40 04 19 N	075 17 57 W	011	0051	67
224	392420501F01	DONORA	WASHINGTON CO	PENNSYLVANIA	40 09 54 N	079 51 45 W	007	0740	40
225	392720100F01	EASTON	NORTHAMPTON CO	PENNSYLVANIA	40 40 34 N	075 13 03 W	010	0300	90
226	393060010F01	ERIE	ERIE CO	PENNSYLVANIA	42 08 22 N	080 02 48 W	012	0660	322
227	393280102F01	FOLCROFT	DELAWARE CO	PENNSYLVANIA	39 53 20 N	075 16 35 W	011	0035	59
228	393880401F01	HARRISBURG	DAUPHIN CO	PENNSYLVANIA	40 14 40 N	076 50 50 W	011	0411	355
229	394460011F01	JOHNSTOWN	CAMBRIA CO	PENNSYLVANIA	40 18 35 N	078 54 57 W	012	1185	345
230	394620001F05	KUTZTOWN	BERKS CO	PENNSYLVANIA	40 30 39 N	075 47 07 W	011	0490	177
231	394660007F01	LANCASTER CITY	LANCASTER CO	PENNSYLVANIA	40 02 38 N	076 17 12 W	012	0325	329
232	395740501F01	MIDLAND	BEAVER CO	PENNSYLVANIA	40 38 06 N	080 26 40 W	007	0840	300
233	396440015F01	NEW CASTLE	LAWRENCE CO	PENNSYLVANIA	40 59 45 N	080 20 48 W	013	0860	332
234	396540013F01	NORRISTOWN	MONTGOMERY CO	PENNSYLVANIA	40 06 44 N	075 18 38 W	012	0175	342
235	396560701F01	NORTHAMPTON	NORTHAMPTON CO	PENNSYLVANIA	40 41 18 N	075 29 31 W	011	0368	16
236	397120301F03	PERRY CO	PERRY CO	PENNSYLVANIA	40 27 28 N	077 10 03 W	007	0415	224
237	397140004H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	40 00 32 N	075 05 54 W	017	0072	88
238	397140014H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	40 03 00 N	075 14 26 W	013	0364	314
239	397140019H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	39 58 43 N	075 05 51 W	013	0023	91
240	397140021H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	39 53 05 N	075 13 51 W	013	0024	87
241	397140023H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	39 54 28 N	075 08 58 W	013	0023	87
242	397140024H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	40 04 35 N	075 00 43 W	013	0093	359
243	397140026H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	39 56 48 N	075 09 50 W	013	0053	91
244	397140027H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	40 00 38 N	075 09 07 W	015	0131	91
245	397140029H01	PHILADELPHIA	PHILADELPHIA CO	PENNSYLVANIA	39 57 26 N	075 10 23 W	050	0081	86

N	SAROAD ID	SITE LOCATION	COUNTY	STATE	LATITUDE	LONGITUDE	PROBE HEIGHT AGL (FT)	PROBE HEIGHT MSL (FT)	#DAYS
246	397260008G01	PITTSBURGH	ALLEGHENY CO	PENNSYLVANIA	40 27 56 N	079 57 40 W	020	0840	217
247	397620009F01	READING	BERKS CO	PENNSYLVANIA	40 19 15 N	075 55 40 W	012	0265	352
248	398040006F01	SCRANTON	LACKAWANNA CO	PENNSYLVANIA	41 26 35 N	075 37 28 W	012	0825	324
249	398040100A05	SCRANTON	LACKAWANNA CO	PENNSYLVANIA	41 35 30 N	076 04 20 W	022	0962	72
250	399430016F01	WILKES-BARRE	LUZERNE CO	PENNSYLVANIA	41 14 16 N	075 53 45 W	012	0540	338
251	399470402F01	WILLIAMSPORT	LYCOMING CO	PENNSYLVANIA	41 14 47 N	076 59 25 W	025	0532	72
252	399470403F01	WILLIAMSPORT	LYCOMING CO	PENNSYLVANIA	41 14 47 N	076 59 27 W	025	0536	70
253	399560008F01	YORK	YORK CO	PENNSYLVANIA	39 57 56 N	076 41 57 W	007	0410	344
254	410140002F03	KENT CO	KENT CO	RHODE ISLAND	41 36 54 N	071 43 11 W	010	0220	328
255	410300012F01	PROVIDENCE	PROVIDENCE CO	RHODE ISLAND	41 49 32 N	071 24 19 W	080	0180	340
256	470140003F01	BURLINGTON	CHITTENDEN CO	VERMONT	44 28 44 N	073 12 43 W	012	0210	50
257	470140006F01	BURLINGTON	CHITTENDEN CO	VERMONT	44 27 24 N	073 11 38 W	012	0330	84
258	470265101A08	GREEN MOUNTAIN NAT.	ADDISON CO	VERMONT	43 56 00 N	073 02 00 W	015		355
259	480080009H01	ALEXANDRIA		VIRGINIA	38 48 38 N	077 02 41 W	036	0076	133
260	480200020G01	ARLINGTON CO	ARLINGTON CO	VIRGINIA	38 51 27 N	077 03 33 W	023	0562	137
261	480710008F01	CHESAPEAKE		VIRGINIA	36 40 00 N	076 19 42 W	013	0028	158
262	480720004F05	CHESTERFIELD CO	CHESTERFIELD CO	VIRGINIA	37 21 26 N	077 35 38 W	015	0252	156
263	481040005G01	FAIRFAX		VIRGINIA	38 50 41 N	077 18 43 W	012	0410	240
264	481060018G01	FAIRFAX CO	FAIRFAX CO	VIRGINIA	38 44 32 N	077 04 39 W	012	0037	150
265	481240003F01	FREDERICKSBURG		VIRGINIA	38 18 40 N	077 26 56 W	010	0200	165
266	481440004F01	HAMPTON		VIRGINIA	37 00 05 N	076 23 57 W	013	0021	262
267	481460001F01	HANOVER CO	HANOVER CO	VIRGINIA	37 38 10 N	077 19 01 W	015	0195	110
268	481500012F01	HENRICO CO	HENRICO CO	VIRGINIA	37 35 25 N	077 30 09 W	020	0220	361
269	481850001G01	MC LEAN	FAIRFAX CO	VIRGINIA	38 56 04 N	077 11 54 W	012	0347	306
270	481920005F01	MARION	SMYTH CO	VIRGINIA	36 50 49 N	081 30 21 W	014	2330	180
271	482180001F01	NORTHAMPTON CO	NORTHAMPTON CO	VIRGINIA	37 17 23 N	075 58 18 W	018	0038	178
272	482520008F05	PRINCE WILLIAM CO	PRINCE WILLIAM	VIRGINIA	38 38 31 N	077 31 42 W			83
273	482800001G01	SALEM		VIRGINIA	37 17 15 N	080 03 10 W	015	1130	360
274	482870004G01	SEVEN CORNERS	FAIRFAX CO	VIRGINIA	38 52 04 N	077 08 35 W	037	0360	263
275	483240007F01	VIRGINIA BEACH		VIRGINIA	36 53 27 N	076 11 02 W	012	0034	338
276	500280004F01	CHARLESTON	KANAWHA CO	WEST VIRGINIA	38 20 31 N	081 37 12 W	020	0605	321
277	500580002A05	GREENBRIER CO	GREENBRIER CO	WEST VIRGINIA	37 46 27 N	080 20 00 W	014	2315	72
278	502120007F01	WHEELING	OHIO CO	WEST VIRGINIA	40 07 14 N	080 41 59 W	035	0695	344
279	090020017I01	WASHINGTON	WASHINGTON	DIST. OF COLUMBIA	38 54 13 N	077 03 06 W	034	0066	244
280	090020025I01	WASHINGTON	WASHINGTON	DIST. OF COLUMBIA	38 58 31 N	077 01 22 W	034	0300	285

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE EPA REGIONAL OXIDANT MODEL: DESCRIPTION AND EVALUATION PLAN		5. REPORT DATE 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Kenneth L. Schere and Allan J. Fabrick		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS (same as 12.)		10. PROGRAM ELEMENT NO. CDWA1A/02-4021 (FY-85) 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS ATMOSPHERIC SCIENCES RESEARCH LABORATORY - RTP, NC OFFICE OF RESEARCH AND DEVELOPMENT U.S. ENVIRONMENTAL PROTECTION AGENCY RESEARCH TRIANGLE PARK, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED In-house 14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The U.S. EPA Regional Oxidant Model (ROM) and NEROS data base are described. The model incorporates a comprehensive description of the physical and chemical processes thought to be important to tropospheric O_3 production on 1000 km scales. The data base employed for the first application of the ROM was collected during the summers of 1979 and 1980 in the Northeast U.S. It contains meteorological and air quality data from regular monitoring networks and from enhanced networks or special field project measurements made during that period. The evaluation procedure that will be used to determine the ROM performance on this data base is outlined. A number of episodes will be simulated from the period July 23 through August 16, 1980, for which performance statistics will be developed. The evaluation of any given day within an episode will proceed in two distinct stages. The first stage will focus on model performance for an individual model realization, irrespective of all other realizations. Model realizations for a given day are functions of the possible flow fields that existed for the day. The second stage will attempt to evaluate model performance using the full probabilistic abilities of the ROM that consider all realizations concurrently. The focus of the evaluation will be on O_3. The exact pathway through the evaluation study will be determined by the resources available at the time.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE