CONTRACT NO. A032-189
FINAL REPORT
DECEMBER 1993

Regional Source-Receptor Relationships for Atmospheric Acidity and Acid Deposition in California

CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

AIR RESOURCES BOARD
Research Division

REGIONAL SOURCE-RECEPTOR RELATIONSHIPS FOR ATMOSPHERIC ACIDITY AND ACID DEPOSITION IN CALIFORNIA

Final Report
Contract No. A032-189

Prepared for

California Air Resources Board
Research Division
2020 L Street
Sacramento, California 95814

LIBRARY-AIR RESOURCES BOARD

Prepared by

Prakash Karamchandani, Christodoulos Pilinis,* and Jitendra Shah†

ENSR Consulting and Engineering
1220 Avenida Acaso
Camarillo, California 93012

*Environmental Quality Laboratory
California Institute of Technology
Pasadena, California 91125

†G2 Environmental, Inc.
3015 Oregon Knolls Drive
Washington, DC 20015

DECEMBER 1993

CONTENTS

ABSTRACT vii
EXECUTIVE SUMMARY ix
1.0 INTRODUCTION 1-1
1.1 Background 1-1
1.2 Objectives of the Study 1-2
1.3 Organization of Report 1-2

2.0 CADMP DRY AND WET DEPOSITION DATA MANAGEMENT 2-1
2.1 Technical Approach 2-1
2.2 Data Gathering and Merging 2-2
2.2.1 Wet Deposition Data 2-2
2.2.2 Dry Deposition Data 2-5
2.3 Data Validation 2-7
2.3.1 Wet Deposition Data Validation 2-14
2.3.2 Dry Deposition Data Validation 2-14
2.4 Data Processing and Distribution 2-19
2.4.1 Wet Deposition Data Processing 2-22
2.4.2 Dry Deposition Data Processing 2-26

3.0 MODEL DESCRIPTION 3-1
3.1 Summary of Model Features 3-1
3.2 Model Formulation 3-2
3.2.1 Governing Equations for Sulfur 3-5
3.2.2 Governing Equations for Nitrogen 3-10
3.3 Calculation of Dry and Wet Deposition 3-10
3.3.1 Dry Deposition 3-10
3.3.2 Wet Deposition 3-11
3.4 Advantages and Disadvantages of the Model 3-12
3.5 Performance History of the Model 3-13

4.0 PREPARATION OF MODEL INPUTS 4-1
4.1 Emissions 4-1
4.2 Upper Air Winds 4-4
4.3 Rainfall Amounts 4-7
4.4 Dry and Wet Deposition Data 4-7

CONTENTS
(Cont'd)

5.0 EVALUATION OF MODEL WITH CADMP DATA 5-1
5.1 Model Parameters 5-1
5.2 Evaluation with Wet Deposition Data 5-3
5.3 Evaluation with Dry Deposition Data 5-15

6.0 SOURCE-RECEPTOR RELATIONSHIPS 6-1

7.0 SENSITIVITY STUDIES 7-1
7.1 Uncertainty Analysis of Model Parameters 7-1
7.1.1 The Latin Hypercube Sampling (LHS) Technique 7-2
7.1.2 Results from Application of LHS 7-7
7.2 Emission Scenario Studies 7-16

8.0 CONCLUSIONS 8-1
8.1 Summary 8-1
8.2 Conclusions and Recommendations 8-2
8.2.1 Model Evaluation Results 8-2
8.2.2 Source-Receptor Relationships 8-3
8.2.3 Sensitivity Studies 8-5

9.0 REFERENCES 9-1

APPENDICES

A ANNUAL AVERAGES FOR DRY AND WET DEPOSITION DATA BY SITE AND YEAR
B SEASONAL SOURCE-RECEPTOR RELATIONSHIPS

CONTENTS
(Cont'd)

LIST OF TABLES

2-1	Structure of the Wet Deposition Database File (WETDEP.DBF)	2-4
2-2	CADMP Database Structure for Ambient Chemical Concentrations (File CPCONnn.DBF)	2-8
2-3	Intermediate Calculation and Validation Flag Definitions for WETDEP.DBF	2-15
2-4	Number of Samples with Validation Flags in WETDEP.DBF	2-16
2-5	Outlier Criteria and Flags in CPCON and CPVAL Files	2-18
2-6	Correlation Coefficients for CPCONO1 Dataset - All Samples Included	2-20
2-7	Correlation Coefficients for CPCONO1 Dataset - Two Sigmas Outliers Excluded	2-21
2-8	Average $PM_{2.5}/PM_{10}$ Concentration Ratios	2-22
2-9	Structure of the Original and Modified Wet Deposition Databases	2-24
2-10	Structure of the File Containing Average Wet Deposition Databases	2-25
2-11	Structure of the Original and Modified Dry Deposition Databases	2-27
4-1	The 17 California Zones Used in the Modeling Study	4-6
5-1	Optimized Parameter Values	5-3
5-2	Model Performance Statistics for Nitrate Concentrations in Rain	5-6
5-3	Model Performance Statistics for Sulfur Concentrations in Rain	5-10
5-4	Results from Linear Regression of $(C_{SO4,\ obs} - C_{SO4,\ est})$ Against C_{Ca}	5-12
6-1	Zones of Influence for the Composite Year (km)	6-3
7-1	Range of Variation of the Five Primary Parameters	7-7
7-2	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) of Parameter Uncertainty Study Results for Annually Averaged Sulfur Concentrations in Rain	7-9
7-3	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) of Parameter Uncertainty Study Results for Annually Averaged Ambient Sulfate Concentrations	7-10
7-4	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) of Parameter Uncertainty Study Results for Annually Averaged Nitrate Concentrations in Rain	7-11

CONTENTS
(Cont'd)

7-5	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) of Parameter Uncertainty Study Results for Annually Averaged Ambient Nitrate Concentrations	7-12
7-6	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) for Source Region Contributions (%) to Annually Averaged Total Sulfur Deposition at Selected Receptor Locations	7-13
7-7	Summary Statistics (Maximum, Minimum, Average, and Standard Deviation) for Source Region Contributions (%) to Annually Averaged Total Nitrate Deposition at Selected Receptor Locations	7-17
7-8	Response of Annual Sulfur Dry Deposition to Changes in SO_x Emissions	7-20
7-9	Response of Annual Sulfur Wet Deposition to Changes in SO_x Emissions	7-21
7-10	Response of Annual Total Sulfur Deposition to Changes in SO_x Emissions	7-22
7-11	Response of Annual Nitrogen Dry Deposition to Changes in NO_x Emissions	7-24
7-12	Response of Annual Nitrogen Wet Deposition to Changes in NO_x Emissions	7-25
7-13	Response of Annual Total Nitrogen Deposition to changes in NO_x Emissions	7-26

CONTENTS
(Cont'd)

LIST OF FIGURES

2-1	Location of CADMP Wet Deposition Monitoring Sites	2-3
2-2	Location of CADMP Dry Deposition Monitoring Sites	2-6
3-1	Idealization of Dispersion of a Plume in the Planetary Boundary Layer	3-4
3-2	Probability Interaction Diagram for Atmospheric Sulfur	3-6
3-3	Probability Interaction Diagram for Atmospheric Nitrogen	3-7
4-1	1985 NO_x Emissions by County	4-2
4-2	1985 SO_x Emissions by County	4-3
4-3	Division of California into 17 Meteorological Zones	4-5
5-1	Comparison of Annually Averaged Observations of Nitrate Concentrations in Rain with Model Estimates for Six Years (1984-1989)	5-5
5-2	Comparison of Estimated Sulfur Concentrations in Rain with Observed Concentrations (with and without Sea Salt Correction) for the Years 1984 and 1985	5-8
5-3	Comparison of Annually Averaged Observations of Non-Sea Salt Concentrations in Rain with Model Estimates for Six Years (1984-1989)	5-9
5-4	Comparison of Seasonally Averaged Observations of Nitrate Concentrations in Rain with Model Estimates for the Winter, Spring, and Fall Seasons of the Years 1985 and 1988	5-13
5-5	Comparison of Seasonally Averaged Observatiosn of Non-Sea Salt Sulfur Concentrations in Rain with Model Estimates for the Winter, Spring, and Fall Seasons of the Years 1985 and 1988	5-14
5-6	Comparison of Observed Non-Sea Salt Concentrations in Rain with Model Estimates Before and After Correction for the Soil Dust Contribution at Six Locations	5-16
5-7	Comparison of Annually Averaged Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for 1988	5-17
5-8	Comparison of Annually Averaged Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for 1989	5-18
5-9	Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Summer 1988	5-19
5-10	Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Fall 1988	5-20

CONTENTS
(Cont'd)

5-11 Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Winter 1989 5-21

5-12 Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Spring 1989 5-22

5-13 Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Summer 1989 5-24

6-1 Contribution (%) of the 17 Source Regions to Annual Total Sulfur (SO_2 + Sulfate) Dry Deposition at Various Receptor Locations 6-4

6-2 Contribution (%) of the 17 Source Regions to Annual Total Sulfur (SO_2 + Sulfate) Wet Deposition at Various Receptor Locations 6-7

6-3 Contribution (%) of the 17 Source Regions to Annual Total Sulfur Deposition at Various Receptor Locations 6-10

6-4 Contribution (%) of the 17 Source Regions to Annual Total Nitrogen (NOx + Nitrate) Dry Deposition at Various Receptor Locations 6-12

6-5 Contribution (%) of the 17 Source Regions to Annual Total Nitrogen (NO_x + Nitrate) Wet Deposition at Various Receptor Locations 6-14

6-6 Contribution (%) of the 17 Source Regions to Annual Total Nitrogen Deposition at Various Receptor Locations 6-16

7-1 Intervals Used with an LHS of Size n = 5 in Terms of the Density Function and Cumulative Distribution Function for a Normal Random Variable 7-4

7-2 Intervals Used with an LHS of Size n = 5 in Terms of the Density Function and Cumulative Distribution Function for a Uniform Random Variable 7-5

7-3 A Two-Dimensional Representation of One Possible LHS of Size 5 Utilizing X_1 and X_2 7-6

ABSTRACT

This report describes the results of a database management and semi-empirical modeling study that was performed to evaluate regional source-receptor relationships (SRRs) for atmospheric acidity and acidic deposition in California. The objectives of the study were to quantify the contributions of the various source regions in California to acidic deposition at selected receptors in the state and to estimate the uncertainties in the derived values.

The study consisted of the following tasks:

- acquiring precipitation chemistry and dry deposition databases from the California Acid Deposition Monitoring Program (CADMP), as well as meteorological and emissions databases for the years of interest (1984 through 1989);
- incorporation of the CADMP data into a dBase IV Database Management System (DBMS) and limited validation of the data;
- preparation of annually and seasonally averaged modeling databases for ENSR's semi-empirical statistical acid deposition model (STATMOD);
- adaptation of STATMOD for application to California;
- application of STATMOD to compute annually and seasonally averaged precipitation and ambient concentrations of sulfate and nitrate and evaluation and optimization of the model using CADMP data;
- computation of source-receptor relationships for selected receptors;
- uncertainty analysis of the derived source-receptor relationships, based on the uncertainty of the model parameters; and
- sensitivity analysis of SO_2 and NO_x emission controls on acidic deposition at sensitive receptor locations.

The model performed well in estimating nitrate concentrations in rain and air, but sulfate concentrations in rain and air at several receptors were consistently underestimated. Total sulfur (SO_2 + sulfate) concentrations were also underestimated, suggesting that the SO_2 emissions used in the simulations were either inaccurate or incomplete. Assuming that there were no major inaccuracies in the anthropogenic SO_2 emissions, this indicated that the model underestimated sulfur concentrations in rain and air because SO_2 emissions from non-anthropogenic sources (e.g., sea-salt, or wind blown soil dust) were not included in the inventory. A correction for the sea-salt contribution improved the performance of the model. However, the model still tended to underestimate the observed sulfur concentrations. A statistical analysis using observed calcium concentrations in rain was performed to demonstrate that model performance could be improved further by estimating the contribution of wind blown soil dust to sulfur concentrations in rain, and SO_2 and sulfate concentrations in air.

The source-receptor relationships for a composite year and composite seasons were developed in the form of plots showing the relative contribution of the various source regions in California to acidic deposition at selected receptor locations, as well as zones of influence for each receptor. Ten receptors from various urban, rural, and remote regions of the state

were selected to create a representative sample for the analysis. The source-receptor analysis showed that the contribution of local sources dominated the total sulfur (SO_2 + sulfate) dry deposition at receptors located in urban regions or close to large sources. The zone of influence for total sulfur dry deposition at these receptors was usually less than 100 km, i.e., dry sulfur deposition at these receptors could be attributed mainly to sources that were less than 100 km away from the receptors. At remote receptors, there was some evidence of long-range transport in the dry deposition of sulfur. The zone of influence for these receptors was of the order of 200 km. Long-range transport played a larger role in total sulfur wet deposition as compared to dry deposition, even at receptors in the vicinity of large sources. The zone of influence for sulfur wet deposition was higher than that for sulfur dry deposition at all receptors. At some receptors, the wet deposition zone of influence was more than twice the dry deposition zone of influence. Long-range transport was found to be a factor in nitrogen dry deposition, even at receptors that were located close to high NO_x sources. The zones of influence for dry deposition of nitrate were always larger than the zones of influence for sulfur dry deposition, indicating that nitrate dry deposition has a more regional nature than sulfur dry deposition. The zones of influence for the wet deposition of sulfur and nitrate were comparable. The total (dry + wet) nitrogen deposition was larger than the total sulfur deposition at most receptor locations. These results are consistent with the relative NO_x and SO_x emission rates in California.

Sensitivity studies were conducted to determine the sensitivity of the derived source-receptor relationships to changes in the model parameters within their expected range of values, and to determine the response of the model to changes in NO_x and SO_x emissions. The Latin Hypercube Sampling (LHS) technique was used to construct 100 parameter sets. The simulations were performed for a composite year. The sensitivity studies showed that the estimated sulfur concentrations in rain were relatively insensitive to changes in the model parameters. The standard deviation was less than 10 percent for all the receptors. The estimated ambient sulfate concentration was more sensitive to changes in the model parameters than sulfur in rain. However, the standard deviation of the estimated values was still small (less than 15%) for all the receptors. Nitrate concentrations in both rain and air were more sensitive to the input parameters. For many of the receptors, the maximum estimated concentrations were about two times larger than the minimum estimated concentrations. The relative contributions of the source regions to total sulfur and nitrate deposition at various receptors were found to be insensitive to the changes in STATMOD parameters.

The effect of five different emission control scenarios on acidic deposition in California was examined. The results of the emission control scenario studies showed that a 50 percent reduction in statewide SO_x emissions did not always result in a 50 percent reduction in the dry, wet, or total sulfur deposition. On the other hand, the response to a 50 percent reduction in NO_x emissions across the entire state was linear. Emission controls in individual source regions showed that larger reductions were estimated at receptors that were directly influenced by these source regions as compared to receptors that did not receive large contributions from these sources. The results were slightly different for sulfur and nitrogen, particularly at receptors located close to the source regions where emissions were reduced in our simulations. The differences in the results for sulfur and nitrogen can be attributed to the differences in their zones of influence for dry deposition, and the fact that wet scavenging of sulfur is limited by the availability of oxidants.

EXECUTIVE SUMMARY

This report describes the results of a database management and semi-empirical modeling study that was performed to evaluate regional source-receptor relationships (SRRs) for atmospheric acidity and acidic deposition in California. The study was motivated by the need to determine the contribution of various source regions in California (e.g., the San Francisco Bay Area and the South Coast Air Basin) to acidic deposition at sensitive receptor locations in urban and remote areas.

The objectives of the study were to quantify the contributions of the various source regions in California to acidic deposition at selected receptors in the state and to estimate the uncertainties in the derived values. The following tasks were performed to accomplish these objectives:

- acquire the California Acid Deposition Monitoring Program (CADMP) precipitation chemistry and dry deposition databases and incorporate the data into a dBase IV Database Management System (DBMS) that allows the manipulation of the data (e.g., to compute annual or seasonal averages);

- acquire meteorological and emissions databases for the years of interest ;

- adapt ENSR's semi-empirical statistical acid deposition model (STATMOD) for application to California;

- apply STATMOD to compute annually and seasonally averaged precipitation and ambient concentrations of sulfate and nitrate and use CADMP data to evaluate and optimize the model;

- compute source-receptor relationships for selected receptors;

- perform an uncertainty analysis of the derived source-receptor relationships, based on the uncertainty of the model parameters; and

- determine the impact of SO_2 and NO_x emission controls on acidic deposition at sensitive receptor locations.

CADMP DATABASE MANAGEMENT

The database management efforts were designed to facilitate data gathering, validation, and data exchange with interested users (e.g., modelers, data analysts). The objectives of the CADMP dry and wet deposition data management task were to:

- acquire the relevant precipitation chemistry and air quality databases to calibrate and evaluate the semi-empirical model;
- incorporate the databases into a Database Management System (DBMS);
- perform a limited validation of these databases; and
- develop programs to generate seasonal and yearly averages of the data to compare against model outputs.

The CADMP dry and wet deposition databases were pooled together into a relational database using dBase IV, a commercially available database management system.

Once the data were obtained, data consistency tests based on the known physical relationships among the observed variables were applied. The data were screened for unusual and inconsistent values, although they were expected to have already gone through two levels of validation. For this study, the data were processed through the third level of validation known as "Level III validation." Questionable data were not edited — instead, they were flagged in the database created for this study.

The validated dry and wet deposition data sets were assimilated and programs were developed to process the data to obtain seasonally or annually averaged precipitation chemistry and air quality input files in the format required by STATMOD.

SEMI-EMPIRICAL MODEL

The model that was used in this study is a nonlinear semi-empirical long-range transport model, referred to as STATMOD. The model is based on the premise that long term (annual or seasonal) averages of concentration and deposition of a pollutant are insensitive to short term fluctuations of the governing processes such as meteorology and chemistry. This assumption allows us to base model estimates on the statistics of the governing processes.

STATMOD incorporates parameterizations for the transport, chemistry, and scavenging (dry and wet deposition) of sulfur and nitrogen species. Transport is treated using straight-line trajectories governed by large scale wind roses derived from upper air winds. To account for the spatial variation in the transport winds, the region of interest is divided into several subregions, each of which is assigned its own wind rose based on upper air wind measurements.

The model uses a stochastic approach in which sulfur (or nitrogen) is assigned to four states: wet and dry primary species (SO_2 or NO_x), and wet and dry secondary species (sulfate or nitrate). Four differential equations govern the temporal evolution of these four states. The transition of species between wet and dry periods is governed by two time scales, which are the average lengths of wet and dry periods.

The dry deposition of sulfur and nitrogen is treated by specifying dry deposition velocities. The wet removal of SO_2, sulfate and nitrate is parameterized in terms of a scavenging rate that is proportional to the rainfall amount at the receptor of interest.

The gas-phase conversion from SO_2 to sulfate (or NO_x to nitrate) is parameterized in terms of first-order conversion rates for wet and dry periods. An oxidant-limited wet scavenging rate of SO_2 in the wet state is used for treating the aqueous-phase oxidation of SO_2 in precipitating clouds, and an oxidant-limited SO_2 oxidation rate in the dry state is used for oxidation in non-precipitating clouds.

All the parameters of the model have physical meaning, and therefore can be assigned values on the basis of our understanding of the relevant processes. However, in the actual application of the model, some of the parameter values are adjusted to obtain an optimum fit between model estimates and the observations being analyzed.

PREPARATION OF MODEL INPUTS

The application of STATMOD required an emission inventory for SO_x and NO_x, upper air winds, and ambient air quality and precipitation chemistry data to evaluate the model and optimize model parameters. These data were required for 1984 through 1989, the period of interest for this study.

The ARB annual summary of emissions for each county in California were used to prepare the STATMOD emission files. Upper air wind speed and direction data for seven sites were obtained from the National Climatic Data Center (NCDC). Upper air data for Fresno were

provided by the ARB. The NCDC and ARB upper air data were supplemented with historical summaries of California upper air data compiled by the ARB.

Annual and seasonal wind roses for the years 1984 through 1989 were constructed for 17 meteorological zones using these data. In addition, wind roses for the composite year (average of all years) and composite seasons were also developed.

The total rainfall amount in each zone was also estimated by interpolating the CADMP precipitation amounts to a grid covering the modeling domain, using the "Kriging" technique described in Venkatram (1988).

Annually and seasonally averaged air quality and precipitation chemistry files for STATMOD were prepared using the CADMP dry and wet deposition dBase IV files developed for the study.

EVALUATION OF MODEL WITH CADMP DATA

STATMOD was applied to California and the model parameters were optimized by comparing model estimates with the annually and seasonally averaged CADMP data. We spent significantly more time evaluating and optimizing the model using the CADMP wet deposition database than with the dry deposition database because the wet deposition database had more years and more sites of data, and was more reliable than the dry deposition database.

Sensitivity studies with uniform winds and the actual wind roses showed that meteorology played a very important role in the performance of the model. We also found an improvement in model performance when we made small adjustments to the wind roses in meteorological zones where the only source of upper air data was the ARB historical summary.

Although the model performed well in estimating nitrate concentrations in rain and air, it was found that sulfate concentrations in rain and air at several receptors were consistently underestimated. Total sulfur (SO_2 + sulfate) concentrations were also underestimated, suggesting that the SO_2 emissions used in the simulations were either inaccurate or incomplete. We assumed that there were no major inaccuracies in the anthropogenic SO_2 emissions, and that the model underestimated sulfur concentrations in rain and air because SO_2 emissions from non-anthropogenic sources (e.g., sea-salt, or wind blown soil dust) were not included in the inventory.

We first attempted to estimate the sea-salt contribution to the observed sulfur concentrations in rain. This was done by using the observed sodium and magnesium concentrations and the chemical composition of sea water. This correction resulted in some improvement in the performance of the model, particularly at coastal receptors. However, there were still several receptors where the model underestimated the observed sulfur concentrations, by an average of about 30 percent. The largest underestimations occurred at Gasquet, S. Lake Tahoe, Bakersfield, Lakeport, and San Rafael.

We then performed a regression analysis of the residual of the seasonally averaged observed and estimated sulfur concentrations in rain against observed calcium concentrations for those receptors where the model performed poorly. The analysis showed that the residuals and observed calcium concentrations were highly correlated and that model performance could be improved by estimating the contribution of wind blown soil dust to sulfur concentrations in rain, and SO_2 and sulfate concentrations in air.

SOURCE-RECEPTOR RELATIONSHIPS

STATMOD was used to assess the contribution of the seventeen meteorological zones to the anthropogenic sulfur and nitrate deposition at various receptor locations. The simulations were performed for a composite year and for composite seasons to minimize the effect of year-to-year fluctuations in meteorology and emissions. The source-receptor relationships are presented as:

- plots showing the relative contribution of the various source regions in California to acid deposition at selected receptor locations; and

- zones of influence for each receptor — the zone of influence is a measure of the area around a given receptor location that predominantly influences the estimated deposition at the location.

We selected the following ten receptors from various urban, rural, and remote regions of the state to create a representative sample for our analysis: Bakersfield, Santa Barbara, Yosemite, Pasadena, Sacramento, San Jose, Sequoia, S. Lake Tahoe, Montague, and Escondido.

The source-receptor analysis showed that the contribution of local sources dominated the total sulfur (SO_2 + sulfate) dry deposition at receptors located in urban regions or close to large sources. This is expected because sulfur dry deposition in these regions is dominated by locally emitted SO_2. The zone of influence for total sulfur dry deposition at these receptors

was usually less than 100 km, i.e., dry sulfur deposition at these receptors could be attributed mainly to sources that were less than 100 km away from the receptors.

On the other hand, for remote receptors, such as Yosemite, there was some evidence of long-range transport in the dry deposition of sulfur. Yosemite had a zone of influence for sulfur dry deposition of 200 km and the largest contributors to the annual total sulfur dry deposition at Yosemite were the Upper San Joaquin Valley and the San Francisco Bay Area.

Long-range transport played a larger role in total sulfur wet deposition as compared to dry deposition, even at receptors in the vicinity of large sources. The zone of influence for sulfur wet deposition was higher than that for sulfur dry deposition at all receptors. At some receptors, the wet deposition zone of influence was more than twice the dry deposition zone of influence.

The annual total sulfur dry deposition was estimated to be larger than the sulfur wet deposition at receptors located near large source regions. In some cases, such as Bakersfield, sulfur dry deposition was more than an order of magnitude larger than the wet deposition. At some remote receptors, the dry and wet deposition of sulfur were comparable, or the wet deposition was larger than the dry deposition.

The differences between dry and wet sulfur deposition were attributed to the following factors:

- the dry deposition rate of SO_2 is about an order of magnitude larger than that for sulfate — thus, sulfur dry deposition is dominated by deposition of SO_2, and tends to be of local origin; on the other hand, sulfur wet deposition is controlled by the rate at which SO_2 is oxidized to sulfate — thus, a significant fraction of the sulfur in rain can be attributed to SO_2 emissions that have been transported over long distances; and

- wet deposition is a function of precipitation patterns, and, given the same amount of SO_2 to sulfate conversion, will be higher in regions experiencing large rainfall amounts than in regions experiencing small rainfall amounts.

Long-range transport was found to be a factor in nitrogen dry deposition, even at receptors that were located close to high NO_x sources. The zones of influence for dry deposition of nitrate were always larger than the zones of influence for sulfur dry deposition, indicating that nitrate dry deposition has a more regional nature than sulfur dry deposition. The reason for this is that nitrogen dry deposition is dominated by dry deposition of nitrate, since NO_x is not dry deposited efficiently. Thus, significant dry deposition of nitrogen oxides can only occur after the NO_x has been converted to nitrate.

The source contributions to nitrate wet deposition were generally similar to the sulfate wet deposition source contributions. The zones of influence for the wet deposition of the sulfur and nitrate were also comparable.

The total (dry + wet) nitrogen deposition was larger than the total sulfur deposition at most receptor locations. These results are consistent with the relative NO_x and SO_x emission rates in California.

SENSITIVITY STUDIES

Studies were conducted to determine the sensitivity of the derived source-receptor relationships to changes in the model parameters within their expected range of values, and to determine the response of the model to changes in NO_x and SO_x emissions.

Because it was not possible to examine the sensitivity of the model to all its parameters, the uncertainty analysis focused on a few important parameters. These parameters were the mixing height, the duration of dry and wet periods, the gas-phase oxidation rates of SO_2 and NO_x for dry and wet conditions, the SO_2 oxidation rate in non-precipitating clouds, and the concentration of the aqueous-phase SO_2 oxidant, H_2O_2. Some of these parameters were expected to be correlated with one another. For these parameters, we varied one independent parameter, and allowed the related parameters to vary proportionately. This reduced the number of primary parameters to five. Table ES-1 shows the range of variation of the five primary parameters.

TABLE ES-1

Range of Variation of the Five Primary Parameters

Parameter	Base Case Value	Range of Variation
z_i (m)	600	400 to 800
τ_d (hr)	100	80 to 120
τ_w (hr)	10	5 to 15
k_{dSO2} (%/hr)	1	0.5 to 1.5
k'_d (%/hr)	2	1 to 3

We used a constrained sampling scheme, referred to as the Latin Hypercube Sampling (LHS) technique, to construct 100 sets of the five primary parameters. After the sets were constructed, we applied STATMOD 100 times for sulfur and nitrogen. To eliminate the influence of year-to-year variations in model inputs (e.g., upper air winds), we performed the simulations for a composite year.

The outputs from these simulations were used to perform a statistical analysis of the estimated ambient concentrations of sulfate and nitrate aerosol and sulfur and nitrate concentrations in rain at each receptor site. A similar analysis was performed for the contribution of each of the seventeen zones to the deposition of acidic species at various receptors.

The sensitivity studies showed that the estimated sulfur concentrations in rain were relatively insensitive to changes in the model parameters. The standard deviation was less than 10 percent for all the receptors. The estimated ambient sulfate concentration was more sensitive to changes in the model parameters than sulfur in rain. However, the standard deviation of the estimated values was still small (less than 15%) for all the receptors.

Nitrate concentrations in both rain and air were more sensitive to the input parameters. For many of the receptors, the maximum estimated concentrations were about two times larger than the minimum estimated concentrations. The sensitivity was more pronounced for receptors located in urban areas, such as Southern California, that were directly affected by local sources.

The relative contribution of the 17 meteorological zones to total sulfur and nitrate deposition at various receptors was found to be insensitive to the changes in STATMOD parameters. For the 10 receptors selected for the analysis, the major contributing source regions and their relative ranks remained the same, regardless of the parameter set used.

We examined the effect of five different emission control scenarios on acidic deposition in California. The first scenario corresponded to a 50 percent reduction in the emissions of both SO_x and NO_x across the entire state. The other four scenarios corresponded to 50 percent reductions of SO_x and NO_x emissions in individual source regions. The four source regions selected for the latter analysis were the South Coast Air Basin, the Lower San Joaquin Valley, the San Francisco Bay Area and the Upper South Central Coast.

The results of the emission control scenario studies showed that a 50 percent reduction in statewide SO_x emissions did not always result in a 50 percent reduction in the dry, wet, or total sulfur deposition. On the other hand, the response to a 50 percent reduction in NO_x

emissions across the entire state was linear. These results are consistent with the formulation of the model which is non-linear for sulfur, but linear for nitrogen.

Emission controls in individual source regions showed that larger reductions were estimated at receptors that were directly influenced by these source regions as compared to receptors that did not receive large contributions from these sources. The results were slightly different for sulfur and nitrogen, particularly at receptors located close to the source regions where emissions were reduced in our simulations. The differences in the results for sulfur and nitrogen can be attributed to the differences in their zones of influence for dry deposition, and the fact that wet scavenging of sulfur is limited by the availability of oxidants.

1.0 INTRODUCTION

This report describes the results of a database management and semi-empirical modeling study that was performed to evaluate regional source-receptor relationships (SRRs) for atmospheric acidity and acidic deposition in California. The study was performed by ENSR Consulting and Engineering and its subcontractors and sponsored by the Research Division of the California Air Resources Board (CARB). As discussed below, the study was motivated by the need to determine the contribution of various source regions in California (e.g., the San Francisco Bay Area and the South Coast Air Basin) to acidic deposition at sensitive receptor locations in urban and remote areas.

1.1 Background

There is increasing concern about the potential effects of acid deposition in California. Acidic deposition is widespread in California — its occurrence has been documented in the San Francisco Bay Area, the San Joaquin Valley, the South Coast Air Basin, the Sierra Nevada Mountains and elsewhere. Concentrations of nitrate and sulfate in rain are of the order of 20 $\mu eq\ \ell^{-1}$ in regions downwind of urban areas in California, typically three to five times higher than background concentrations in remote areas of the state. While these concentration levels are generally lower than those found in precipitation in the eastern United States, several sensitive receptor locations, such as the high elevation lakes in the Sierra Nevada Mountains, have little alkalinity to buffer acidic inputs. Thus, even relatively small levels of acidic ions in precipitation can have a harmful effect on these sensitive receptors.

Because of this concern, the California Legislature adopted the Atmospheric Acidity Protection Act in 1988 to continue and expand the Air Resources Board's research and monitoring program established in 1982. The Act established the Atmospheric Acidity Protection Program of 1988 (AAPP), which requires the Air Resources Board to determine the relationships between sources of acidic pollutants and the deposition of these species in receptor areas.

In addition, in 1984, the Air Resources Board established a program, called the California Acid Deposition Monitoring Program (CADMP), to monitor a variety of species that are relevant to acidic deposition. The program consists of a precipitation chemistry network, a dry deposition network, and a fog monitoring network. The precipitation network consists of 35 samplers distributed throughout the state and provides measurements of concentrations in precipitation of sulfate, nitrate, chloride, ammonium, sodium, magnesium, potassium,

calcium, pH, and conductivity. The dry deposition network consists of 10 sites that provide measurements of sulfate, nitrate, chloride, ammonium, sodium, magnesium, potassium, and calcium in aerosols, as well as concentrations of sulfur dioxide, nitrogen oxides, ammonia, and nitric acid in the gas phase. This database can be used by models in order to derive source-receptor relationships.

1.2 Objectives of the Study

The overall objectives of this study were to quantify the contributions of selected source regions to acidic deposition at selected receptors and to estimate the uncertainties in the derived values. The following tasks were performed to accomplish these objectives:

- acquire the CADMP precipitation chemistry and dry deposition databases and incorporate the data into a dBase IV Database Management System (DBMS) that allows the manipulation of the data (e.g., to compute annual or seasonal averages);
- acquire meteorological and emissions databases for the years of interest;
- adapt ENSR's semi-empirical statistical acid deposition model (STATMOD) for application to California;
- apply STATMOD to compute annually and seasonally averaged precipitation and ambient concentrations of sulfate and nitrate and compare model results with observations;
- compute source-receptor matrices for selected receptors;
- perform an uncertainty analysis of the derived source-receptor relationships, based on the uncertainty of the input parameters; and
- determine the impact of SO_2 and NO_x emission controls on acidic deposition at sensitive receptor locations.

1.3 Organization of Report

Section 2 of the report describes the management of the CADMP dry and wet deposition databases. Section 3 describes the semi-empirical model, STATMOD, that was used in our analysis, while Section 4 describes the preparation of the seasonal and annual databases required for the application and evaluation of the model in California. Section 5 discusses the

application and calibration of the model parameters using the CADMP database. Sections 6 and 7 present the source-receptor relationships and the results of the uncertainty and emission control sensitivity studies, respectively. Finally, Section 8 summarizes the results and conclusions from the study.

2.0 CADMP DRY AND WET DEPOSITION DATA MANAGEMENT

The database management efforts were designed to facilitate data gathering, validation, and data exchange with interested users (e.g., modelers, data analysts). Traditional data validation and exchange protocols were modified and implemented for the study.

The objectives of the dry and wet deposition data management task were to:

- acquire the relevant precipitation chemistry and air quality databases to calibrate and evaluate the semi-empirical model;
- incorporate the databases into a Database Management System (DBMS);
- perform a limited validation of these databases; and
- develop programs to generate seasonal and yearly averages of the data to compare against model outputs.

In the following section, we describe our technical approach to performing the database management task, and the various steps that were taken to accomplish the above objectives.

2.1 Technical Approach

We obtained the CADMP precipitation chemistry and dry deposition databases from the ARB and the Desert Research Institute (DRI). Dr. Blanchard, a contractor to ARB on a related project, was also a useful source of data and information.

The databases were pooled together into a relational database using dBase IV. dBase IV is a relational database management system that is commercially available for implementation on microcomputers. It can handle 255 fields of 4,000 characters per record and up to one billion records per file. dBase IV can be implemented on most IBM PC-compatible desk top computers. dBase IV files can also be read directly by a variety of popular statistical, plotting, database, and spreadsheet programs.

Our original plan was to create a menu system in dBase IV as a framework to build upon during the project. However, it was discovered that an overall system would not be as useful because of the diverse nature of the databases. Since our focus was to obtain averages of

the wet and dry deposition data, it was more efficient to develop these programs to run on their own. The menu format created can be developed further, if needed.

The various tasks that were performed as part of the data management effort for this study were:

- data gathering and merging;
- data assessment and validation; and
- data processing and distribution.

Each of these tasks is described in more detail below. Due to wet and dry deposition data quality issues, the data assessment task was only partially performed.

2.2 Data Gathering and Merging

2.2.1 Wet Deposition Data

Weekly precipitation samples, using Aerochem Metric Model 301 samplers, have been collected since 1983 at the 35 CADMP sites shown in Figure 2-1. The wet-side bucket samples are generally collected on Tuesday and are packed in blue ice and transported within 24 hours of collection. Field measurements at the time of sample collection include rain gauge, sample volume, pH, and specific conductance readings. Conductance and pH are measured and ion analyses are performed on all samples at the laboratory. Chloride, nitrate, and sulfate ion concentrations are determined by ion chromatography while sodium, potassium, calcium, and magnesium ion concentrations are determined by atomic absorption. Ammonium ion concentrations are determined by automated colorimetry. A detailed description of the CADMP precipitation chemistry network is available in CARB (1988).

The wet deposition database was first obtained from the ARB and subsequently from Dr. Blanchard, a contractor to ARB who had already integrated the most recent information (CARB, 1988; Blanchard, 1992). Collocated datasets for wet deposition at four locations (Sacramento, Montague, Sequoia - Giant Forest, and Tanbark Flats) were also obtained from Dr. Blanchard and processed for the data validation/assessment task. Table 2-1 describes the wet deposition database structure, the species included, and their units. Note that fields 21 to 30 were added for this project and are not part of the original CADMP database.

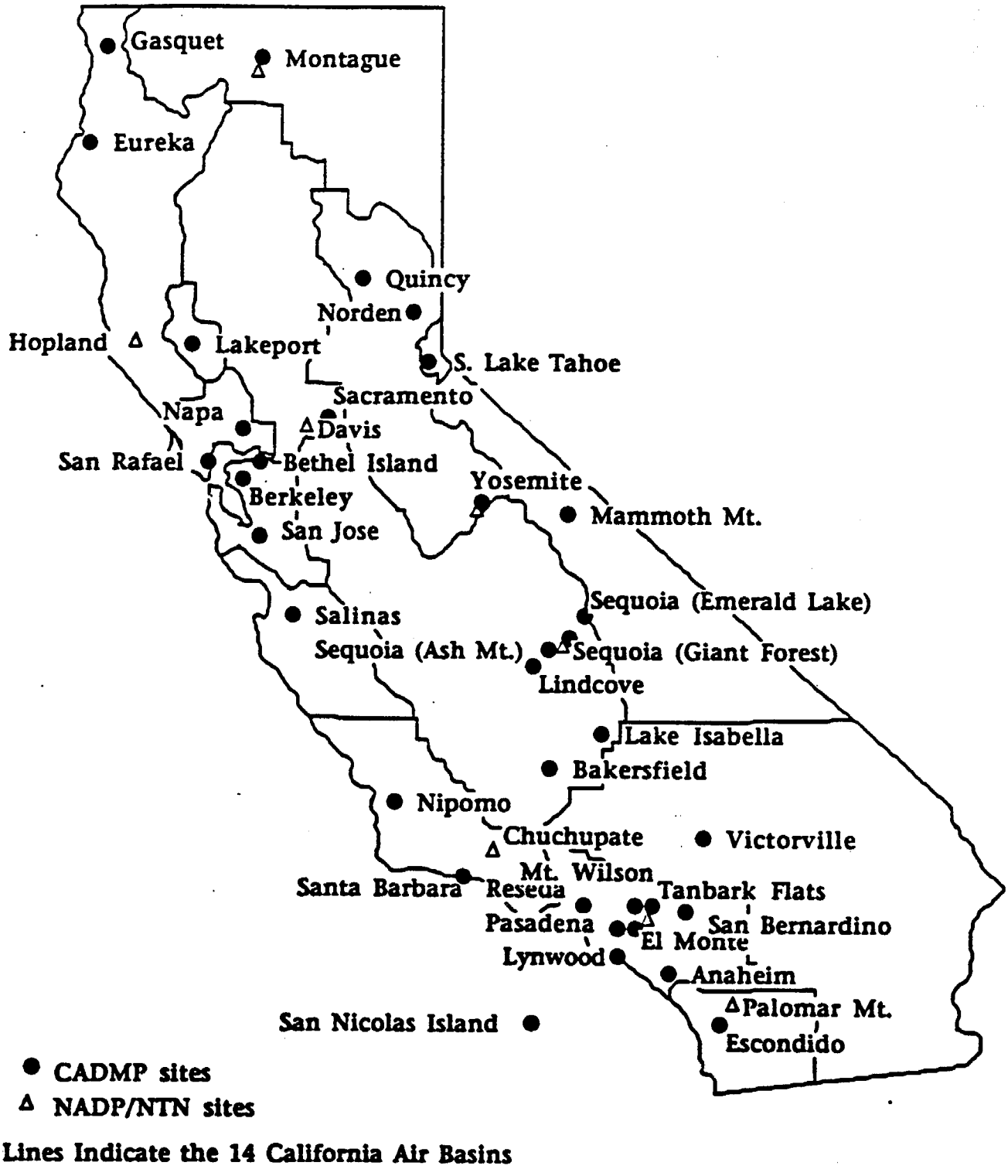


FIGURE 2-1. Location of CADMP Wet Deposition Monitoring Sites

TABLE 2-1

Structure of the Wet Deposition Database File (WETDEP.DBF)

Number of data records: 9,487 Date of last update: 10/26/92

Field	Field Name	Type	Width	Dec	Index	Definitions
1	SITE	Numeric	7		N	Site code
2	START	Numeric	6		N	Starting date, YYMMDD
3	STOP	Numeric	6		N	Stopping date, YYMMDD
4	RAIN_GAUGE	Numeric	5	2	N	Rain gauge in inches
5	VOL	Numeric	5		N	Volume of rain in ml
6	PH_F	Numeric	4	2	N	Field pH
7	PH_L	Numeric	4	2	N	Laboratory pH
8	COND_F	Numeric	5	1	N	Field conductivity, µs/cm
9	COND_L	Numeric	5	1	N	Laboratory conductivity, µs/cm
10	NA	Numeric	6	3	N	Concentrations in mg/l
11	K	Numeric	6	3	N	Concentrations in mg/l
12	CA	Numeric	5	3	N	Concentrations in mg/l
13	MG	Numeric	5	3	N	Concentrations in mg/l
14	NH4	Numeric	5	3	N	Concentrations in mg/l
15	CL	Numeric	6	2	N	Concentrations in mg/l
16	NO3	Numeric	6	2	N	Concentrations in mg/l
17	SO4	Numeric	6	2	N	Concentrations in mg/l
18	R1	Character	1		N	Remark column 1
19	R2	Character	1		N	Remark column 2
20	R3	Character	1		N	Remark column 3
21	H	Numeric	10	6	N	See Table 2-3
22	CATION_EQ	Numeric	6	1	N	See Table 2-3
23	ANION_EQ	Numeric	6	1	N	See Table 2-3
24	COND_PRED	Numeric	5	1	N	See Table 2-3
25	EQ_FLAG	Character	1		N	See Table 2-3

TABLE 2-1 (Cont'd)

Structure of the Wet Deposition Database File (WETDEP.DBF)

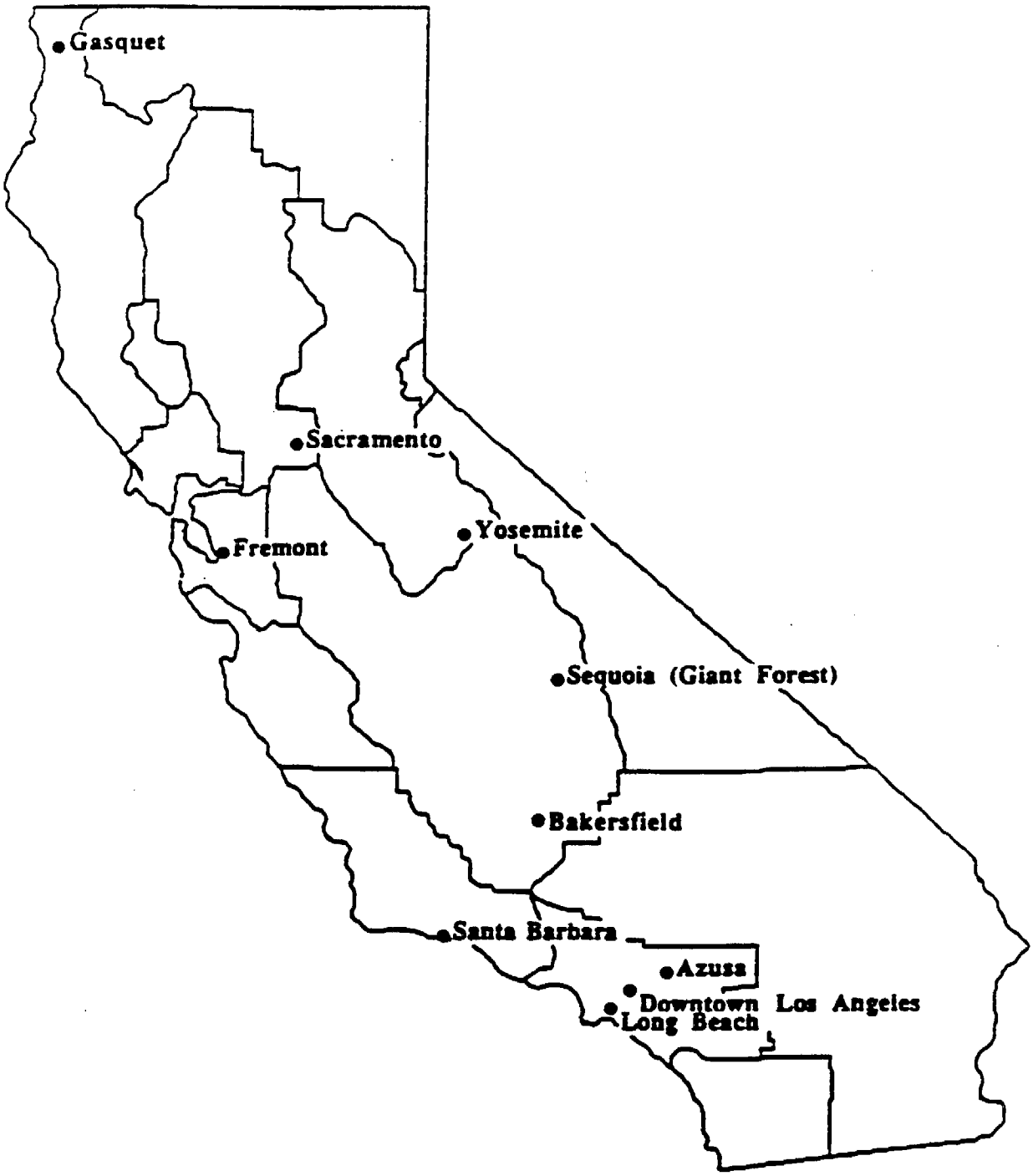
Number of data records: 9,487 Date of last update: 10/26/92

Field	Field Name	Type	Width	Dec	Index	Definitions
26	PCND_FLAG	Character	1		N	See Table 2-3
27	PH_FLAG	Character	1		N	See Table 2-3
28	COND_FLAG	Character	1		N	See Table 2-3
29	VOL_FLAG	Character	1		N	See Table 2-3
30	DUR_FLAG	Character	1		N	See Table 2-3
		TOTAL	129			

2.2.2 Dry Deposition Data

The CADMP dry deposition samples were collected at 10 sites, shown in Figure 2-2, with a specially constructed sampler that collects particles and selected gases on four different filter packs. The four different filter packs are designated as TK, DN, TN, and GT. The TK filter pack consists of a membrane Teflon filter for collecting PM_{10} particles, followed by a citric acid impregnated filter for collecting NH_3 gas, followed by a potassium carbonate impregnated filter for collecting SO_2 gas. The DN filter pack consists of a nylon filter placed downstream of a nitric acid denuder for collecting total particulate NO_3^-. The TN filter pack consists of a Teflon membrane filter for collecting $PM_{2.5}$ particles followed by a nylon filter for collecting volatilized particle NO_3^- plus nitric acid. The GT filter pack consists of a glass fiber filter to remove particles, followed by a triethyleneamine (TEA) impregnated filter for collecting NO_2.

Daytime (0600–1800) and nighttime (1800–0600) samples were collected every sixth day following the EPA-recommended schedule. $PM_{2.5}$ and PM_{10} particles were analyzed for mass by gravimetry; Cl^-, NO_3^-, and $SO_4^=$ were analyzed by ion chromatography; NH_4^+ by automated colorimetry; and Na^+, Mg^{++}, K^+, and Ca^{++} by atomic absorption spectrophotometry. HNO_3 was determined as the sum of the Teflon NO_3^- plus the backup nylon NO_3^-, minus the denuded nylon NO_3^-, times a NO_3^- to HNO_3 molecular formula conversion factor.



Lines Indicate the 14 California Air Basins

FIGURE 2-2. Location of CADMP Dry Deposition Monitoring Sites

SO_2, NH_3, and NO_2 were determined by analyzing the potassium carbonate impregnated filters for $SO_4^=$, the citric acid impregnated filters for NH_4^+, and the TEA impregnated filters for NO_3^-, respectively. The ions were converted to their gaseous species equivalent by the appropriate molecular formula conversion constant.

Samples were analyzed in batches consisting of between one month's and one quarter's worth of samples. The concentration database was constructed in a series of steps using dBase IV programs designed for the study, as summarized below. Files of mass and each chemical species were made with concentration and concentration uncertainty in μg/filter. Concentration uncertainty was determined for mass and each chemical species as the standard deviation of differences between original measurements and replicate measurements. Field data, and mass and chemical data were merged into a single file for each batch of samples, with concentrations in μg/filter. A separate file of average and standard deviation of field blank values was constructed. A file of concentrations in $\mu g/m^3$ was constructed by subtracting the field blank values from the concentrations in μg/filter and dividing by sampling volume. A file of validation ratios, consisting of ratios between various species was also made. Concentration uncertainties were propagated through each calculation step.

Data from May 1988 through September 1989 were gathered and processed by DRI, and were distributed as CPCON01.DBF and CPVAL01.DBF files. Subsequent data processing was conducted by the ARB, and data from October 1989 through June 1991 were distributed as CPCON02.DBF and CPVAL02.DBF files.

Table 2-2 describes the dry deposition database structure (including fields added for this project), species included, and their units. All missing or invalid data are represented by a -99 located in place of the value. Detailed descriptions of the data monitoring and chemical analysis are available in DRI (1991).

2.3 Data Validation

Data validation is one of the most important functions of data management. Data validation is the process whereby data are filtered and accepted or rejected based on a set of criteria (U.S. EPA, 1976). To create a usable data set for the analyst, data sets are processed to eliminate invalid data and to supply flags and documentation for the remainder to facilitate further analysis and judgement. The data reliability may depend on the extent or level of validation a data set has undergone.

TABLE 2-2

CADMP Database Structure for Ambient Chemical Concentrations (File CPCONnn.DBF)

Field	Name	Data Type	Width	Explanation
1	IDGT	Character	10	Sample ID, GT filter pack
2	IDTK	Character	10	Sample ID, TK filter pack
3	IDDN	Character	10	Sample ID, DN filter pack
4	IDTN	Character	10	Sample ID, TN filter pack
5	SITE	Character	2	Sampling site
6	SAMPID	Numeric	2	Sampler ID number
7	DATE	Date	8	Sampling date
8	STRTIM	Numeric	4	Sample start time, HH:MM
9	PERIOD	Character	1	Sample period D=day, N=night
10	FGTTFLG	Character	5	Field flag, GT filter pack
11	FTKTFLG	Character	5	Field flag, TK filter pack
12	FDNFFLG	Character	5	Field flag, DN filter pack
13	FTNFFLG	Character	5	Field flag, TN filter pack
14	MAGTFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter mass
15	MAGTTF	Character	5	Analysis flag, PM_{10} Teflon filter mass
16	NAATFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter Na^+
17	NAATTF	Character	5	Analysis flag, PM_{10} Teflon filter Na^+
18	MGATFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter Mg^{++}
19	MGATTF	Character	5	Analysis flag, PM_{10} Teflon filter Mg^{++}
20	KPATFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter K^+
21	KPATTF	Character	5	Analysis flag, PM_{10} Teflon filter K^+
22	CAATFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter Ca^{++}
23	CAATTF	Character	5	Analysis flag, PM_{10} Teflon filter Ca^{++}
24	N4CTTF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter NH_4^+
25	N4CTTF	Character	5	Analysis flag, PM_{10} Teflon filter NH_4^+
26	ANITFF	Character	5	Analysis flag, $PM_{2.5}$ Teflon filter Cl^-, NO_3^-, $SO_4^=$

TABLE 2-2 (Cont'd)

CADMP Database Structure for Ambient Chemical Concentrations (File CPCONnn.DBF)

Field	Name	Data Type	Width	Explanation
27	ANITTF	Character	5	Analysis flag, PM_{10} Teflon filter Cl^-, NO_3^-, $SO_4^=$
28	N3CNFF	Character	5	Analysis flag, nylon filter NO_3^-
29	HNCNGF	Character	5	Analysis flag, nylon filter HNO_3
30	NHCCGF	Character	5	Analysis flag, citric acid filter NH_3
31	SOIKGF	Character	5	Analysis flag, K_2CO_3 filter SO_2
32	NOCEGF	Character	5	Analysis flag, TEA filter NO_2
33	VOLGTC	Numeric	6.2	Volume, m^3, GT filter pack
34	VOLGTU	Numeric	6.2	Volume uncertainty, m^3, GT filter pack
35	VOLTKC	Numeric	6.2	Volume, m^3, TK filter pack
36	VOLTKU	Numeric	6.2	Volume uncertainty, m^3, TK filter pack
37	VOLDNC	Numeric	6.2	Volume, m^3, DN filter pack
38	VOLDNU	Numeric	6.2	Volume uncertainty, m^3, DN filter pack
39	VOLTNC	Numeric	6.2	Volume, m^3, TN filter pack
40	VOLTNU	Numeric	6.2	Volume uncertainty, m^3 TN filter pack
41	MAGTFC	Numeric	10.4	$PM_{2.5}$ Teflon filter mass, $\mu g/m^3$
42	MAGTFU	Numeric	10.4	$PM_{2.5}$ Teflon filter mass uncertainty, $\mu g/m^3$
43	MAGTTC	Numeric	10.4	PM_{10} Teflon filter mass, $\mu g/m^3$
44	MAGTTU	Numeric	10.4	PM_{10} Teflon filter mass uncertainty, $\mu g/m^3$
45	NAATFC	Numeric	10.4	$PM_{2.5}$ Teflon filter Na^+, $\mu g/m^3$
46	NAATFU	Numeric	10.4	$PM_{2.5}$ Teflon filter Na^+ filter uncertainty, $\mu g/m^3$
47	NAATTC	Numeric	10.4	PM_{10} Teflon filter Na^+, $\mu g/m^3$
48	NAATTU	Numeric	10.4	PM_{10} Teflon filter Na^+ filter uncertainty, $\mu g/m^3$
49	MGATFC	Numeric	10.4	$PM_{2.5}$ Teflon filter Mg^{++}, $\mu g/m^3$
50	MGATFU	Numeric	10.4	$PM_{2.5}$ Teflon filter Mg^{++} filter uncertainty, $\mu g/m^3$
51	MGATTC	Numeric	10.4	PM_{10} Teflon filter Mg^{++}, $\mu g/m^3$
52	MGATTU	Numeric	10.4	PM_{10} Teflon filter Mg^{++} filter uncertainty, $\mu g/m^3$

TABLE 2-2 (Cont'd)

CADMP Database Structure for Ambient Chemical Concentrations (File CPCONnn.DBF)

Field	Name	Data Type	Width	Explanation
53	KPATFC	Numeric	10.4	$PM_{2.5}$ Teflon filter K^+, $\mu g/m^3$
54	KPATTU	Numeric	10.4	$PM_{2.5}$ Teflon filter K^+ filter uncertainty, $\mu g/m^3$
55	KPATTC	Numeric	10.4	PM_{10} Teflon filter K^+, $\mu g/m^3$
56	KPATTU	Numeric	10.4	PM_{10} Teflon filter K^+ filter uncertainty, $\mu g/m^3$
57	CAATFC	Numeric	10.4	$PM_{2.5}$ Teflon filter Ca^{++}, $\mu g/m^3$
58	CAATFU	Numeric	10.4	$PM_{2.5}$ Teflon filter Ca^{++} filter uncertainty, $\mu g/m^3$
59	CAATTC	Numeric	10.4	PM_{10} Teflon filter Ca^{++}, $\mu g/m^3$
60	CAATTU	Numeric	10.4	PM_{10} Teflon filter Ca^{++} filter uncertainty, $\mu g/m^3$
61	N4CTFC	Numeric	10.4	$PM_{2.5}$ Teflon filter NH_4^+, $\mu g/m^3$
62	N4CTTC	Numeric	10.4	$PM_{2.5}$ Teflon filter NH_4^+ filter uncertainty, $\mu g/m^3$
63	N4CTTC	Numeric	10.4	PM_{10} Teflon filter NH_4^+, $\mu g/m^3$
64	N4CTTC	Numeric	10.4	PM_{10} Teflon filter NH_4^+ filter uncertainty, $\mu g/m^3$
65	CLITFC	Numeric	10.4	$PM_{2.5}$ Teflon filter Cl^-, $\mu g/m^3$
66	CLITFU	Numeric	10.4	$PM_{2.5}$ Teflon filter Cl^- filter uncertainty, $\mu g/m^3$
67	CLITTC	Numeric	10.4	PM_{10} Teflon filter Cl^-, $\mu g/m^3$
68	CLITTU	Numeric	10.4	PM_{10} Teflon filter Cl^- filter uncertainty, $\mu g/m^3$
69	N3ITFC	Numeric	10.4	$PM_{2.5}$ Teflon filter NO_3^-, $\mu g/m^3$
70	N3ITFU	Numeric	10.4	$PM_{2.5}$ Teflon filter NO_3^- filter uncertainty, $\mu g/m^3$
71	N3ITTC	Numeric	10.4	PM_{10} Teflon filter NO_3^-, $\mu g/m^3$
72	N3ITTU	Numeric	10.4	PM_{10} Teflon filter NO_3^- filter uncertainty, $\mu g/m^3$
73	S4ITFC	Numeric	10.4	$PM_{2.5}$ Teflon filter $SO_4^=$, $\mu g/m^3$
74	S4ITFU	Numeric	10.4	$PM_{2.5}$ Teflon filter $SO_4^=$ filter uncertainty, $\mu g/m^3$
75	S4ITTC	Numeric	10.4	PM_{10} Teflon filter $SO_4^=$, $\mu g/m^3$
76	S4ITTU	Numeric	10.4	PM_{10} Teflon filter $SO_4^=$ filter uncertainty, $\mu g/m^3$
77	N3CNFC	Numeric	10.4	Nylon filter NO_3^-, $\mu g/m^3$
78	N3CNFU	Numeric	10.4	Nylon filter NO_3^- uncertainty, $\mu g/m^3$

TABLE 2-2 (Cont'd)

CADMP Database Structure for Ambient Chemical Concentrations (File CPCONnn.DBF)

Field	Name	Data Type	Width	Explanation
79	HNCNGC	Numeric	10.4	Nylon filter HNO_3, $\mu g/m^3$
80	HNCNGU	Numeric	10.4	Nylon filter HNO_3 uncertainty, $\mu g/m^3$
81	NHCCGC	Numeric	10.4	Citric acid filter NH_3, $\mu g/m^3$
82	NHCCGU	Numeric	10.4	Citric acid filter NH_3 uncertainty, $\mu g/m^3$
83	SOIKGC	Numeric	10.4	K_2CO_3 filter SO_2, $\mu g/m^3$
84	SOIKGU	Numeric	10.4	K_2CO_3 filter SO_2 uncertainty, $\mu g/m^3$
85	NOCEGC	Numeric	10.4	TEA filter NO_2, $\mu g/m^3$
86	NOCEGU	Numeric	10.4	TEA filter NO_2 uncertainty, $\mu g/m^3$
87	HNDDGC	Numeric	10.4	Denuder difference HNO_3, $\mu g/m^3$
88	HNDDGU	Numeric	10.4	Denuder difference HNO_3 uncertainty, $\mu g/m^3$
89	SUMIFC	Numeric	10.4	Sum of $PM_{2.5}$ chemical concentrations, $\mu g/m^3$
90	SUMIFU	Numeric	10.4	Sum uncertainty of $PM_{2.5}$ chemical concentrations, $\mu g/m^3$
91	SUMITC	Numeric	10.4	Sum of PM_{10} chemical concentrations, $\mu g/m^3$
92	SUMITU	Numeric	10.4	Sum uncertainty of PM_{10} chemical concentrations, $\mu g/m^3$
93	NOTE	Character	60	Comments
94	VALFLAGS	Numeric	3	Summary validation flag
95	MAGTRFMAGTRF	Character	2	See Table 2-5 for flag definition
96	NAATRF	Character	2	See Table 2-5 for flag definition
97	MGATRF	Character	2	See Table 2-5 for flag definition
98	KPATRF	Character	2	See Table 2-5 for flag definition
99	CAATRF	Character	2	See Table 2-5 for flag definition
100	N4CTRF	Character	2	See Table 2-5 for flag definition
101	CLITRF	Character	2	See Table 2-5 for flag definition
102	N3ITRF	Character	2	See Table 2-5 for flag definition
103	S4ITRF	Character	2	See Table 2-5 for flag definition
104	IONRFF	Character	2	See Table 2-5 for flag definition

TABLE 2-2 (Cont'd)

CADMP Database Structure for
Ambient Chemical Concentrations (File CPCONnn.DBF)

Field	Name	Data Type	Width	Explanation
105	IONRTF	Character	2	See Table 2-5 for flag definition
106	CHBALFF	Character	2	See Table 2-5 for flag definition
107	CHBALTF	Character	2	See Table 2-5 for flag definition
108	N4PRFF	Character	2	See Table 2-5 for flag definition
109	N4PRTF	Character	2	See Table 2-5 for flag definition
110	HNO3F	Character	2	See Table 2-5 for flag definition
111	DNO3PM10RF	Character	2	See Table 2-5 for flag definition
112	DNO3TOTLRF	Character	2	See Table 2-5 for flag definition

The traditional validation steps (e.g., Mueller et al., 1983) were modified for this study. Once the data were obtained, data consistency tests based on the known physical relationships among observables were applied. The data were screened for unusual and inconsistent values as discussed below. Most of the data to be included in the database were expected to have already gone through such screening. Any discrepancies found were discussed with ARB staff and the original data gatherer.

The philosophy governing data validation for this study was to minimize data editing. We did not edit questionable data — instead, they were flagged in the database created for this study. Since the number and size of flags are finite and small, additional comments in data validation reports were provided. These reports, which provide supplemental information regarding data reliability, were written during the course of the project and have been previously provided to the ARB Project Manager (G2E, 1992).

Before performing the validation, we spent some time examining the documentation provided by the ARB and DRI. The DRYDEP and MET programs and data files originally provided were thoroughly examined and tested. The methods of processing and validation were examined and confirmed.

All data obtained for this study have been gathered by the ARB or others under previous projects. In general, they should have, at the minimum, gone through two levels of validation

(DRI, 1990). For this study, we processed the data through the third level of validation known as "Level III validation." Some Level III data validation steps will be similar to Level II for redundancy.

The data validation steps that were utilized for this project are discussed below. The problems identified are included in the final database as flags and are briefly discussed below. Details of our analysis have been provided during the course of the project in our quarterly progress reports as well as in memoranda to the ARB Project Manager.

The following validation procedures were conducted on the dry and wet deposition databases. The exact limits or values used for these procedures depended on the data itself, and were determined during the process or upon analysis of the data as discussed later in this section.

Screening analysis: All data that were less than the species precision (used as a lower quantifiable limit) were flagged if not done so already. For each species of interest, the outliers were identified by determining observations greater than the mean plus three times the standard deviation. These outliers were further investigated and flagged as appropriate.

Ratio analysis: All species concentrations that were not flagged were further processed to generate ratios among observables. The upper and lower limits of the ratios were developed using the database.

Time series analysis: Time series plots of the observable of interest at each site were generated using the VOYAGER software package. These plots were used to verify that the screening and ratio analyses described above were implemented correctly in the dataset.

Scatter plot and correlation analysis: Various scatter plots and correlations between species concentrations at a given site and among sites, along with temporal comparison plots, were made to further identify outliers or anomalies. Consistent pattern variations were also identified using time series plots. The other relationships that were tested include the following:

- particle size relationship to ensure that $PM_{2.5}$ measurements were less than PM_{10} (within measurement uncertainty);
- ratio of sum of ions to mass for $PM_{2.5}$;
- ratio of sum of ions to mass for PM_{10};

- charge balance for $PM_{2.5}$;
- charge balance for PM_{10};
- predicted/measured NH_4^+ concentration;
- HNO_3 less than the Minimum Detectable Concentration (MDC);
- daytime vs. nighttime relationship; and
- sampling site differences to check consistency from time to time and from site type to site type.

2.3.1 Wet Deposition Data Validation

The procedure for flagging questionable data in the wet deposition database was developed in consultation with ARB staff and Dr. Charles Blanchard. The dBase file, WETDEP.DBF, was constructed from the ASCII wet deposition files. Missing data values are identified as -99 and concentrations below the detection limit are negative values, with the value being the detection limit. The field names are generally self explanatory, except that PH_F is field measured pH, PH_L is lab measured pH, COND_F is field measured conductivity, COND_L is lab measured conductivity, and R1, R2, and R3 are field data flags as present in the original ASCII files (Blanchard, 1992). All units are the same as in the original ASCII files.

Intermediate concentration fields and validation flag fields were added to WETDEP.DBF (see Table 2-1). Intermediate calculation and validation flag definitions for the database are listed in Table 2-3. The number of samples flagged for each flagged field are shown in Table 2-4.

2.3.2 Dry Deposition Data Validation

In order to provide a general understanding of the two-part dry deposition dataset, files CPCON01 and CPCON02 were explored graphically using the "VOYAGER" software. All of the samples flagged as suspect and obvious outliers were examined. Average concentrations and their standard deviations for each species were calculated separately for the two files. Characteristics such as concentration ranges for each species, trends in concentration, relationships between species, and spatial relationships were examined.

TABLE 2-3

Intermediate Calculation and Validation Flag Definitions for WETDEP.DBF

Field Name	Description	Definition [a, b]
H	H^+ concentration, mg/l	$10^{(-PH_L)} * 1.00794 * 1000$
CATION_EQ	Sum of cation concentrations, µeq/l	1000 * (H/1.00794 + NA/22.98977 + MG/12.1525 + K/39.0983 + CA/20.039 + NH4/18.0385)
ANION_EQ	Sum of anion concentrations, µeq/l	1000 * (CL/35.453 + NO3/62.0049 + SO4/48.0318)
COND_PRED	Predicted conductivity, µs/cm	350 * H/1.00794 + 50.1 * NA/22.98977 + 53.0 * MG/12.1525 + 73.5 * K/39.0983 + 59.5 * CA/20.039 + 73.5 * NH4/18.0385 + 76.3 * CL/35.453 + 71.4 * NO3/62.0049 + 80.0 * SO4/48.0318
EQ_FLAG	Charge balance flag; 'C' = cations high, 'A' = anions high	Flag = 'C' if (CATION_EQ-ANION_EQ) / ((CATION_EQ + ANION_EQ)/2) > 0.3 Flag = 'A' if (ANION_EQ-CATION_EQ) / ((CATION_EQ + ANION_EQ)/2) > 0.3
PCND_FLAG	Predicted conductivity flag; 'H' if predicted high, 'L' if predicted low	Flag = 'H' if COND_PRED > 1.3 * COND_L Flag = 'L' if COND_PRED < 0.7 * COND_L
PH_FLAG	Lab vs. field pH flag; 'F' if field higher than lab by 0.1 pH unit or 5 µeq/l (whichever is less stringent), 'L' if lab higher than field	Flag = 'F' if PH_F - PH_L > 0.1 and $(10^{(-PH_L)} * 10^6) - (10^{(-PH_F)} * 10^6) > 5$ Flag = 'L' if PH_L - PH_F > 0.1 and $(10^{(-PH_F)} * 10^6) - (10^{(-PH_L)} * 10^6) > 5$
COND_FLAG	Lab vs. field conductivity flag; 'F' if field high, 'L' if lab high	Flag = 'F' if (COND_F-COND_L) / ((COND_F + COND_L)/2) > 0.3 Flag = 'L' if (COND_L-COND_F) / ((COND_F + COND_L)/2) > 0.3
VOL_FLAG	Volume flag; 'R' if rain gauge differs from wet bucket, 'L' if volume is low	Flag = 'R' if VOL * 0.0006 / RAIN_GAUGE < 0.7 or VOL * 0.0006 / RAIN_GAUGE > 1.3 Flag = 'L' if VOL < 200 or RAIN_GAUGE < 0.118
DUR_FLAG	Duration flag; 'S' if short sample, 'L' if long sample	Flag = 'S' if STOP - START < 7 Flag = 'L' if STOP - START > 7

a Factors for conversion of mg/l to µeq/l based on gram equivalent weights.

b Factors for conversion of µeq/l to predicted conductivity taken from Kenni James, "1989 Quality Assurance Report, NADP/NTN Deposition Monitoring, Laboratory Operations, Central Analytical laboratory, January 1989 through December 1989," Illinois State Water Survey, April 1991.

TABLE 2-4

Number of Samples with Validation Flags in WETDEP.DBF

Flag Field	Flag	Number of Flags	Number of Data Points
EQ_FLAG	C	1,164	3,625
	A	17	3,625
PCND_FLAG	H	16	3,624
	L	208	3,624
PH_FLAG	F	26	2,185
	L	305	2,185
COND_FLAG	F	133	2,220
	L	25	2,220
VOL_FLAG	R	164	7,264
	L	6,610	7,264
DUR_FLAG	S	642	9,487
	L	563	9,487

Note: See Table 2-3 for flag definition.

Our analysis of the CPCON02 database revealed some inconsistencies in the dry deposition data collected after October 1989. For example, a much larger number of negative HNO_3 concentrations were observed for the CPCON02 dataset than for the CPCON01 dataset. We discussed these inconsistencies with ARB staff and were informed that the version of the CPCON02 database we received contained sample calibration errors of 20 to 100 percent, and should not be used in our analysis. A revalidated version of the database was being prepared under a separate ARB contract. However, we could not use the revised version since it was not available in time for our study. Thus, we only used the CPCON01 dataset for the remainder of our study.

Records with sample durations outside of the prescribed limits were invalidated. Since sampling time is not indicated in the CPCON files, sampling volumes were used. All samples with volumes between 0 and 1/2 nominal in both CPCON01 were already flagged as invalid before we received the data. Nominal sampling volumes are 14.4 m^3 for the TK, DN, and TN

filter packs, and 1.4 m^3 for the GT filter pack. Sampling volumes less than 1/2 nominal or greater than 2 times nominal were selected as limits. These limits are intended to remove samples for which insufficient mass would be collected for an adequate analysis (less than 1/2 normal volume), or which may indicate incorrectly recorded elapsed time meter readings (greater than 2 times normal volume).

The final data validation tasks for the dry deposition data set included identifying and flagging outliers, and calculating correlation coefficients. Several criteria for identifying outliers were developed. These criteria included:

- $PM_{2.5}$ to PM_{10} ratios;
- sum of ions to mass ratio;
- charge balance;
- ratio of predicted to measured NH_4^+ (the calculation of predicted NH_4^+ is described below); and
- negative HNO_3 values.

Outliers were also identified based on tests of certain concentration relationships among species. In each test, the concentration uncertainties were considered in order to avoid classifying samples as outliers when the concentration uncertainties were large with respect to the concentrations. Uncertainties were calculated by propagating concentration uncertainties using normal error propagation methods.

Outliers were classified as one sigma, two sigma, or three sigma outliers based on uncertainty levels. Table 2-5 shows the outlier criteria, the corresponding flags, and the flag field names. The original database files, CPVAL01 and CPCON01, were modified to include these new fields. The factor "n" in each equation in Table 2-5 is 1, 2, or 3, and corresponds to the one, two or three sigma outlier classes. That is, n = 1 corresponds to flag R1, n = 2 corresponds to flag R2, etc.

TABLE 2-5

Outlier Criteria and Flags in CPCON and CPVAL Files

Outlier Criteria [a]	Flags	Flag Fields
$PM_{2.5}/PM_{10}$ ratio > 1 + n * uncertainty	R1, R2, R3	MAGTRF (mass) NAATRF (Na^+) MGATRF (Mg^{++} KPATRF (K^+) CAATRF (Ca^{++}) N4CTRF (NH_4^+) CLITRF (Cl^-) N3ITRF (NO_3^-) S4ITRF ($SO_4^=$)
Sum of ions/total mass > 1 + n * uncertainty	R1, R2, R3	IONRFF ($PM_{2.5}$) IONRTF (PM_{10})
Cation equiv./anion equiv. > 2 + n * uncertainty	H1, H2, H3	CHBALFF ($PM_{2.5}$) CHBALTF (PM_{10})
Cation equiv./anion equiv. < 0.5 - n * uncertainty	L1, L2, L3	CHBALFF ($PM_{2.5}$) CHBALTF (PM_{10})
Predicted/measured NH_4^+ > 1.5 + n * uncertainty	H1, H2, H3	N4PRFF ($PM_{2.5}$) N4PRTF (PM_{10})
Predicted/measured NH_4^+ < 0.67 - n * uncertainty	L1, L2, L3	N4PRFF ($PM_{2.5}$) N4PRTF (PM_{10})
HNO_3 < 0 - n * uncertainty	N1, N2, N3	HNO3F
Denuded $PM_{2.5}$ NO_3^-/PM_{10} NO_3^- > 1 + n * uncertainty	R1, R2, R3	DNO3PM10RF
Denuded/(Tef. + Backup) $PM_{2.5}$ NO_3^- > 1 + n * uncertainty	R1, R2, R3	DNO3TOTLRF
Summary of all flags	FLAGS	

a n = 1, 2, or 3 and corresponds to flags R1, R2, and R3, respectively.

Predicted NH_4^+ was calculated in two ways, first by assuming that the ammonium was associated with ammonium bisulfate (NH_4HSO_4) and ammonium nitrate (NH_4NO_3), and second by assuming association with ammonium sulfate [$(NH_4)_2SO_4$] and NH_4NO_3. Of these two, the value providing a predicted/measured NH_4^+ ratio closest to 1 was used for further analyses. This "best" ratio was then flagged as an outlier if it did not fall between 0.67 and 1.5. These bounds were chosen because they were found empirically to contain over 90 percent of the data points.

Flags were defined for each outlier test criteria and a summary flag for each sample was created. Fields for calculated outlier ratios and ratio uncertainties are included along with the

flags in file CPVAL01. Their field names are the same as the flag field names, except that they end in C (for concentration) and U (for uncertainty) instead of F. Also included in the CPCON and CPVAL files is a summary validation flag field named VALFLAGS. This field is a 3-digit field, "nmm", where the two right-most digits "mm" indicate the number of flags applied to that record, and the left digit "n" denotes the highest sigma of all flags for that record. The VALFLAGS field is 0 for records with no validation flags.

Correlation coefficients were calculated separately for the CPCON01 data set, first using all non-missing data, and again with the two sigma outliers excluded. Tables 2-6 and 2-7 show the results. In general, correlations improved when outliers were excluded. For example, the correlation of $PM_{2.5}$ nitrate with $PM_{2.5}$ ammonium in CPCON01 changed from 0.67 to 0.88. In many cases, there were large differences in correlations between the two data sets — HNO_3 to backup NO_3^- correlations changed from 0.58 to 0.81 in CPCON01.

The ratios of $PM_{2.5}$ to PM_{10} concentrations were calculated again with two sigma outliers removed. Ratios with all samples included and with two sigma outliers excluded are given in Table 2-8. The ratios did not change significantly for the CPCON01 data except for Cl^-.

Concentration statistics including monthly average, maximum, minimum, standard deviation and number of samples were calculated for the CPCON01 data set. Time series plots for each species showing day and night time concentrations separately and for monthly means were prepared.

2.4 Data Processing and Distribution

The validated data sets were assimilated and programs were developed to process the data to obtain seasonally or annually averaged precipitation chemistry and air quality input files in the format required by STATMOD.

The dry and wet deposition databases prepared for distribution to the ARB and other interested users have the following features. Each observable in the database is identified by a field name which follows a pattern for that type of observable. Each measurement method is associated with a separate validation field to document the sample validity for that method. Data contained in different dBase IV files can be linked by indexing on and relating to common attributes in each file. In general, sampling site, sampling date and period, and sampling substrate IDs are the common fields among various data files which can be used to relate data in one file to the corresponding data in another file.

TABLE 2-6

Correlation Coefficients for CPCONO1 Dataset
All Samples Included

	FMass	TMass	FNa+	TNa+	FMg++	TMg++	FK+	TK+	FCa++	TCa++	FNH4+	TNH4+	FCl-	TCl-	FNO3-	TNO3-	FSO4=	TSO4=	DNO3-	BNO3-	NH3	GSO2	GNO2	GHNO3
FMass	1.00																							
TMass	0.88	1.00																						
FNa+	0.10	0.20	1.00																					
TNa+	-0.02	0.19	0.77	1.00																				
FMg++	0.32	0.33	0.63	0.54	1.00																			
TMg++	0.13	0.38	0.78	0.94	0.65	1.00																		
FK+	0.83	0.52	0.17	-0.06	0.42	0.09	1.00																	
TK+	0.23	0.29	0.11	0.09	0.18	0.16	0.33	1.00																
FCa++	0.46	0.31	0.26	-0.05	0.67	0.10	0.43	0.13	1.00															
TCa++	0.56	0.73	0.30	0.21	0.54	0.45	0.36	0.22	0.60	1.00														
FNH4+	0.75	0.72	-0.04	-0.00	0.02	0.09	0.35	0.14	0.04	0.27	1.00													
TNH4+	0.73	0.73	-0.07	-0.00	-0.02	0.06	0.33	0.14	0.01	0.26	0.96	1.00												
FCl-	0.21	0.15	0.18	0.06	0.12	0.06	0.23	0.06	0.01	-0.02	0.18	0.16	1.00											
TCl-	-0.11	-0.02	0.52	0.66	0.34	0.60	-0.04	0.04	-0.06	-0.04	-0.10	-0.09	0.31	1.00										
FNO3-	0.61	0.59	-0.06	-0.09	-0.03	-0.01	0.29	0.10	0.05	0.24	0.65	0.67	0.13	-0.07	1.00									
TNO3-	0.57	0.68	0.06	0.17	0.05	0.24	0.24	0.12	-0.02	0.33	0.65	0.71	0.07	0.00	0.89	1.00								
FSO4=	0.45	0.53	0.23	0.29	0.21	0.39	0.13	0.09	0.06	0.35	0.62	0.57	-0.02	-0.06	0.20	0.37	1.00							
TSO4=	0.42	0.55	0.23	0.35	0.18	0.43	0.10	0.10	0.02	0.36	0.58	0.59	-0.03	-0.01	0.21	0.42	0.89	1.00						
DNO3-	0.74	0.79	-0.01	0.02	0.07	0.14	0.33	0.15	0.10	0.45	0.66	0.66	0.12	-0.11	0.89	0.72	0.45	0.46	1.00					
BNO3-	0.42	0.56	0.19	0.20	0.24	0.36	0.13	0.11	0.16	0.56	0.42	0.36	-0.04	-0.14	0.07	0.23	0.55	0.51	0.56	1.00				
GNH3	0.41	0.50	0.16	0.06	0.34	0.21	0.30	0.15	0.39	0.57	0.26	0.24	-0.06	-0.11	0.22	0.26	0.23	0.24	0.37	0.40	1.00			
GSO2	0.36	0.49	0.15	0.12	0.26	0.27	0.12	0.11	0.24	0.56	0.30	0.29	-0.04	-0.06	0.26	0.33	0.26	0.32	0.36	0.36	0.47	1.00		
GNO2	0.50	0.64	0.13	0.16	0.16	0.32	0.26	0.16	0.14	0.53	0.42	0.43	0.05	0.01	0.45	0.56	0.30	0.36	0.54	0.34	0.39	0.36	1.00	
GHNO3	0.03	0.15	0.19	0.16	0.21	0.27	-0.04	0.02	0.12	0.37	-0.11	-0.17	-0.12	-0.09	0.05	0.11	0.26	0.21	-0.07	0.56	0.21	0.26	0.10	1.00

Note:
Species prefix of "F" denotes PM2.5
Species prefix of "T" denotes PM10
Species prefix of "G" denotes gas
Species prefix of "D" denotes denuded (NO3-)
Species prefix of "B" denotes backup nylon (NO3-)

TABLE 2-7

Correlation Coefficients for CPCONO1 Dataset
Two Sigma Outliers Excluded

	FMass	TMass	FNa+	TNa+	FMg++	TMg++	FK+	TK+	FCa++	TCa++	FNH4+	TNH4+	FCl-	TCl-	FNO3-	TNO3-	FSO4=	TSO4=	DNO3-	BNO3-	NH3	GSO2	GNO2	GHNO3
FMass	1.00																							
TMass	0.93	1.00																						
FNa+	0.04	0.17	1.00																					
TNa+	0.01	0.21	0.79	1.00																				
FMg++	0.12	0.26	0.91	0.72	1.00																			
TMg++	0.15	0.37	0.79	0.95	0.78	1.00																		
FK+	0.63	0.55	0.13	-0.04	0.16	0.04	1.00																	
TK+	0.43	0.46	0.18	0.14	0.21	0.21	0.61	1.00																
FCa++	0.25	0.30	0.22	0.02	0.46	0.17	0.22	0.21	1.00															
TCa++	0.50	0.69	0.26	0.26	0.43	0.47	0.26	0.34	0.65	1.00														
FNH4+	0.78	0.78	0.01	0.06	0.05	0.17	0.39	0.29	0.07	0.40	1.00													
TNH4+	0.78	0.79	-0.00	0.07	0.03	0.17	0.37	0.28	0.04	0.40	0.97	1.00												
FCl-	0.27	0.22	0.19	0.05	0.13	0.09	0.32	0.25	0.03	0.09	0.15	0.13	1.00											
TCl-	-0.04	0.06	0.59	0.72	0.54	0.68	0.01	0.10	-0.01	0.07	-0.07	-0.07	0.34	1.00										
FNO3-	0.77	0.74	-0.06	-0.06	-0.01	0.05	0.42	0.30	0.10	0.38	0.90	0.88	0.19	-0.04	1.00									
TNO3-	0.76	0.81	0.05	0.15	0.09	0.25	0.37	0.31	0.07	0.47	0.91	0.93	0.15	0.01	0.93	1.00								
FSO4=	0.43	0.60	0.30	0.37	0.29	0.43	0.15	0.17	0.04	0.31	0.63	0.62	-0.05	-0.00	0.27	0.42	1.00							
TSO4=	0.40	0.50	0.31	0.41	0.26	0.47	0.12	0.16	0.01	0.32	0.61	0.63	-0.03	0.03	0.26	0.44	0.94	1.00						
DNO3-	0.75	0.78	-0.02	0.03	0.04	0.14	0.35	0.27	0.12	0.47	0.90	0.91	0.12	-0.07	0.95	0.95	0.38	0.38	1.00					
BNO3-	0.27	0.45	0.20	0.23	0.30	0.34	0.00	0.10	0.29	0.56	0.33	0.36	-0.10	-0.11	0.20	0.36	0.44	0.46	0.41	1.00				
GNH3	0.31	0.43	0.21	0.14	0.31	0.26	0.22	0.20	0.36	0.48	0.27	0.27	0.01	0.00	0.26	0.32	0.16	0.17	0.33	0.36	1.00			
GSO2	0.37	0.46	0.17	0.17	0.28	0.31	0.07	0.14	0.34	0.59	0.38	0.36	0.03	0.00	0.35	0.40	0.26	0.30	0.40	0.42	0.40	1.00		
GNO2	0.60	0.68	0.18	0.21	0.22	0.34	0.31	0.30	0.20	0.57	0.60	0.57	0.18	0.10	0.59	0.64	0.36	0.34	0.60	0.25	0.39	0.44	1.00	
GHNO3	0.25	0.34	0.15	0.11	0.29	0.22	0.09	0.12	0.32	0.46	0.24	0.21	-0.00	-0.06	0.17	0.21	0.29	0.29	0.21	0.61	0.30	0.36	0.16	1.00

Note:
Species prefix of "F" denotes PM2.5
Species prefix of "T" denotes PM10
Species prefix of "G" denotes gas
Species prefix of "D" denotes denuded (NO3-)
Species prefix of "B" denotes backup nylon (NO3-)

TABLE 2-8

Average $PM_{2.5}/PM_{10}$ Concentration Ratios (All Samples vs. Two Sigma Outliers Excluded)

Species	CPCON01	
	All	Two Sigma
Mass	0.52	0.52
Na^+	0.32	0.31
Mg^{++}	0.29	0.28
K^+	0.59	0.59
Ca^{++}	0.33	0.32
NH_4^+	0.90	0.88
Cl^-	0.57	0.44
NO_3^-	0.52	0.49
$SO_4^=$	0.88	0.83

The data processing procedures that were employed are described below. The annually averaged dry and wet deposition data for each site and year are provided in Appendix A.

2.4.1 Wet Deposition Data Processing

A program was written to convert the wet deposition data sets provided by the ARB (and subsequently by Dr. Blanchard) to dBase format. The ARB data sets are in ASCII format and one file is provided for each year. Since there are no flags in the original ARB files, all the data provided were retained. However, we had to make some modifications to accommodate the collocated sites.

The dBase IV format files for each year were then merged into a single file called "TOTALWET.DBF". After merging all years of data, a method was established to place each record of data into a specific weekly format for proper use in a time line. First, four new fields were created: STARTDATE, STOPDATE, SMPLPERIOD, and TIME. This effectively placed

each record into a fixed seven-day period and determined the actual time lapse as TIME. The file structure changes are shown in Table 2-9.

We also developed programs to generate time-averages of the wet deposition variables for a specified averaging period. The mean rain depth and volume are simple averages over the duration of the period. Note that there is not always a value for rain depth even though there may have been a measured volume. The averages for all other fields, which are determined as "precipitation-weighted" averages (see below), are only relevant when there are corresponding precipitation measurements. The count of records used to determine the average is included in the field "WETCNT".

The method used to determine precipitation-weighted averages is based on the methodology described in the "NADP/NTN Annual Data Summary, 1991" report (section V, page 18):

$$\overline{C_w} = \frac{\sum_{i=1}^{n} C_i P_i}{\sum_{i=1}^{n} P_i} \qquad (2\text{-}1)$$

where $\overline{C_w}$ is the precipitation-weighted concentration, C_i is an individual concentration measurement, and P_i is the corresponding precipitation amount (in inches of rain).

When a measured volume of rain was provided without a corresponding rain depth in inches, the depth was calculated as "Volume in ml/1724" (CARB, 1988).

The precipitation-weighted average acidity was determined by converting the pH to H^+ concentrations, then determining the average hydrogen ion concentration using Equation (2-1), and finally converting the average hydrogen ion concentration back to average pH. The conductivity fields were treated in the same way as the concentrations.

The averaging period and averaged values can be obtained in ASCII format by using the dBase IV "LIST" command to output the desired fields. Table 2-10 shows the structure of a typical output file for the averaged wet deposition data. This ASCII file can then be used as input for STATMOD simulations or can be compiled into VOYAGER format for creating graphic representations (including time lines) of the data.

TABLE 2-9

Structure of the Original and Modified Wet Deposition Databases

Structure for Original Database				Structure for Modified Database			
Field	**Field Name**	**Type**	**Width**	**Field**	**Field Name**	**Type**	**Width**
1	STATION	Character	8	1	STATION	Character	8
2	START	Character	6	2	START	Character	6
				3	STARTDATE	Date	8
3	STOP	Character	6	4	STOP	Character	6
				5	STOPDATE	Date	8
				6	SMPLPERIOD	Date	8
				7	TIME	Numeric	4
4	RAIN	Numeric	5	8	RAIN	Numeric	5
5	VOL_F	Numeric	5	9	VOL_F	Numeric	5
6	PH_F	Numeric	5	10	PH_F	Numeric	5
7	PH_1	Numeric	5	11	PH_1	Numeric	5
8	COND_F	Numeric	5	12	COND_F	Numeric	5
9	COND_1	Numeric	5	13	COND_1	Numeric	5
10	NA	Numeric	6	14	NA	Numeric	6
11	K	Numeric	6	15	K	Numeric	6
12	CA	Numeric	5	16	CA	Numeric	5
13	MG	Numeric	5	17	MG	Numeric	5
14	NH4	Numeric	5	18	NH4	Numeric	5
15	CL	Numeric	6	19	CL	Numeric	6
16	NO3	Numeric	6	20	NO3	Numeric	6
17	SO4	Numeric	6	21	SO4	Numeric	6

TABLE 2-10

Structure of File Containing Average Wet Deposition Data

Field Name	Description	Units
STATION	ARB site number	
COUNT	Count of records available	
WETCNT	Records with measurable precipitation	
TIME	Sample duration	days
RAIN	Rain gauge reading	inches
VOL_F	Sample volume	mls
PH_F	Field pH measurement	
PH_L	Lab pH measurement	
COND_F	Field conductivity	μs/cm
COND_L	Lab conductivity	μs/cm
NA	Sodium ion concentration	mg/l
K	Potassium ion concentration	mg/l
CA	Calcium ion concentration	mg/l
MG	Magnesium ion concentration	mg/l
NH4	Ammonium ion concentration	mg/l
CL	Chloride ion concentration	mg/l
NO3	Nitrate ion concentration	mg/l
SO4	Sulfate ioin concentration	mg/l

2.4.2 Dry Deposition Data Processing

The dry deposition database was received as a single file called CPCON.DBF. To use this data in a time line effectively, some means of placing each record into a fixed time interval needed to be established. For this, two new fields were created in dBase IV, one named SCHEDDATE to indicate an arbitrary six-day period, and the other named INTERVAL to indicate the actual time lapse between samples. With the CPCON file open and indexed on SITE, DATE, and STRTIM, a utility (called FIXDRY.PRG) was run to assign an appropriate SCHEDDATE for each record except the first record of each SITE. These "first records" were then hand entered. Next, the INTERVAL's were calculated with a utility named MKPRDDRY.PRG.

The file structure changes are shown in Table 2-11. We used the original flags in the data that we received from the ARB to eliminate totally invalid records. For the purposes of averaging and certain time line plots, it was desirable to separate the data into daytime and nighttime periods. These daytime and nighttime sub-files (CPCOND and CPCONN, respectively) were created manually using the appropriate filters on the database.

We developed a program (DRYAVG.PRG) to generate a report of averages of each field of the dry deposition (CPCON) file for any provided range of dates. This program can be used to generate seasonal and annual averages for model use. We also developed a utility (OUTLIER.PRG) to identify outliers. Both programs can either print directly to a printer or create an ASCII file using the dBase IV "LIST" command.

TABLE 2-11

Structure of the Original and Modified Dry Deposition Databases

Structure for Original Database				Structure for Modified Database			
Field	**Field Name**	**Type**	**Width**	**Field**	**Field Name**	**Type**	**Width**
1	IDGT	Character	10	1	IDGT	Character	12
2	IDTK	Character	10	2	IDTK	Character	12
3	IDDN	Character	10	3	IDDN	Character	12
4	IDTN	Character	10	4	IDTN	Character	12
4	SITE	Character	2	5	SITE	Character	2
6	SAMPID	Numeric	2	6	SAMPID	Numeric	2
				7	SCHEDDATE	Date	8
7	DATE	Date	8	8	DATE	Date	8
8	STRTIM	Numeric	4	9	STRTIM	Numeric	4
9	PERIOD	Character	1	10	PERIOD	Character	1
				11	INTERVAL	Numeric	2
10	FGTTFLG	Character	5	12	FGTTFLG	Character	5
11	FTKTFLG	Character	5	13	FTKTFLG	Character	5
12	FDNFFLG	Character	5	14	FDNFFLG	Character	5
13	FTNFFLG	Character	5	15	FTNFFLG	Character	5

3.0 MODEL DESCRIPTION

The model used in this study is a nonlinear semi-empirical long-range transport model (STATMOD) that scientists at ENSR have developed and refined over the last decade (Venkatram et al., 1982; 1990; 1993; Venkatram and Pleim, 1985). The model is nonlinear in the sense that it accounts for the role of oxidant limitation in the aqueous-phase oxidation of SO_2 in both precipitating and non-precipitating clouds.

STATMOD was developed initially to allow the estimation of long-term concentrations and depositions of sulfur compounds. It was subsequently expanded to treat the long-range transport of nitrogen compounds. The model is based on the premise that long term (annual or seasonal) averages of concentration and deposition of a pollutant are insensitive to short term fluctuations of the governing processes such as meteorology and chemistry. This assumption allows us to base model estimates on the statistics of the governing processes.

The scientific credibility of STATMOD is determined by several factors. First, the components of the model should be consistent with the corresponding physical processes. Second, the model should explain a large fraction of the variance of the observations it is being fitted against. Furthermore, the values of the model parameters obtained by fitting model estimates to observations should be within the range of measurements corresponding to the physical processes that they represent.

3.1 Summary of Model Features

STATMOD incorporates parameterizations for the transport, chemistry, and scavenging (dry and wet deposition) of sulfur and nitrogen species. Transport is treated using straight-line trajectories governed by large scale wind roses derived from upper air winds. To account for the spatial variation in the transport winds, the region of interest is divided into several subregions, each of which is assigned its own wind rose. Transport between a given source and receptor is determined by the wind rose in the subregion containing the source.

The model uses a stochastic approach in which sulfur (or nitrogen) is assigned to four states: wet and dry primary species (SO_2 or NO_x), and wet and dry secondary species (sulfate or nitrate). Four differential equations govern the temporal evolution of these four states. The transition of species between wet and dry periods is governed by two time scales, which are the average lengths of wet and dry periods.

The transition from SO_2 to sulfate (or NO_x to nitrate) occurs through gas-phase conversion, which is parameterized in terms of conversion rates for wet and dry periods. However, SO_2 is also oxidized (primarily by H_2O_2) to sulfate in the aqueous-phase in both precipitating and non-precipitating clouds — these processes are accounted for by specifying an oxidant-limited wet scavenging rate of SO_2 for the wet state (i.e., precipitating clouds) and specifying an oxidant-limited SO_2 oxidation rate for the dry state (i.e., non-precipitating clouds).

The dry deposition of sulfur and nitrogen is treated by specifying dry deposition velocities. The wet removal of SO_2, sulfate and nitrate is parameterized in terms of a scavenging rate that is proportional to the rainfall amount at the receptor of interest. As discussed above, the wet scavenging of SO_2 is actually a surrogate process for the aqueous-phase oxidation of SO_2 in precipitating clouds.

All the parameters of the model have physical meaning, and therefore can be assigned values on the basis of our understanding of the relevant processes. However, in the actual application of the model, some of the parameter values are adjusted to obtain an optimum fit between model estimates and the observations being analyzed. This process is carried out either by using a non-linear optimization program or by examining the scatterplots comparing model estimates and observations, and manually adjusting the parameter values on the basis of the comparison and the underlying physics. Because the model parameters are associated with physical processes, it is necessary to ensure that their values resulting from this exercise are consistent with our prior understanding of these processes. Thus, in the optimization procedure, parameters are not allowed to take on values outside their expected range.

The formulation of the parameterizations and the governing differential equations are described in more detail in the following section.

3.2 Model Formulation

The premise that long-term averages of concentrations and depositions are primarily a function of the statistics of the governing processes allows us to write differential equations for the evolution of the probabilities of the pollutants existing in specific states. It is reasonable to assume that, as we sample at a location (or travel time, t) over a sufficiently long time, we can observe sulfur or nitrogen in these different states. The states and associated probabilities that are used in the model (Venkatram and Pleim, 1985) are defined below.

$G_d(t)$ = probability that sulfur (or nitrogen) exists as SO_2 (or NO_x) in the dry state at time t,

$G_w(t)$ = probability that sulfur (or nitrogen) exists as SO_2 (or NO_x) in the wet state at time t,

$S_d(t)$ = probability that sulfur (or nitrogen) exists as sulfate (SO_4) [or nitrate (NO_3)] in the dry state at time t, and

$S_w(t)$ = probability that sulfur (or nitrogen) exists as sulfate (SO_4) [or nitrate (NO_3)] in the wet state at time t.

Sulfur (or nitrogen) is said to be in the "wet" state, when it is embedded in a synoptic scale rain system.

The simple model illustrated in Figure 3-1 can be used to derive the relevant differential equations. The pollutants are taken to be well mixed throughout the depth of the boundary layer, and horizontal spread is characterized by a constant angle of spread, θ. Then, the flux of material at any distance is given by

$$F = u\theta r z_i C, \tag{3-1}$$

where F is the flux (e.g., moles/s), u is the mean wind speed (m/s), r is the downwind distance (m), z_i is the mixing height (m) and C is the concentration of the pollutant (e.g., moles/m^3) at distance r. Note that the flux, as defined in Equation (3-1), is different from the conventional definition of flux (which is defined as rate of mass transported per unit area). Our definition is actually the conventional flux (=uC) times the area across which transport occurs at distance "r" (=$\theta r z_i$).

Over the long term, the fluxes of the pollutants during wet and dry periods are identical to the probabilities defined earlier, so that

$$G_d = f_d r C_{dSO2}\ \theta\ z_i u \tag{3-2a}$$

$$G_w = f_w r C_{wSO2}\ \theta\ z_i u \tag{3-2b}$$

$$S_d = f_d r C_{dSO4}\ \theta\ z_i u \tag{3-2c}$$

$$S_w = f_w r C_{wSO4}\ \theta\ z_i u \tag{3-2d}$$

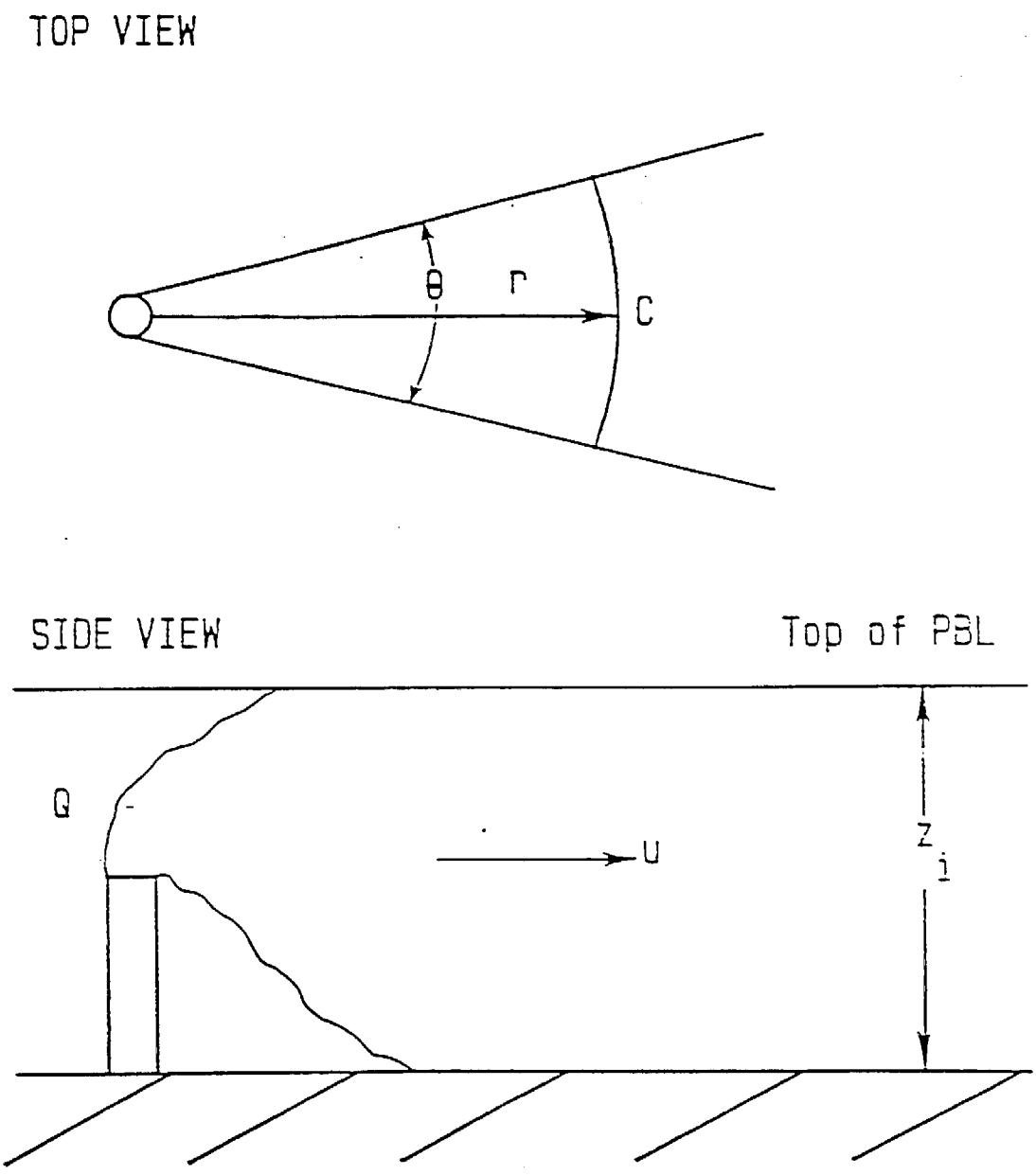


FIGURE 3-1. Idealization of Dispersion of a Plume in the Planetary Boundary Layer

where f_d and f_w are the fractions of dry and wet periods at the receptor, respectively, and are given by

$$f_d = \tau_d/(\tau_d + \tau_w) \quad (3\text{-}3a)$$

$$f_w = 1 - f_d \quad (3\text{-}3b)$$

where τ_d and τ_w are the lengths of dry and wet periods, respectively. Note that we have used SO_2 and SO_4 in Equation (3-2) for convenience. The probabilities for the four nitrogen states can be defined in an analogous manner.

Figure 3-2 is a probability interaction diagram for the four sulfur states. The rates at which SO_2 and sulfate undergo transition from the wet state to the dry state and vice versa are denoted by $1/\tau_w$ and $1/\tau_d$. The conversion of SO_2 to sulfate by gas-phase chemistry is represented by k_d and k_w for the dry and wet states, respectively. The aqueous-phase oxidation of SO_2 in non-precipitating clouds (which occurs only in the dry state) is represented by k_d'. The dry deposition removal rate constants for SO_2 and sulfate are represented by λ_d and $\overline{\lambda}_d$, respectively. Note that the dry scavenging rate, λ_d, is nothing but the deposition velocity divided by the mixing height (v_d/z_i). The wet scavenging coefficients of SO_2 and sulfate are represented by λ_w and $\overline{\lambda}_w$, respectively. Recall that the SO_2 wet scavenging coefficient implicitly accounts for aqueous-phase SO_2 oxidation in precipitating clouds.

The probability interaction diagram for nitrogen species, shown in Figure 3-3, is similar to that for the sulfur species. The primary difference is that NO_x to nitrate conversion is assumed to occur solely in the gas phase. Thus, there is no pathway in Figure 3-3 equivalent to that represented by k_d' in Figure 3-2.

The following sections describe the development of the differential equations that represent the evolution of the four sulfur and nitrogen states, and the calculation of dry and wet depositions from the computed probabilities.

3.2.1 Governing Equations for Sulfur

The gas-phase oxidation of SO_2 to sulfate is assumed to be a linear process. The model accounts for the non-linear effect of oxidant limitation on the wet scavenging of SO_2 by assuming that the scavenging rate is proportional to the oxidant concentration when the SO_2 concentration is larger than the oxidant concentration (on a ppb or molar basis). The oxidant concentration is one of the parameters of the model. We assume that the primary aqueous-

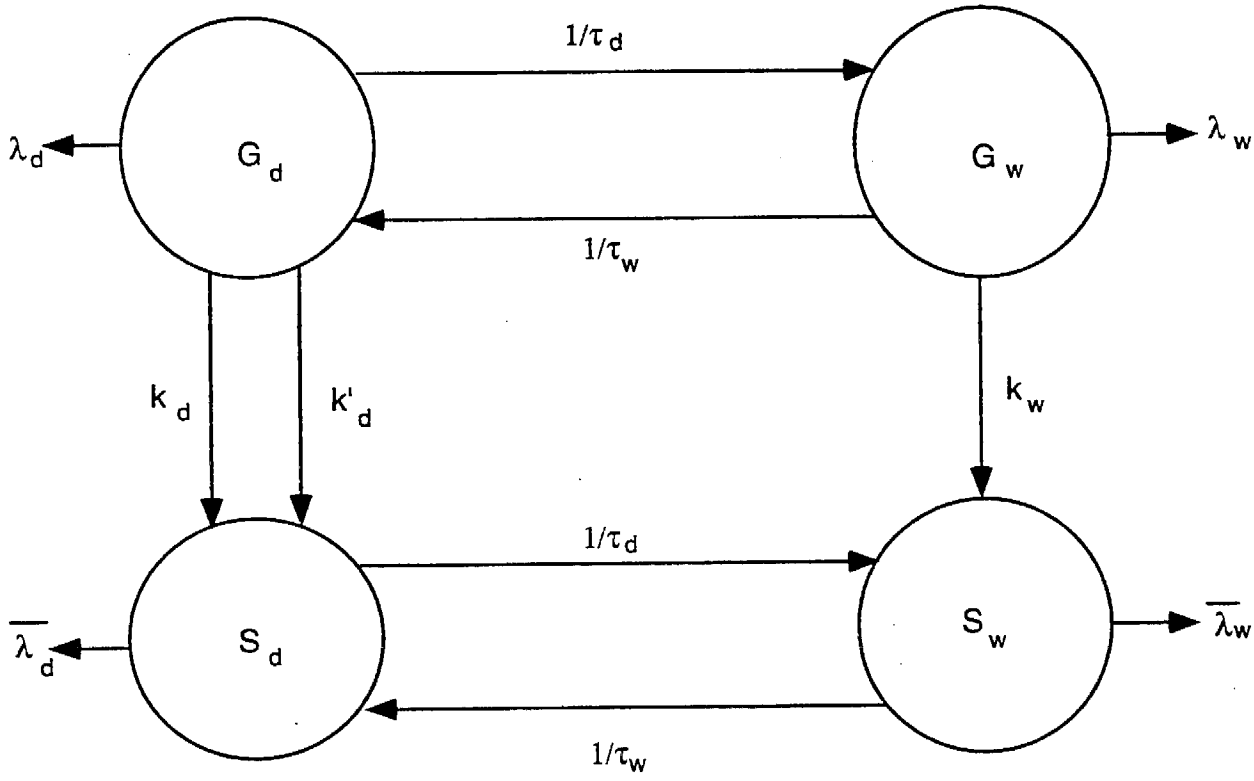


FIGURE 3-2. Probability Interaction Diagram for Atmospheric Sulfur

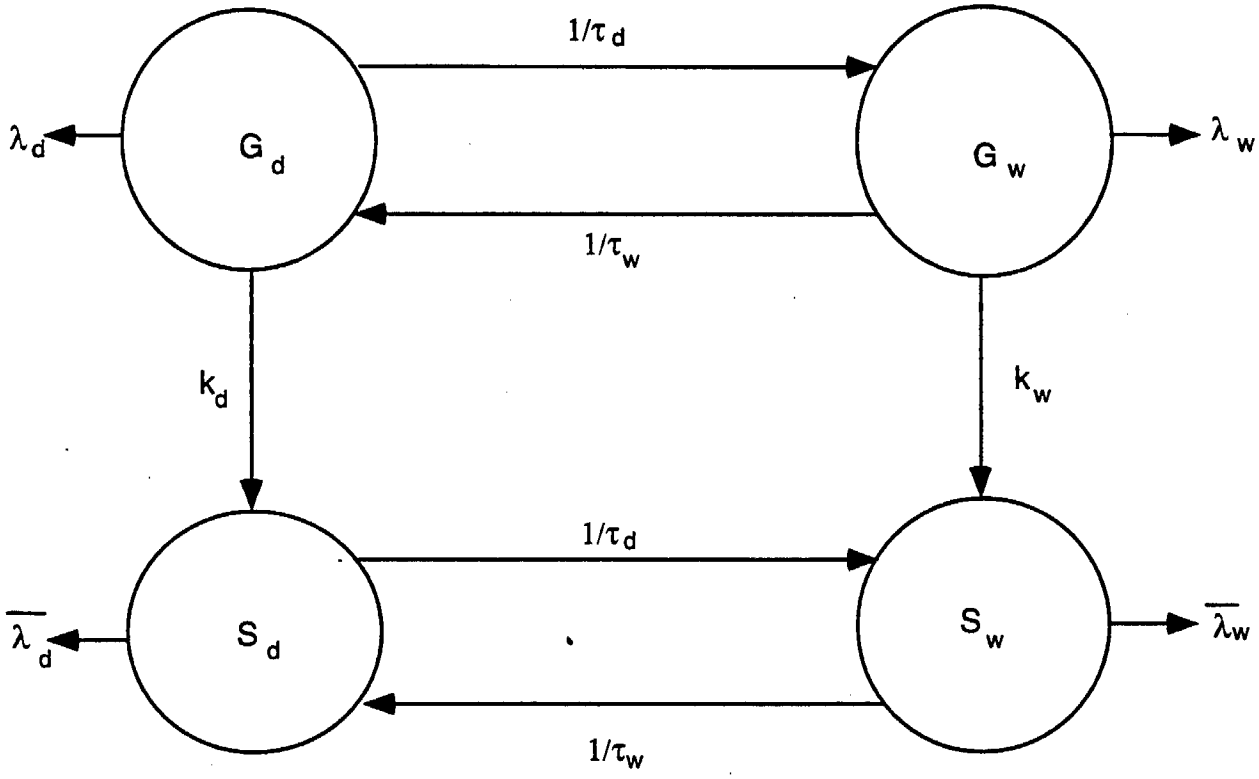


FIGURE 3-3. Probability Interaction Diagram for Atmospheric Nitrogen

phase oxidant is hydrogen peroxide (H_2O_2). The basis for this assumption is that the rates for the other aqueous-phase SO_2 oxidation pathways (oxidation by ozone, trace metal catalyzed oxidation) are generally smaller than the rate for SO_2 oxidation by H_2O_2 under typical atmospheric conditions. When the SO_2 concentration is lower than the oxidant concentration, the scavenging rate is proportional to the SO_2 concentration — this corresponds to linear aqueous-phase oxidation. Thus, the wet removal term for SO_2 is given by

$$D_w = \lambda_w \chi_0 \text{ when } C_{SO2} > \chi_0 \quad (3\text{-}4)$$

where χ_o is the oxidant concentration, and

$$D_w = \lambda_w C_{SO2} \text{ when } C_{SO2} < \chi_0 \quad (3\text{-}5)$$

The formulation of the equations then requires that two regimes be considered: in the region close to a source, where high SO_2 concentrations are expected, the oxidant-limited equations apply; at larger distances downwind, the linear model applies. This can be stated formally by defining a limiting travel time, t_s, or travel distance, r_s, where the SO_2 concentration is just equal to the oxidant concentration (an iterative scheme is used to determine the limiting travel time for each source-receptor combination).

SO_2 can also be oxidized to sulfate in non-precipitating clouds (e.g., Karamchandani and Venkatram, 1992). Because this process is also oxidant-limited, we use the same technique that is used for the wet scavenging of SO_2 — the rate at which SO_2 in the dry state is oxidized in non-precipitating clouds is assumed to be proportional to its concentration only when the SO_2 concentration is less than that of the oxidant; otherwise, the rate of oxidation is proportional to the oxidant concentration. As in the wet scavenging case, this formulation requires the definition of an oxidant-limited regime and a regime where the linear rate applies.

The equations that describe the evolution of the four sulfur states can then be written as

$$\frac{dG_d}{dt} = -k_d G_d - k_d' \chi_o r f_d - \frac{1}{\tau_d} G_d + \frac{1}{\tau_w} G_w \quad (for\ t < t_{s1}) \tag{3-6a}$$

$$\frac{dG_d}{dt} = -(k_d + k_d') G_d - \lambda_d G_d - \frac{1}{\tau_d} G_d + \frac{1}{\tau_w} G_w \quad (for\ t > t_{s1}) \tag{3-6b}$$

$$\frac{dG_w}{dt} = -k_w G_w - \lambda_d G_w - \frac{1}{\tau_w} G_w - \lambda_w \chi_o r f_w + \frac{1}{\tau_d} G_d \quad (for\ t < t_{s2}) \tag{3-7a}$$

$$\frac{dG_w}{dt} = -k_w G_w - \lambda_d G_w - \frac{1}{\tau_w} G_w - \lambda_w G_w + \frac{1}{\tau_d} G_d \quad (for\ t > t_{s2}) \tag{3-7b}$$

$$\frac{dS_d}{dt} = k_d G_d + k_d' \chi_o r f_d - \overline{\lambda_d} S_d - \frac{1}{\tau_d} S_d + \frac{1}{\tau_w} S_w \quad (for\ t < t_{s1}) \tag{3-8a}$$

$$\frac{dS_d}{dt} = (k_d + k_d') G_d - \overline{\lambda_d} S_d - \frac{1}{\tau_d} S_d + \frac{1}{\tau_w} S_w \quad (for\ t > t_{s1}) \tag{3-8b}$$

$$\frac{dS_w}{dt} = k_w G_w - \overline{\lambda_d} S_w + \frac{1}{\tau_d} S_d - \frac{1}{\tau_w} S_w - \overline{\lambda_w} S_w \tag{3-9}$$

For convenience, the constant term, $\theta z_i u$, has been canceled from both sides of the above set of differential equations. In Equations (3-6) through (3-8), t_{s1} is the limiting travel time for oxidant-limited SO_2 oxidation in non-precipitating clouds (dry state only), while t_{s2} is the limiting travel time for oxidant-limited wet scavenging of SO_2 (wet state only).
The initial conditions for Equations (3-6) through (3-9) are:

$$G_d(0) = f_d Q_{SO2} / \theta z_i u \tag{3-10a}$$

$$G_w(0) = f_w Q_{SO2} / \theta z_i u \tag{3-10b}$$

$$S_d(0) = f_d Q_{SO4} / \theta z_i u \tag{3-10c}$$

$$S_w(0) = f_w Q_{SO4} / \theta z_i u \tag{3-10d}$$

where Q_{SO2} and Q_{SO4} are the emission rates for SO_2 and sulfate, respectively.

3.2.2 Governing Equations for Nitrogen

The differential equations governing the evolution of the four nitrogen states can be derived in a similar manner as those for the sulfur states. As pointed out earlier, the primary difference between the two is that there is no mechanism to convert NO_x to nitrate in the aqueous-phase. Then, the governing equations can be written as:

$$\frac{dG_d}{dt} = -k_d G_d - \lambda_d G_d - \frac{1}{\tau_d} G_d + \frac{1}{\tau_w} G_w \quad (3\text{-}11)$$

$$\frac{dG_w}{dt} = -k_w G_w - \lambda_d G_w - \frac{1}{\tau_w} G_w - \lambda_w G_w + \frac{1}{\tau_d} G_d \quad (3\text{-}12)$$

$$\frac{dS_d}{dt} = k_d G_d - \overline{\lambda_d} S_d - \frac{1}{\tau_d} S_d + \frac{1}{\tau_w} S_w \quad (3\text{-}13)$$

$$\frac{dS_w}{dt} = k_w G_w - \overline{\lambda_d} S_w - \frac{1}{\tau_w} S_w + \frac{1}{\tau_d} S_d - \overline{\lambda_w} S_w \quad (3\text{-}14)$$

The initial conditions for Equations (3-11) through (3-14) are similar to those for sulfur [shown in Equation (3-10)], with the NO_x and nitrate emission rates replacing the SO_2 and sulfate emission rates.

3.3 Calculation of Dry and Wet Deposition

We can derive analytical solutions to Equations (3-6) through (3-14) for each source-receptor combination if the parameters are assumed to be independent of concentrations. Once the equations are solved, the dry and wet deposition of sulfur and nitrogen species for a given source-receptor combination can be determined as described below.

3.3.1 Dry Deposition

The dry deposition, D^{d}_{ijp}, of primary species (SO_2 or NO_x) at receptor "j" due to emission at source "i" is given by

$$D^{d}_{ijp} = z_i \lambda_d \left[G_d(t_{ij}) + G_w(t_{ij}) \right] T_{ij} \quad (3\text{-}15)$$

Similarly, the dry deposition, D^{d}_{ijs}, of secondary species (sulfate or nitrate) is given by

$$D_{ijz}^{d} = z_i \overline{\lambda_d} \, [S_d(t_{ij}) + S_w(t_{ij})] T_{ij} \qquad (3\text{-}16)$$

In Equations (3-15) and (3-16), t_{ij} is the time of travel between the source and the receptor, and T_{ij} represents the large scale dispersion. Notice that dry deposition occurs under both dry and wet conditions.

The vertical dispersion can be treated by assuming that the pollutants are well-mixed throughout the depth of the boundary layer. Past experience (Fisher, 1978; Venkatram and Pleim, 1985) indicates that it is possible to make reasonable estimates of the horizontal dispersion by assuming that the trajectories from the source are approximately straight. This assumption allows the following simple expressions for T_{ij} and t_{ij}:

$$T_{ij} = \frac{f_\theta}{2\pi r_{ij} u_\theta z_i} \; ; \; t_{ij} = r_{ij} / u_\theta \qquad (3\text{-}17)$$

where f_θ is the relative frequency with which the large scale wind blows in the direction of the source to the receptor, u_θ is the wind speed along r, and $r_{ij} = | r |$. The parameters, f_θ and u_θ, can be derived from rawinsonde observations of upper-air wind speeds and directions. Notice that the wind rose corresponds to the averaging period of interest.

The total dry deposition at the receptor is then the sum of the contribution of all the sources. The model computes the dry deposition of both primary and secondary species, as well as the sum of the two, i.e., the dry deposition of total sulfur (SO_2 + sulfate) and total nitrogen (NO_x + nitrate).

3.3.2 Wet Deposition

The computation of wet deposition is similar to that for dry deposition. The primary difference is that wet deposition is computed only for pollutants in the wet state. Another difference is that the total sulfur (primary + secondary) wet deposition is computed when comparing model estimates with observations of sulfate concentrations in rain. This is consistent with the observation that sulfur in rain occurs primarily as sulfate and with our interpretation of the SO_2 wet scavenging rate as an aqueous-phase SO_2 oxidation rate. Then, the total sulfur wet deposition, D_{ij}^{w}, at receptor "j" due to emission at source "i" is given by

$$D_{ij}^{w} = z_i \, [\lambda_w G_w(t_{ij}) + \overline{\lambda_w} S_w(t_{ij})] \, T_{ij} \qquad (3\text{-}18)$$

A similar equation applies for total nitrogen wet deposition. In this case, the contribution of the primary species, NO_x, to total wet deposition is small because NO_x is relatively insoluble. This insolubility is reflected in the value used for the NO_x wet scavenging rate coefficient.

3.4 Advantages and Disadvantages of the Model

The primary advantage of STATMOD is that its simplicity allows us to easily determine the relative importance of the governing processes. Furthermore, it is easy to use and its computer and input data requirements are small. Thus, it can economically provide estimates of long-term (annual or seasonal) source-receptor relationships. Its primary disadvantage is also a consequence of its simplicity — STATMOD represents the complex physical and chemical processes governing the deposition of acidic species in a parameterized fashion. It cannot treat the details of these processes — for example, the complex sequence of atmospheric reactions leading to the formation of SO_2 and NO_x oxidants and subsequent oxidation of SO_2 and NO_x are parameterized as first-order processes with specified rate constants.

Comprehensive models, such as the Regional Acid Deposition Model (RADM — Chang et al., 1987), and the Acid Deposition and Oxidant Model (ADOM — Venkatram et al., 1988), represent the governing processes in more detail. However, they are not easy to use and require large computational and labor resources to develop long-term source-receptor relationships and to conduct sensitivity studies. A semi-empirical model, such as STATMOD, is the most cost-effective tool for this purpose. Comprehensive models are best suited for episodic applications, and for providing a mechanistic explanation of the governing processes. Both semi-empirical and comprehensive models can play complementary roles in enhancing our understanding of the acidic deposition system in California.

The utility of STATMOD is determined by the consistency of its formulation with the corresponding physical processes, and its performance history. The model should also explain a large fraction of the variance of the observations it is being fitted against. Furthermore, the values of the model parameters obtained by fitting model estimates to observations should be within the range of measurements corresponding to the physical processes that they represent. Finally, sensitivity studies conducted by varying model parameters can provide useful information on the uncertainty of model estimates. Sections 5 and 7 describe the performance of the model with the CADMP database and results from the parametric sensitivity studies, respectively. Below, we provide a brief description of results obtained from applications of the model to other situations.

3.5 Performance History of the Model

Venkatram et al. (1982) initially applied STATMOD to estimate wet deposition of sulfur over the northeastern United States and Canada. Venkatram and Pleim (1985) used the model to explain observations of annual wet sulfur deposition at 62 stations in northeastern America. They showed that it was necessary to include an efficient SO_2 removal mechanism by rain to explain the observed values. This is consistent with our understanding that in-cloud oxidation of SO_2 is a significant source of sulfate in rain.

Venkatram et al. (1990) used an improved version of STATMOD, which included a treatment of oxidant limitation on SO_2 removal by rain (see Section 3.2.1), to explain observations of sulfur wet deposition over eastern North America corresponding to the period 1982–1985. The model was found to provide an excellent description of the spatial patterns of annual averages of sulfur concentrations in rain. The model parameters obtained by fitting model estimates to observations were consistent with their expected values. Furthermore, the importance of including oxidant limitation in the formulation was underlined by this study — when oxidant limitation was removed (by specifying a large value for the oxidant concentration), model performance was adversely affected, and optimum values of other model parameters were found to be unreasonable.

In a study performed for the Electric Power Research Institute (EPRI), STATMOD was applied to estimate source-receptor relationships for sulfur deposition at 15 receptor sites in eastern North America (UAPSP, 1992). The United States and Canada were divided into 26 sulfur emission source regions for this study. Model parameters were optimized in order to obtain the best fit between seasonal observations of sulfur concentrations in rain and model estimates for the aggregate (or composite) year, 1982–1985. The model (with the optimized parameters) was then independently tested by comparing model estimates with seasonal observations of sulfur concentrations in rain for the years 1986 and 1987.

In all the studies described above, sulfur concentrations in rain were primarily used to optimize STATMOD parameters. However, in a recent study for EPRI, ambient SO_2 and sulfate concentrations were also used to optimize the model (Venkatram et al., 1993). Model estimates of sulfur concentrations in rain, and ambient SO_2 and sulfate concentrations were compared with observations from the Eulerian Model Evaluation Field Study and the National Dry Deposition Network (NDDN) for the period from July to September, 1988. The study showed that it was necessary to include a treatment for sulfate formation in nonprecipitating clouds to explain observed SO_2 and sulfate concentrations (Venkatram et al., 1993)

In the same study, the model was evaluated with multiple independent datasets (Venkatram et al., 1993). Specifically, optimized parameters from previous studies were used to explain ambient SO_2 and sulfate concentrations measured during the Sulfate Regional Experiment (SURE)) and the Eastern Region Air Quality Study (ERAQS) for the following aggregated periods: winter 1977 and 1978; fall 1977 and 1978; spring 1978 and 1979; summer 1978; and July to September, 1978. Precipitation chemistry data from the Acid Deposition System (ADS) database for these periods were also used to compare model estimates of sulfur concentrations in rain with observed concentrations. No further optimization of model parameters was conducted for these independent datasets. In general, the model performance was reasonable, suggesting that the model is much more than an empirical fit to data, and can be used with confidence to develop source-receptor relationships for periods other than those used to optimize the model parameters.

STATMOD has also been used to explain observations of nitrate concentrations in rain in the eastern United States (ENSR, 1989). In this study, sponsored by EPRI, STATMOD estimates of annually and seasonally averaged sulfur and nitrate concentrations in rain were compared to observed concentrations at 74 locations for the years 1982–1985. The optimized model explained 70 percent of the variance in the observed nitrate concentrations.

4.0 PREPARATION OF MODEL INPUTS

The application of STATMOD requires an emission inventory for the species of interest (SO_x and NO_x in this case), upper air winds, and ambient air quality and precipitation chemistry data to evaluate the model and optimize model parameters. These data were required for 1984 through 1989, the period of interest for this study.

In this section, we describe the emissions and meteorological datasets that were used in our study (the CADMP database, which provides the air quality and precipitation chemistry data, has already been discussed in Section 2). We also discuss how these data were processed to create STATMOD input files.

4.1 Emissions

We obtained a summary of emissions from the ARB for each county in California for the inventory years 1985, 1987, and 1989. The summary includes annual emissions of total organic gases (TOG), reactive organic gases (ROG), carbon monoxide (CO), oxides of nitrogen (NO_x), and sulfur (SO_x). However, for this study, we only required the NO_x and SO_x emissions.

The information obtained from the ARB provides emission totals for each county in California. However, in order to calculate source-receptor distances in STATMOD, we had to associate each source region (county) with a location. Because emissions within a source region are usually not distributed uniformly, emission-weighted centers for each source region were calculated.

To calculate the emission-weighted source location for each county or source region, a gridded inventory for the whole state is needed. However, except for a few regions which have been extensively modeled (e.g., the South Coast Air Basin or the San Joaquin Valley), gridded inventories for the entire state are presently unavailable. For those regions where gridded inventories were available (e.g., the SARMAP inventory for the SJVAQS/AUSPEX domain, and the SCAQS inventory for the South Coast Air Basin), we used the information to estimate the emission-weighted source locations. For the other source regions, we assumed that the source was located at either the geographical center (for rural areas) or at a large population center (for urban areas). Figure 4-1 shows the emission-weighted locations and relative magnitudes of annual NO_x emissions for 1985. Figure 4-2 provides the same information for SO_x emissions.

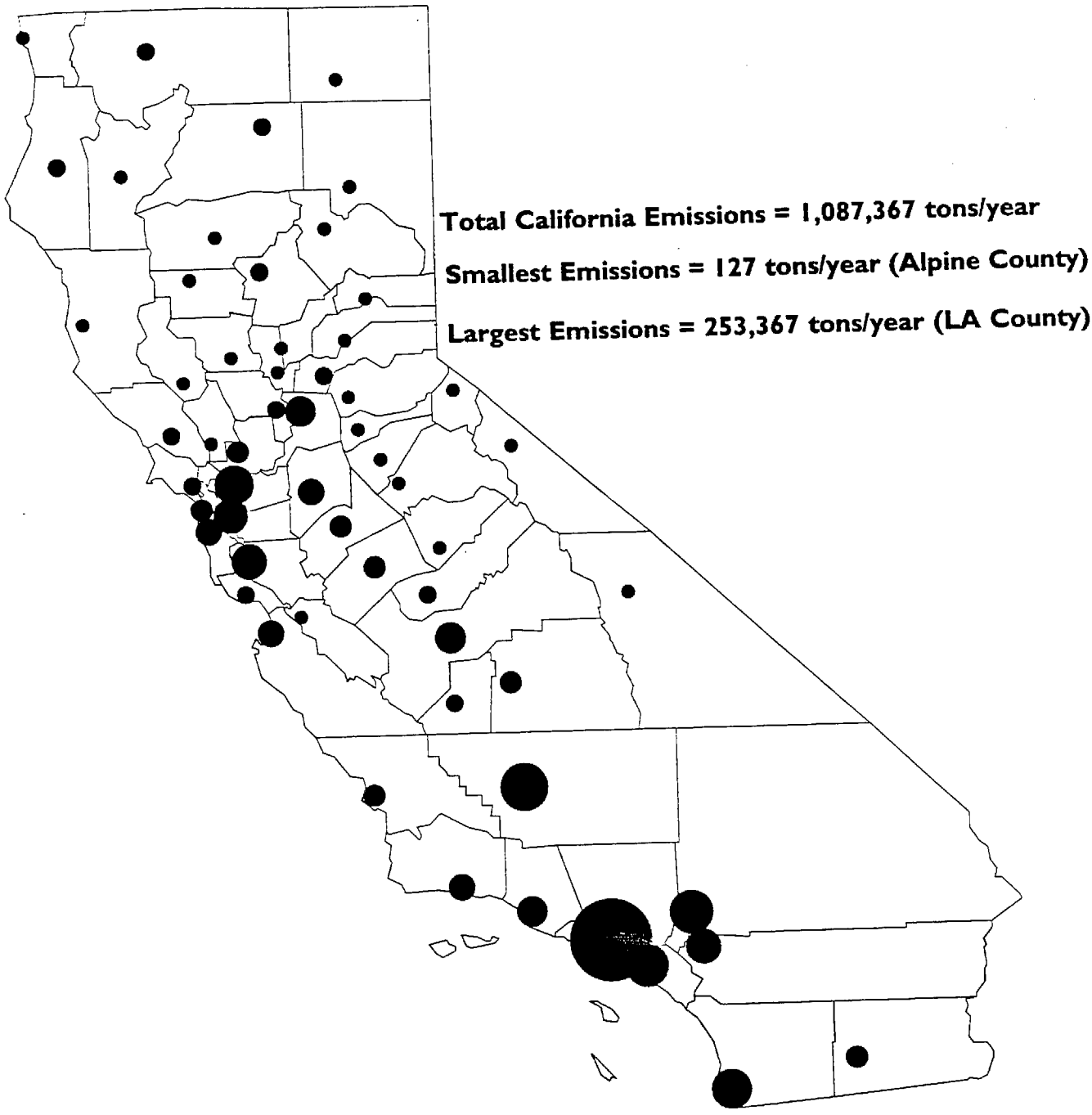


FIGURE 4-1. 1985 NO_x Emissions by County

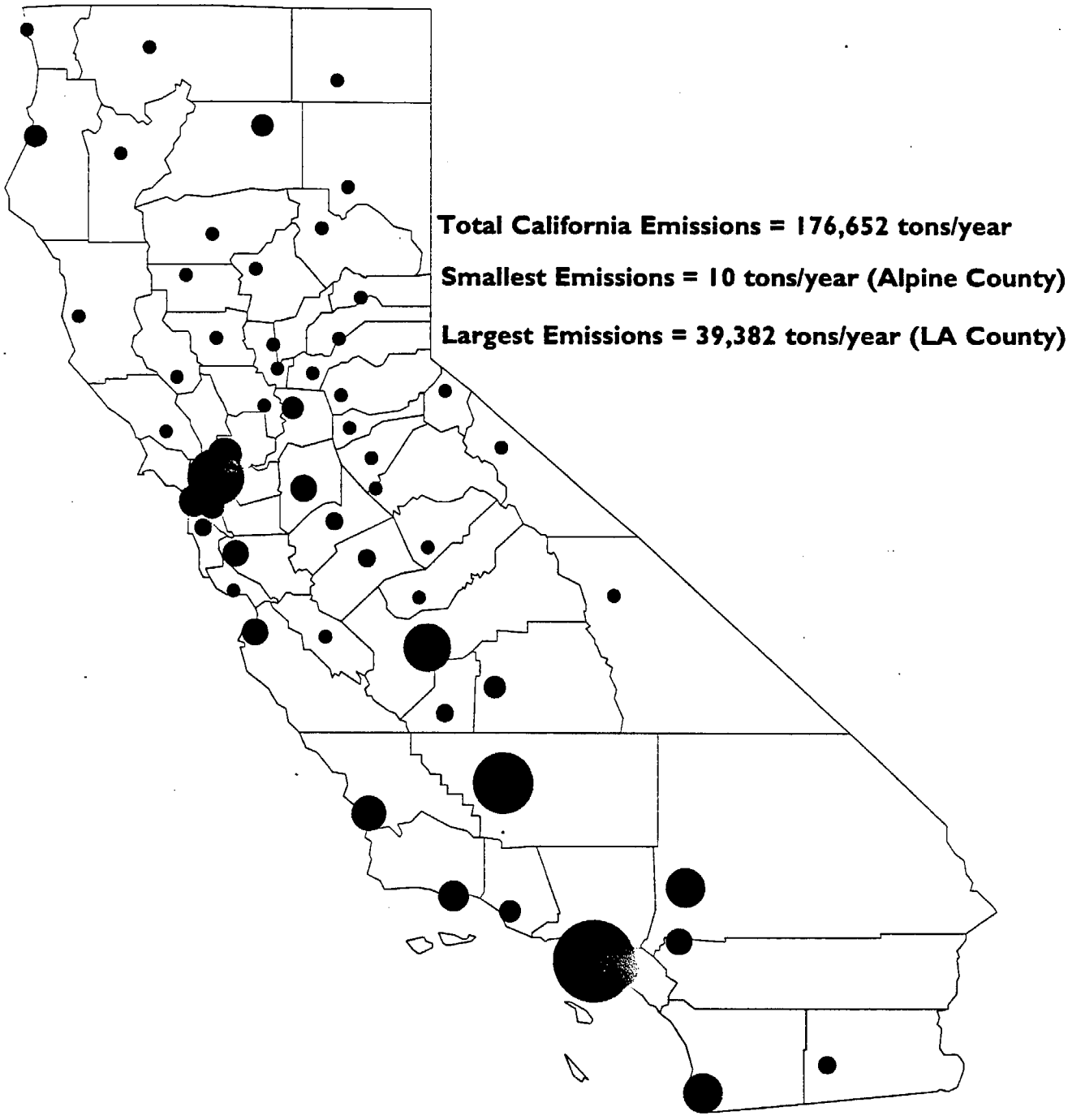


FIGURE 4-2. 1985 SO_x Emissions by County

The county summaries and the emission-weighted source locations were used to create SO_x and NO_x input files in STATMOD format for the years 1985, 1987, and 1989. Since emission summaries were not available for the other years of interest (1984, 1986, and 1988), we had to use the available information to generate input files for these years. We used the 1985 emission inventory for the 1984 simulations, while for 1986 and 1988, we used the averages of the emission inventories of the neighboring years. We also developed a composite-year emission inventory by taking the average of the 1985, 1987, and 1989 emission inventories.

4.2 Upper Air Winds

Upper air wind speed and direction data for seven sites were obtained from the National Climatic Data Center (NCDC). These sites are: Salem, OR; Medford, OR; Winnemuca, NV; Ely, NV; Desert Rock, NV; Oakland, CA; and San Diego, CA. Upper air data are available for all the years of interest (1984 through 1989) at the San Diego and Desert Rock stations only. Data for 1987 are not available at the other sites. The NCDC database includes upper air measurements at various pressure levels at 12-hour intervals.

We also obtained upper air data for several locations in California from the ARB. However, a reasonably complete database for the period of interest was available only at the Fresno site. The other sites reported observations for only a few months during each year. Thus, we only used information at the Fresno site to prepare meteorological input files for STATMOD.

Because of the large differences in wind flow patterns in the various air basins of California, we had to divide California into several meteorological regions to obtain an adequate representation of the meteorology. We defined seventeen meteorological zones, which, for the most part, correspond to the California air basins. Figure 4-3 shows the 17 zones, while Table 4-1 lists the zones as well as the upper air data that were used to develop wind roses for the zones.

Since upper air data for the years of interest for this study are not available to define the wind flow patterns in all the zones, we have supplemented the available upper air data discussed above with historical summaries of California upper air data compiled by the ARB (CARB, 1979). Note that this approach introduces the limitation that year-to-year variability in the upper air winds cannot be resolved for some regions.

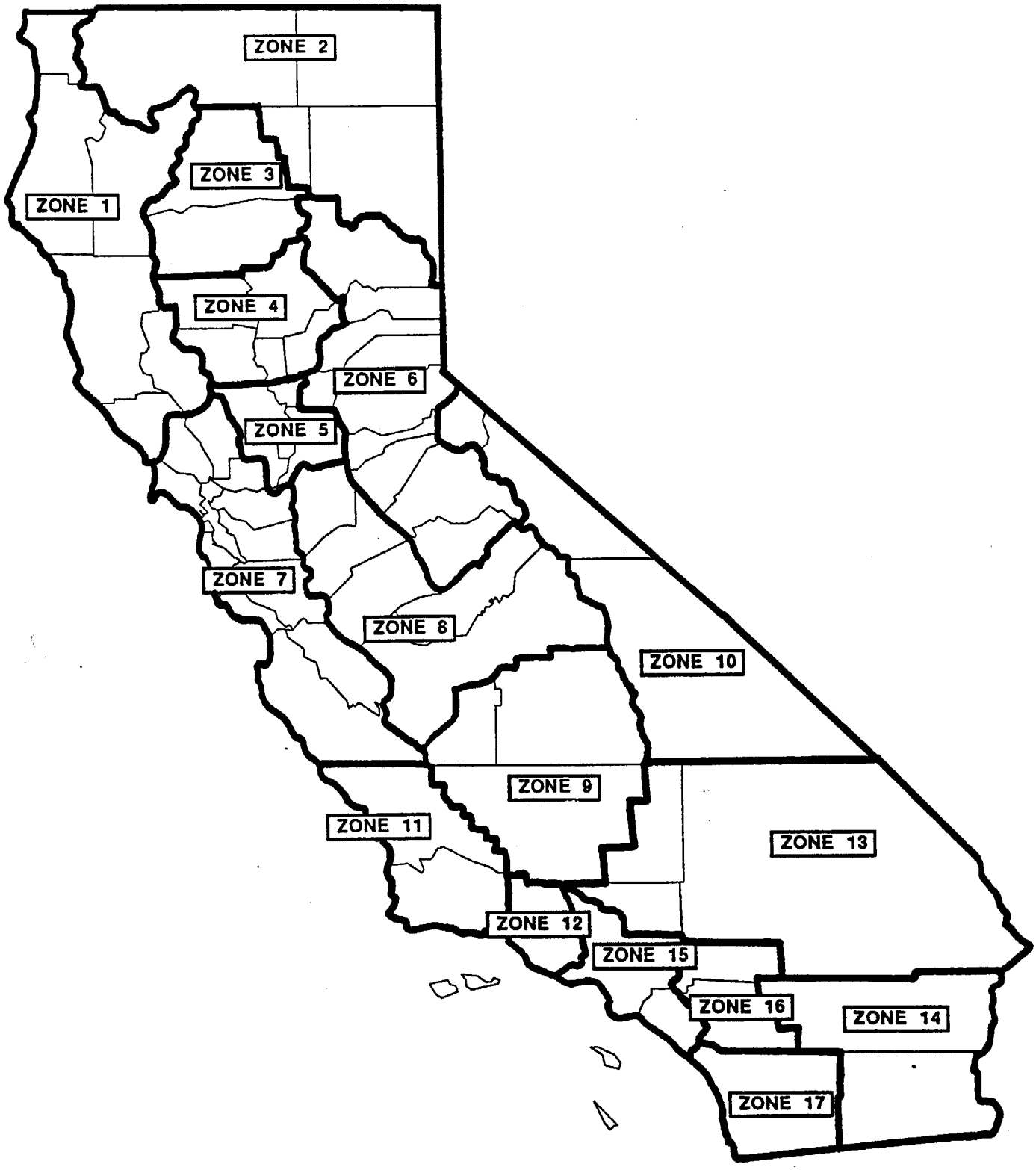


FIGURE 4-3. Division of California into 17 Meteorological Zones

TABLE 4-1

The 17 California Zones Used in the Modeling Study

Zone	County	Upper Air Data Source
1	North Coast and Lake County	Medford, OR (NCDC)
2	Northeast Plateau	Winnemuca, NV (NCDC)
3	Upper Sacramento Valley	ARB historical summary
4	Middle Sacramento Valley	ARB historical summary
5	Lower Sacramento Valley	ARB historical summary
6	Mountain Counties and Lake Tahoe	Ely, NV (NCDC)
7	San Francisco Bay Area	Oakland, CA (NCDC)
8	Upper San Joaquin Valley	Fresno, CA (ARB)
9	Lower San Joaquin Valley	ARB historical summary
10	Great Basin Valleys	Desert Rock, NV (NCDC)
11	Upper South Central Coast	ARB historical summary
12	Lower South Central Coast	ARB historical summary
13	Upper Southeast Desert	ARB historical summary
14	Lower Southeast Desert	ARB historical summary
15	Western South Coast	ARB historical summary
16	Eastern South Coast	ARB historical summary
17	San Diego	San Diego, CA (NCDC)

We developed the necessary software to read the upper air meteorological data files and construct annual and seasonal wind roses for the years 1984 through 1989. In addition, we developed wind roses for the composite year and composite seasons.

4.3 Rainfall Amounts

In addition to defining the upper air winds for the 17 meteorological zones described above, it is necessary to specify the total (annual or seasonal, depending on the averaging period of interest) rainfall amount in each zone. These rainfall amounts are used to scale the wet scavenging coefficients used in STATMOD.

To determine the rainfall amounts, we first extracted the annual or seasonal precipitation measurements from the CADMP database using the dBase programs, described in Section 2, that calculate averages or totals for a specified averaging period. We then interpolated the CADMP precipitation amounts to a grid covering the modeling domain, using the "Kriging" technique described in Venkatram (1988). Once the gridded rainfall amounts were determined, the representative rainfall amount for each zone was estimated by averaging over all grid cells contained in the zone. We also calculated the composite annual and composite seasonal total rainfall amounts, in order to use them in the composite year and seasonal simulations.

4.4 Dry and Wet Deposition Data

Although the dry and wet deposition data are not needed for applying the model, they are required for evaluating the performance of the model and optimizing model parameters. The annually and seasonally averaged CADMP deposition data for the years 1984 through 1989 were extracted from the dBase files, using the dBase programs described in Section 2. Because most of California is dry during the summer, we did not create summer averages of the wet deposition data. In addition to the individual year data, we also prepared files for the composite year and composite seasons by averaging the data for all seasons and all years in the database.

5.0 EVALUATION OF MODEL WITH CADMP DATA

In this section, we describe the application and evaluation of STATMOD in California for the years 1984 through 1989. The primary objective of this phase of the study was to use the CADMP data to optimize the model parameters before using the model to develop source-receptor relationships for California.

We modified the model for its application to California. We also developed versions of the model that could run on a PC or on an IBM RISC 6000 workstation. The primary reason for developing a workstation version was to exploit the graphical capabilities of the workstation to visualize model results. We adapted ENSR's visualization package for this study to:

- generate scatter plots of model estimates versus observations, with a statistical summary of model performance;
- generate spatial distributions of model performance on a map of California by placing color coded symbols at measurement points; and
- display source-receptor relationships.

The visualization of the results allowed us to assess the performance of the model, to identify outliers, and to isolate receptors where the model consistently performed poorly.

Before we describe the model results, it is useful to discuss the model parameters, their expected values, and the values that resulted in the best overall performance against the CADMP data.

5.1 Model Parameters

All the model parameters have physical significance, because they correspond to physical or chemical processes. In general, the parameters depend on the season, the region where the model is being applied, and the averaging period for model computations. We can assign initial values to most of these parameters on the basis of our past experience in applying the model as well as our knowledge of the underlying processes and the modeling region. However, because of regional variations, we expect the model to perform better if some of the parameters are adjusted (within their allowable or expected range of values) to reflect these variations. Note that it is not necessary to adjust all the model parameters — only those

parameters which are expected to show a large regional variation or to have the largest influence on the results need to be examined.

The relevant model parameters have already been discussed in our description of STATMOD (Section 3). In this section, we discuss the parameter values that were used in our simulations. Note that some of the parameters were kept fixed at the initial specified values for all simulations, because our experience with STATMOD has shown that the model results are not very sensitive to these parameters. Other parameters were adjusted based on the comparison between model estimates and observations. The sensitivity of the model to the parameters is described separately in Section 7.

For this study, we did not attempt to optimize the dry deposition velocities or the wet scavenging coefficients of the various species. These parameters were kept constant for all the simulations. The dry deposition velocity of SO_2 was set to 0.5 cm/s. This value is half the expected daytime value of about 1 cm/s, and represents an average of daytime and nighttime values. Similarly, the deposition velocity of sulfate particles was set to 0.05 cm/s, about half the daytime deposition velocity of submicron particles. The deposition velocities of NO_x and nitrates were set to 0.1 and 0.8 cm/s, respectively. The nitrate deposition velocity represents a combination of the dry deposition velocity of gas-phase nitric acid and aerosol nitrate.

We also kept the wet scavenging rates fixed for all the simulations. The values of the rates we used are consistent with our understanding of the solubility and reactivity of the various species and with our past experience in applying the model to the eastern United States (ENSR, 1989). The wet scavenging rates for both SO_2 and $SO_4^=$ were set to 10^{-4} s^{-1}. The relatively high scavenging rate for SO_2 implicitly accounts for the fact that dissolved SO_2 in clouds is rapidly converted to sulfate by available H_2O_2. Recall from the discussion in Section 3 that the availability of H_2O_2 is itself accounted for in the model through the formulation of the governing equations. Because NO_x is relatively insoluble, its wet removal rate was set to 10^{-9} s^{-1}. The nitrate wet removal rate was set to 6×10^{-5} s^{-1}.

The gas-phase SO_2 oxidation rate in the wet state, k_w, was fixed at one percent/hr for all the simulations. Similarly, the SO_2 oxidation rate in non-precipitating clouds, k_d' was also fixed at two percent/hr (however, the aqueous-phase oxidant, H_2O_2 was allowed to have a seasonal variation, as described below).

The rest of the model parameters were optimized by comparing model estimates with observations. Table 5-1 presents the optimized parameter values. Note that these values are well within their expected range. For example, the gas-phase SO_2 conversion rate is of the

order of 1 percent/hr (with some seasonal variations). This rate is consistent with typical OH concentrations in the atmosphere (e.g., Seinfeld, 1986) and the rate constant for the SO_2 + OH reaction, the predominant pathway for sulfate production in the gas-phase. Similarly, the summertime concentration of the aqueous-phase oxidant, H_2O_2, is 5 $\mu g/m^3$, or approximately 3.5 ppb, while the winter value is 0.8 $\mu g/m^3$, or approximately 0.6 ppb. The spring and fall values are in between the winter and summer values. These values and seasonal variation are consistent with atmospheric H_2O_2 measurements (e.g., Sakugawa et al., 1990).

TABLE 5-1

Optimized Parameter Values

Parameter	Year	Winter	Spring	Summer	Fall
Z_i (m)	600	550	650	670	670
τ_d (hrs)	100	80	80	100	100
τ_w (hrs)	10	20	20	8	20
χ_o ($\mu g/m^3$)	4.0	0.8	2.0	5.0	2.0
k_{dSO2} (%/hr)	1	0.8	0.9	1.2	0.9
k_{dNOx} (%/hr)	8	6	9	9	9

It is important to note that we did not change model parameters from one year to another to obtain the best fit for each year separately. The parameter values selected provided the best overall fit for the six years that we simulated. Once we had decided on the final parameter values, these values were used for simulating all the years. Similarly, the seasonal parameters were not allowed to vary from year to year.

5.2 Evaluation with Wet Deposition Data

We spent significantly more time evaluating and optimizing the model using the CADMP wet deposition database than with the dry deposition database. There were several reasons for this. First, we had six years of wet deposition data (1984 through 1989), and only two years of dry deposition data (1988 and 1989). Second, there were fewer sites in the dry deposition database (less than 10) as compared to the wet deposition data (approximately 30). Finally, the database management task (see Section 2) had revealed some inconsistencies in the dry

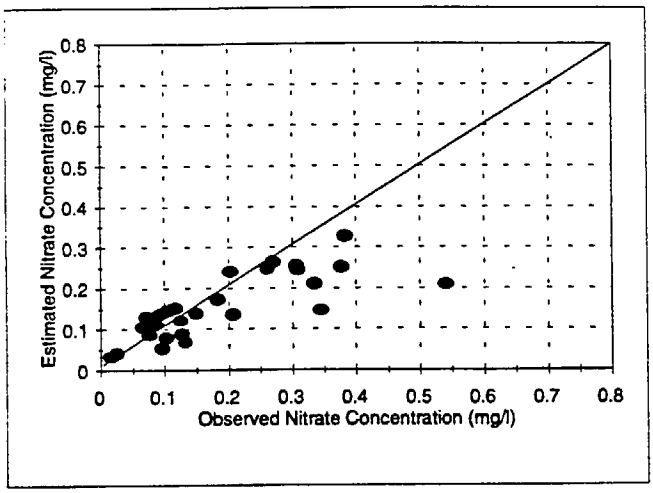
deposition database. Thus, we will first describe the evaluation of the model with the CADMP wet deposition data.

In our first application of the model, we assumed uniform winds (i.e., constant wind speeds and equal frequency from all directions) and precipitation in all the 17 meteorological zones discussed in Section 4. Then, we repeated the simulations using the actual meteorology of the 17 zones. The objective of these studies was to determine the importance of meteorology for long-term source-receptor relationships. The results from this exercise, described in one of our quarterly progress reports, showed that meteorology played a very important role in the performance of the model. The model results with actual wind roses were much better than those with uniform winds. This is not surprising in view of the fact that there are large variations in upper air winds over California.

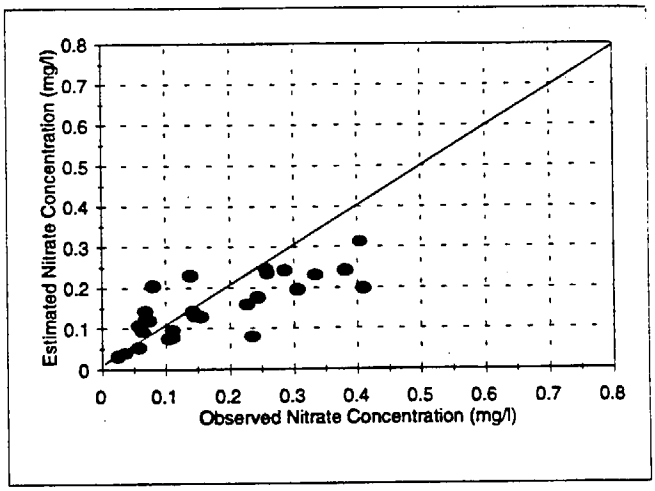
Our examination of the results with the actual meteorology showed that model estimates of pollutant concentrations in rain deviated more from observations at receptors that were downwind of major sources in meteorological zones where upper air data were not available from NCDC or other sources. Recall from Section 4 that we used the ARB historical summaries of upper air data for these zones.

Because our first exercise had shown the important role played by meteorology in the performance of the model, we inferred that model performance could be improved by making slight adjustments to the wind roses in zones where the historical summaries were used. We also adjusted some of the model parameters (see Table 5-1 for the final parameter values) to obtain a better fit against the observed concentrations. The visualization software that we developed on the workstation was invaluable in this task and in examining the role of sea-salt in influencing sulfur concentrations in rain, described later in this section.

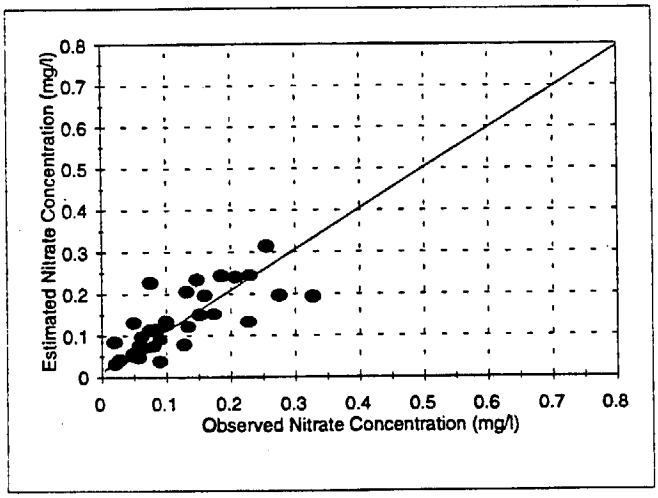
Figure 5-1 shows the comparison of annually averaged observations of nitrate concentrations in rain with model estimates for the six years (1984 through 1989) that were simulated in this study. The results of the simulations for nitrate are presented in Figure 5-1. As shown in the figure, there is good agreement between model estimates and observations at most sites for all the years that were simulated. The largest disagreement occurs at only one site, Victorville, where the nitrate concentration was underestimated in three out of the six years. The reasons for the large underestimation at the Victorville site are not clear at this time. However, it is useful to note that the observed nitrate concentrations in rain at the Victorville site are much larger than the nitrate concentrations measured elsewhere in the CADMP network for the three years when the model underestimates nitrate concentrations at Victorville.



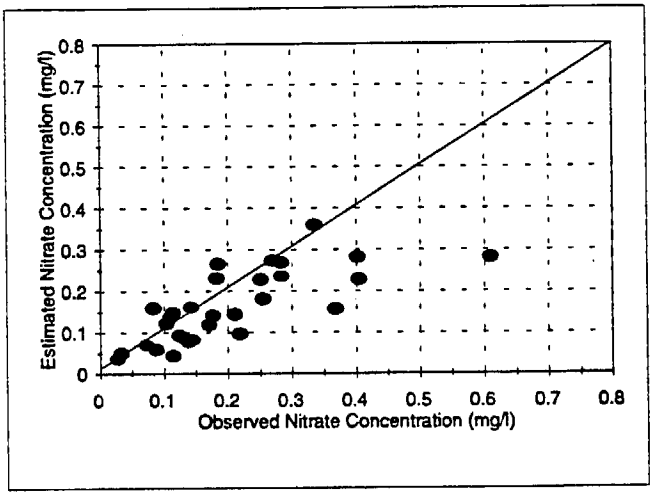
1984



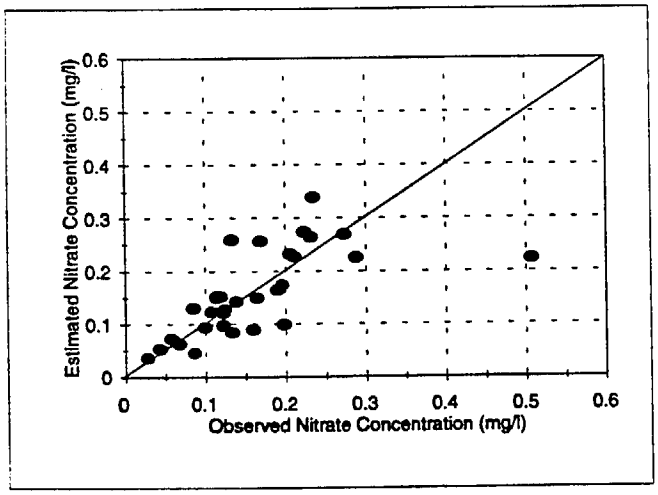
1985



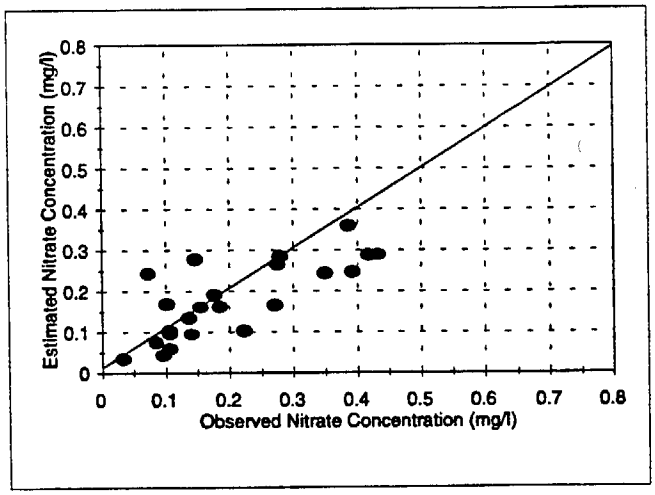
1986



1987



1988



1989

FIGURE 5-1. Comparison of Annually Averaged Observations of Nitrate Concentrations in Rain with Model Estimates for Six Years (1984-1989)

Table 5-2 provides the model performance statistics for annually averaged nitrate concentrations in rain. We see that the model explains more than 50 percent of the variance in the observations for five out of the six years. The model performance statistics improve if we remove the Victorville site, where the model consistently underestimates the observednitrate concentrations in rain. For example, the r^2 for 1988 increases from 0.44 to 0.66 if we exclude the Victorville site.

TABLE 5-2

Model Performance Statistics for Nitrate Concentrations in Rain

Year	$\bar{C}_o$	$\bar{C}_e$	$\bar{C}_r$	a	b	r^2	S.E.
1984	0.19	0.15	0.04	0.60	0.046	0.73	0.04
1985	0.18	0.15	0.03	0.48	0.064	0.61	0.05
1986	0.12	0.13	-0.01	0.70	0.047	0.55	0.05
1987	0.20	0.16	0.04	0.50	0.058	0.55	0.06
1988	0.16	0.16	0.00	0.58	0.063	0.44	0.06
1989	0.21	0.18	0.03	0.57	0.058	0.56	0.06

$$C_e = a\,C_o + b$$

$$\bar{C}_o = \text{Mean Obs. Conc.} = \frac{1}{n}\sum_{i=1}^{n} C_{oi}$$

$$b = \text{Intercept} = \frac{1}{n}\left(\sum_{i=1}^{n} C_{ei} - a\sum_{i=1}^{n} C_{oi}\right)$$

$$\bar{C}_e = \text{Mean Est. Conc.} = \frac{1}{n}\sum_{i=1}^{n} C_{ei}$$

$$r = \text{Corr. Coeff.} = a\,\frac{\textit{standard deviation observed}}{\textit{standard deviation predicted}}$$

$$\bar{C}_r = \text{Mean Residual} = \frac{1}{n}\sum_{i=1}^{n}(C_{oi} - C_{ei})$$

$$\text{S.E.} = \text{Standard Error} = \frac{s}{\left(\sum_{i=1}^{n} C_{oi}^2\right)^{1/2}}$$

$$a = \text{Slope} = \frac{\sum_{i=1}^{n} C_{oi}C_{ei} - \frac{1}{n}\sum_{i=1}^{n} C_{oi}\sum_{i=1}^{n} C_{ei}}{\sum_{i=1}^{n} C_{oi}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} C_{oi}\right)^2}$$

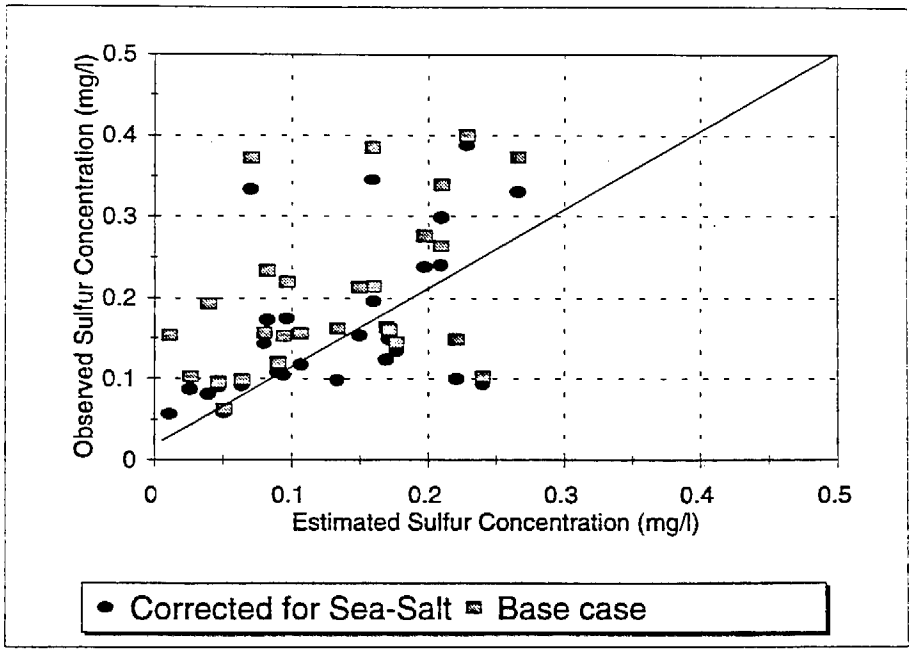
$$s^2 = \text{Variance} = \frac{1}{n-1}\sum_{i=1}^{n}(C_{oi}-C_{ei})^2$$

We found that sulfate concentrations in rain were consistently underestimated for all the six years. The performance was not improved even by making large changes to the model parameters. Moreover, a comparison of estimated ambient concentrations of SO_2 and sulfate concentrations with observed concentrations, described later in this section, indicated that both ambient SO_2 and sulfate, and thus total sulfur, were being underestimated.

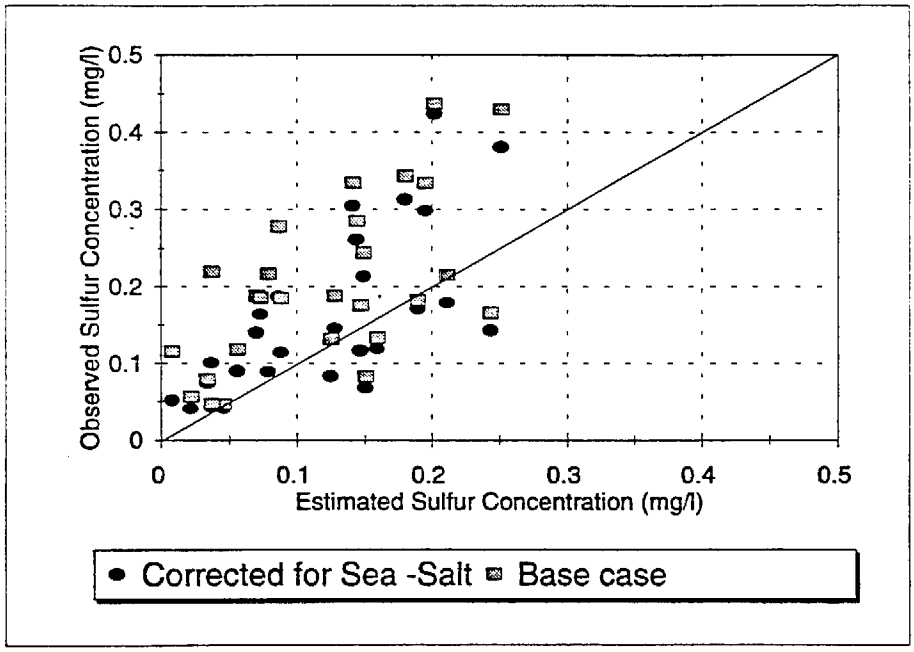
The performance of the model for nitrate concentrations suggested that pollutant transport was being handled correctly in the model. Thus, the finding that total sulfur was underestimated indicated to us that the SO_2 emissions used in the model were probably underestimated. Assuming that the California anthropogenic SO_2 emission inventory provided by the ARB was reasonably accurate, this suggested that sources not included in the inventory, such as natural sources (e.g., sea-salt or wind blown soil dust) or anthropogenic sources from outside the modeling domain (e.g., other states or Mexico) were contributing to observed sulfur concentrations in rain and air.

We first accounted for the contribution of sea-salt to sulfate in precipitation by examining the concentrations of magnesium and sodium ions. We used the chemical composition of sea-water to relate the concentrations of these ions to sulfate ion concentrations. After estimating the sea-salt sulfate contributions, we subtracted it from the observed sulfate concentrations for comparison against model estimates. Figure 5-2 shows a comparison of observed and estimated concentrations of annually averaged sulfur concentrations in rain for the years 1984 and 1985, with and without the sea-salt correction. As can be seen in the figure, the agreement between observed and estimated sulfur concentrations in rain is better when the sea-salt contribution is accounted for. As expected, the sea-salt contribution was found to be more important at coastal sites than at inland sites. The results for the other years are similar.

Figure 5-3 compares the estimated sulfur concentrations in rain with observed concentrations (minus the sea-salt contribution) for all the years. The performance statistics are shown in Table 5-3. Although accounting for the sea-salt contribution improves model performance, we see that the model still tends to underestimate the observed concentrations, by an average of about 30 percent. Further examination of the results showed that the model tended to underestimate concentrations at receptors in agricultural regions (e.g., in the San Joaquin Valley), and at remote receptors in northern California near the borders of California with Nevada and Oregon. The largest underestimates occurred at Gasquet, S. Lake Tahoe, Bakersfield, Lakeport, and San Rafael.

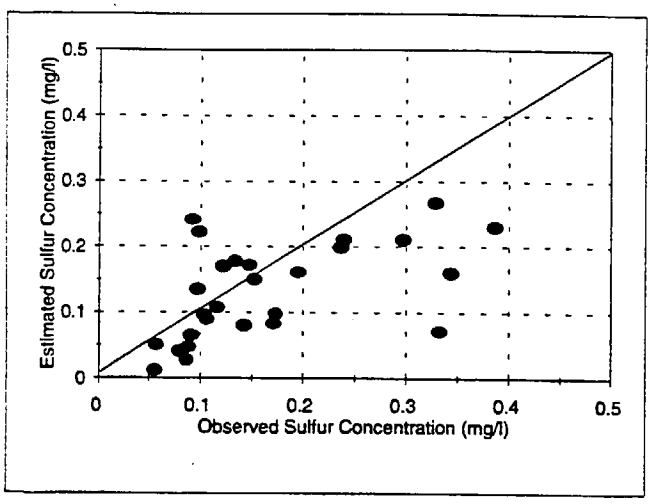


1984

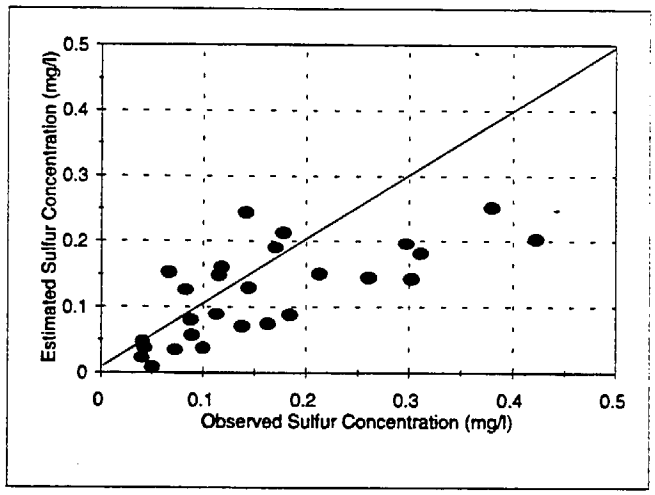


1985

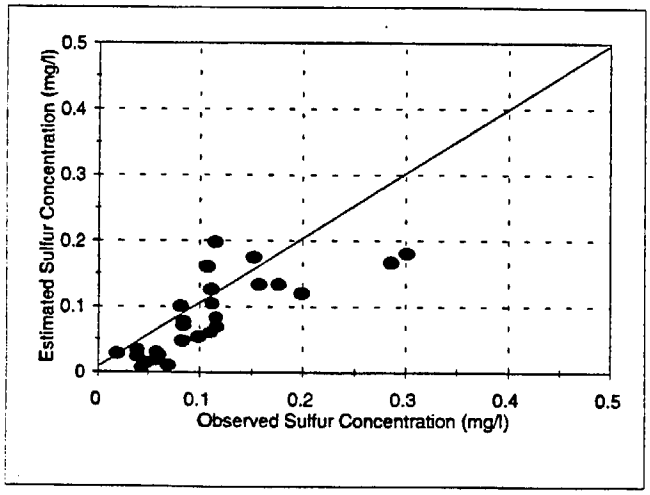
FIGURE 5-2. Comparison of Estimated Sulfur Concentrations in Rain with Observed Concentrations (with and without Sea Salt Correction) for the Years 1984 and 1985



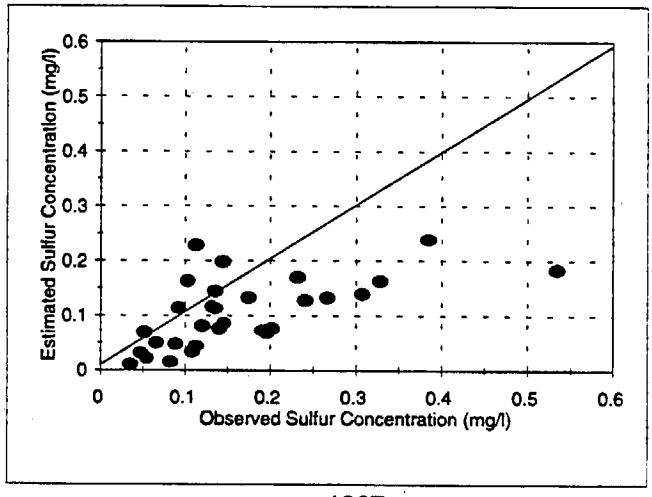
1984



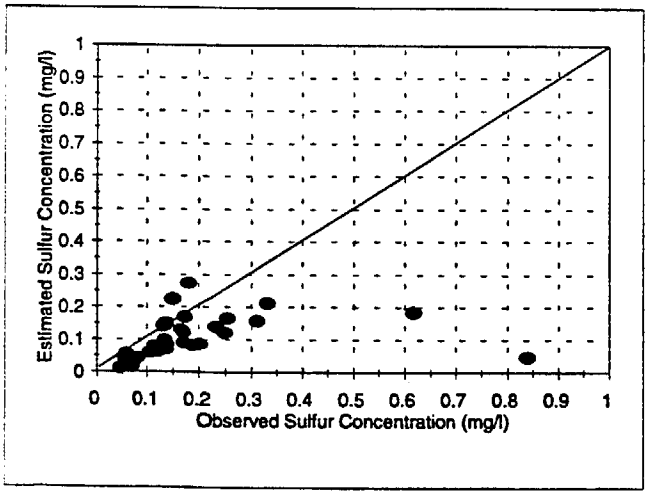
1985



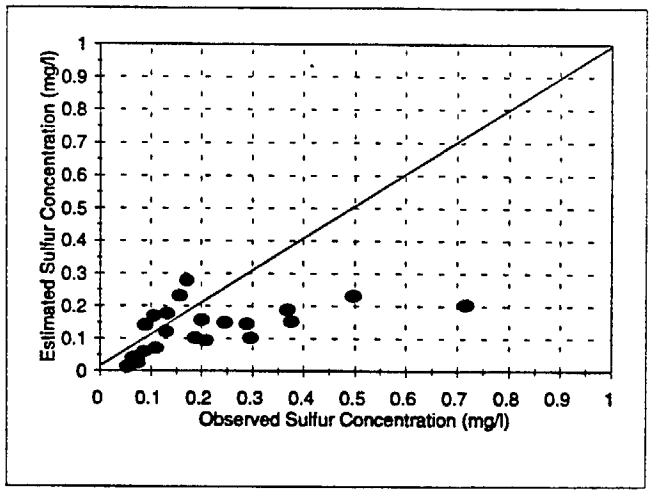
1986



1987



1988



1989

FIGURE 5-3. Comparison of Annually Averaged Observations of Non-Sea Salt Sulfur Concentrations in Rain with Model Estimates for Six Years (1984-1989)

TABLE 5-3

Model Performance Statistics for Sulfur Concentrations in Rain

Year	$\bar{C}_o$	$\bar{C}_e$	$\bar{C}_r$	a	b	r^2	S.E.
1984	0.17	0.13	0.04	0.41	0.060	0.29	0.06
1985	0.16	0.12	0.04	0.46	0.046	0.49	0.05
1986	0.12	0.09	0.03	0.64	0.015	0.64	0.04
1987	0.17	0.10	0.07	0.44	0.045	0.39	0.05
1988	0.20	0.11	0.09	0.11	0.088	0.07	0.06
1989	0.21	0.13	0.08	0.26	0.073	0.32	0.06

$$C_e = a\,C_o + b$$

$$\bar{C}_o = \text{Mean Obs. Conc.} = \frac{1}{n}\sum_{i=1}^{n} C_{oi}$$

$$b = \text{Intercept} = \frac{1}{n}\left(\sum_{i=1}^{n} C_{ei} - a\sum_{i=1}^{n} C_{oi}\right)$$

$$\bar{C}_e = \text{Mean Est. Conc.} = \frac{1}{n}\sum_{i=1}^{n} C_{ei}$$

$$r = \text{Corr. Coeff.} = a\,\frac{\textit{standard deviation observed}}{\textit{standard deviation predicted}}$$

$$\bar{C}_r = \text{Mean Residual} = \frac{1}{n}\sum_{i=1}^{n}(C_{oi} - C_{ei})$$

$$\text{S.E.} = \text{Standard Error} = \frac{s}{\left(\sum_{i=1}^{n} C_{oi}^2\right)^{1/2}}$$

$$a = \text{Slope} = \frac{\sum_{i=1}^{n} C_{oi}C_{ei} - \frac{1}{n}\sum_{i=1}^{n} C_{oi}\sum_{i=1}^{n} C_{ei}}{\sum_{i=1}^{n} C_{oi}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} C_{oi}\right)^2}$$

$$s^2 = \text{Variance} = \frac{1}{n-1}\sum_{i=1}^{n}(C_{oi} - C_{ei})^2$$

The results for the receptors near the border suggest that emissions from the neighboring states may be contributing to the observed concentrations at these receptors. Furthermore, the observed concentrations at some of these receptors are small (of the order of 0.1 mg/ℓ of sulfur), and appear to be background values. These receptors are apparently unaffected by anthropogenic emissions of SO_2 in California.

Of more concern is the underestimation of concentrations at sites well within the modeling domain, such as Bakersfield, where sulfur concentrations in rain are consistently underestimated by the model. An analysis of the wet deposition data shows a high degree of correlation between sulfate and calcium ion concentrations at these sites. These results suggest that local emissions of natural soil dust (containing $CaSO_4$), which are not included in the emission inventory used in our simulations, may be contributing to observed sulfur concentrations in rain.

Although we used some statistical analysis techniques, such as principal component analysis (PCA), to examine the role of wind-blown soil dust on sulfur concentrations in rain, this analysis was inconclusive. Other methods, such as receptor modeling, are likely to yield more information if accurate source profiles are available to apply these methods. However, such data are not readily available. Furthermore, a detailed receptor modeling study was outside the scope of this project. Thus, we addressed the issue by regressing the difference between the observed and estimated sulfur concentrations in rain with the measured calcium concentration at each receptor site. Since there were only six years of data available, this regression could not be performed with the annual averages. We performed the regression using the seasonally averaged concentrations, described below.

For the seasonal simulations, we used the following convention: winter corresponded to December, January, and February; spring corresponded to March, April, and May; summer corresponded to June, July, and August; and fall corresponded to September, October, and November. We created the seasonally averaged wind roses for regions where upper air data were available, and used the annual ARB historical summaries for the other regions. As in the case of the annual simulations, we found that model performance improved when we adjusted the wind roses for the regions where upper air data for the years of interest were not available.

Note that because information on the seasonal variation of the SO_2 and NO_x emissions was not available, we used the annual emission rates in our simulations. Although this is a possible source of error, we did not feel that seasonal variations in emissions would be large enough to influence model results.

The results of the simulations (after optimization of model parameters) for the winter, spring, and fall seasons of the years 1985 and 1988 for nitrate and sulfate concentrations in rain, are presented in Figures 5-4 and 5-5, respectively. As discussed in Section 4, we did not have enough precipitation chemistry data in summer to evaluate the model results.

Figure 5-4 shows that, as in the case of the annually averaged data sets, there is generally good agreement between STATMOD estimates and observed concentrations of nitrate in rain, except in spring 1985 and fall 1988, when some of the higher observed nitrate concentrations are severely underestimated.

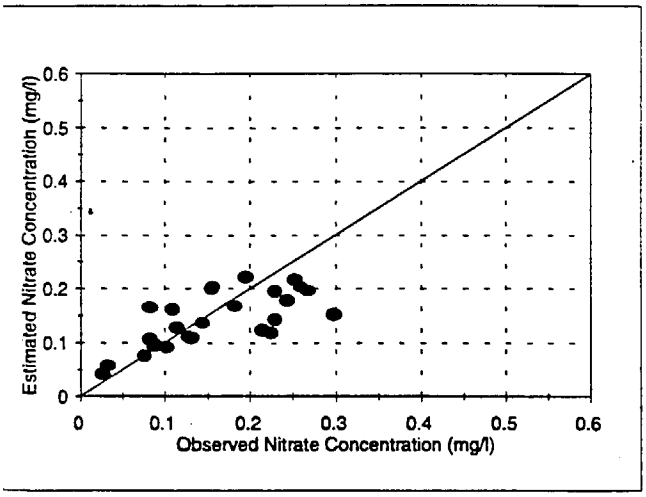
Figure 5-5 shows that the model does not perform well in estimating the seasonally averaged sulfur concentrations in rain. The estimated concentrations are consistently lower than the observed concentrations at several locations. As in the case of the annual averages, we suspect that part of the underestimation can be explained by the contribution of local natural sources not included in the emission inventory. These sources are likely to play a larger role in the seasonal simulations than in the annual simulations, because of the shorter averaging period in the former.

To estimate the contribution of wind-blown soil dust to the seasonally averaged sulfate concentrations in rain, we regressed the residuals of the observed and estimated precipitation sulfate concentrations ($C_{SO4,obs}$ - $C_{SO4,est}$) against the observed calcium concentrations in rain at those sites where the influence of wind blown soil dust appeared to be high. The results of the regression analysis are presented in Table 5-4.

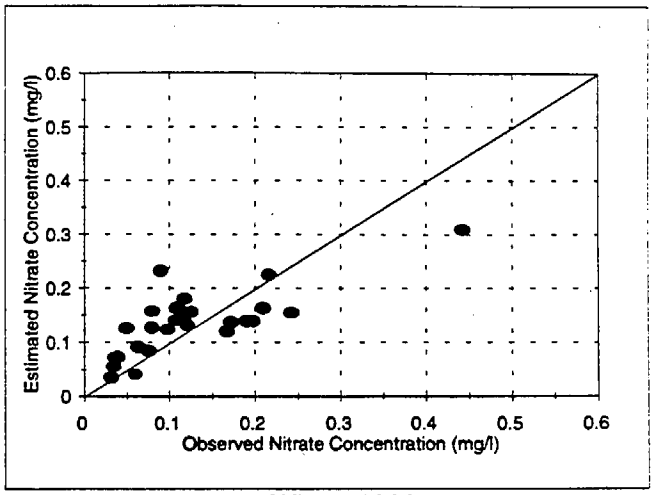
TABLE 5-4

Results from Linear Regression of ($C_{SO4,obs}$ - $C_{SO4,est}$) Against C_{Ca}

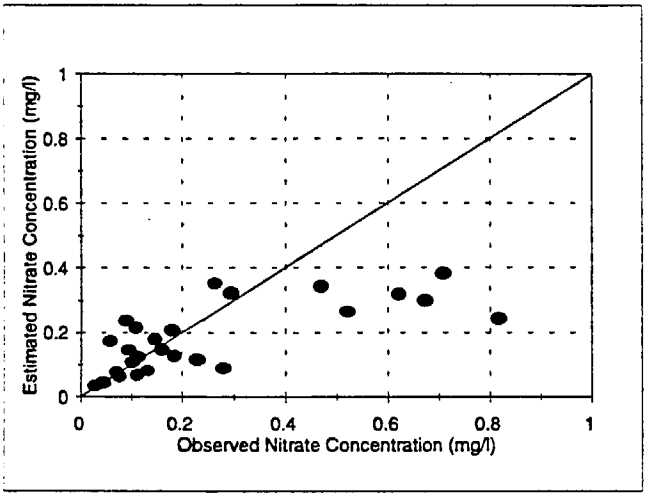
Site	Slope	Intercept (μM)	Correl. Coeff.
Ash Mountain	0.42	0.48	0.69
San Nicholas Island	1.22	-2.11	0.91
San Bernardino	0.69	-0.44	0.95
Sacramento	0.38	0.98	0.66
Lakeport	0.90	-0.96	0.80
Eureka	0.41	0.35	0.64



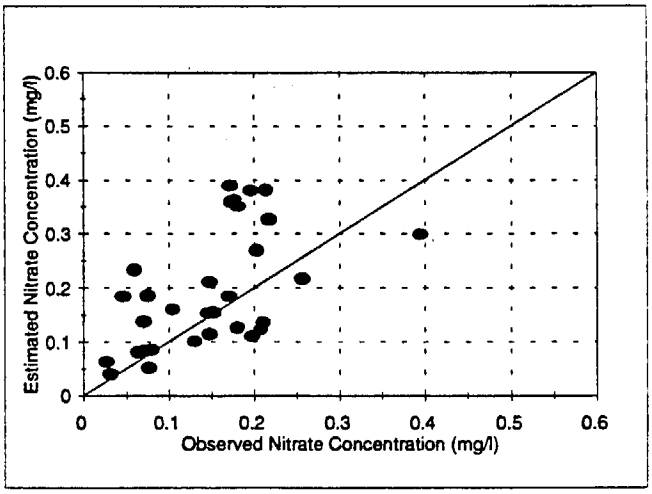
Winter 1985



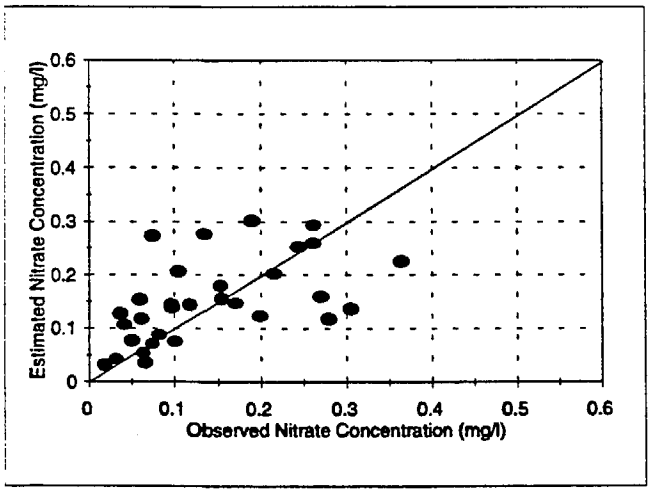
Winter 1988



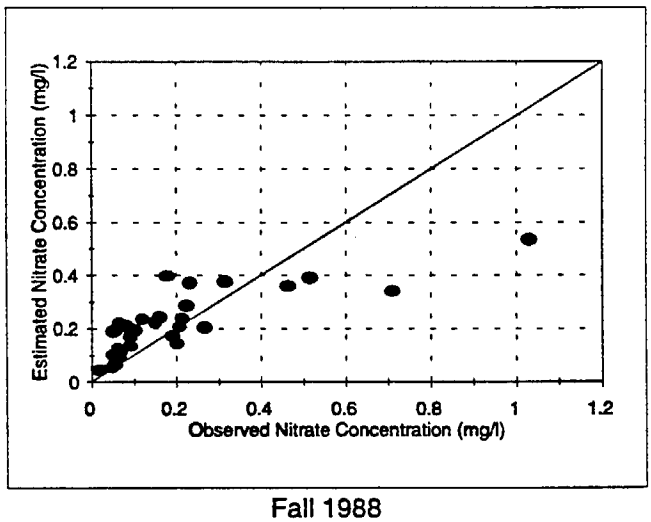
Spring 1985



Spring 1988

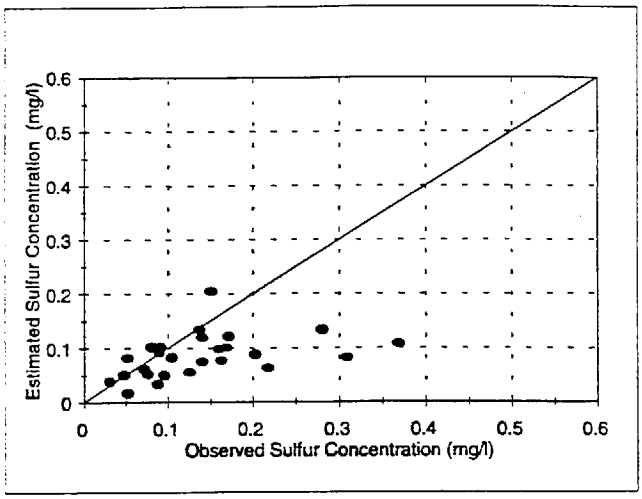


Fall 1985

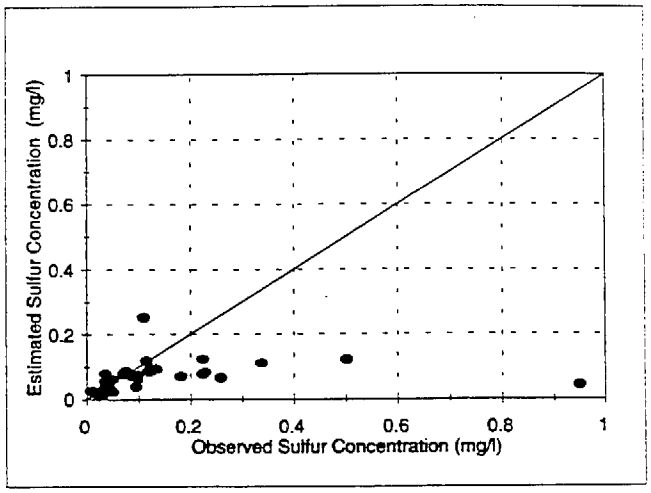


Fall 1988

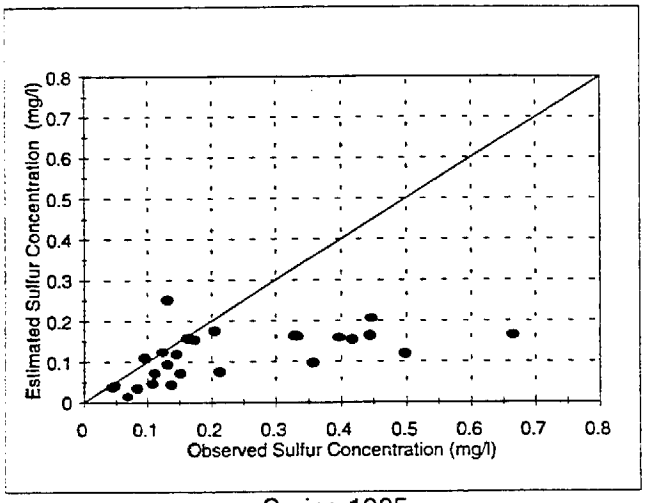
FIGURE 5-4. Comparison of Seasonally Averaged Observations of Nitrate Concentrations in Rain with Model Estimates for the Winter, Spring, and Fall Seasons of the Years 1985 and 1988



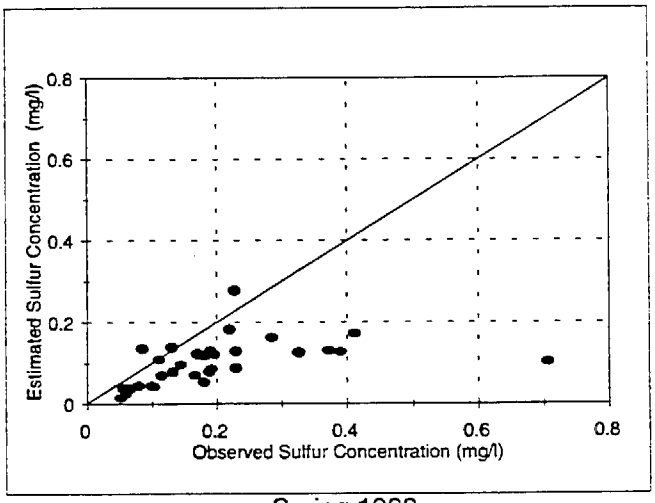
Winter 1985



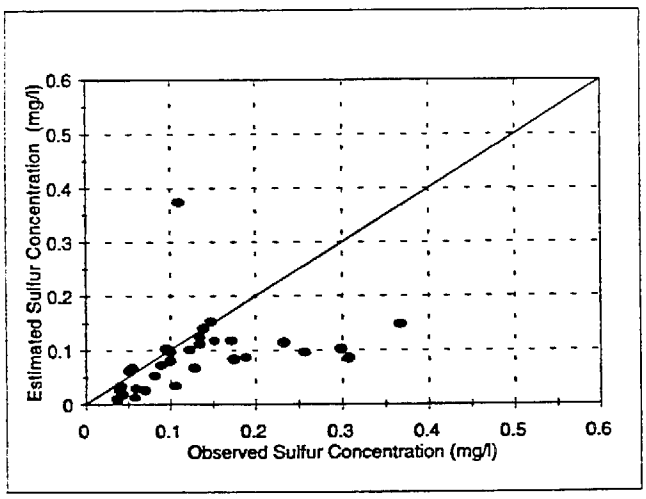
Winter 1988



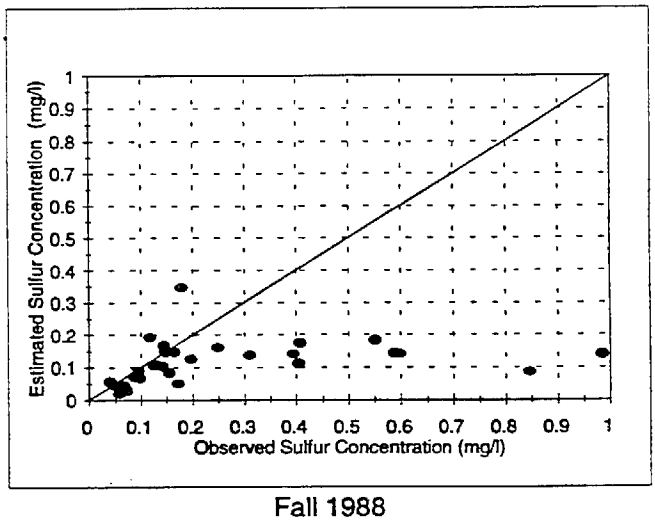
Spring 1985



Spring 1988



Fall 1985



Fall 1988

FIGURE 5-5. Comparison of Seasonally Averaged Observations of Non-Sea Salt Sulfur Concentrations In Rain with Model Estimates for the Winter, Spring, and Fall Seasons of the Years 1985 and 1988

The ideal fit would have a slope equal to 1 and intercept close to zero. As shown in Table 5-4, this is not the case in the receptor sites we examined. The correlation coefficient is high, which indicates that $CaSO_4$ is indeed a strong local source of sulfates, but the slope is less than one, suggesting that there may be other sources responsible for the high observed sulfate concentrations in rain.

Figure 5-6 compares observed sulfur concentrations in rain at the six sites with model estimates (with and without the estimated dust contribution) for all the seasons. There is an improvement in model performance, particularly for the higher observed concentrations, when the estimated dust contribution is added. A possible explanation for this behavior is that unusually high sulfate concentrations in rain may be the result of high dust concentrations, probably due to the higher wind speeds that accompany precipitation events. In most of the other cases, however, the contribution of the anthropogenic sources to the sulfate concentration in rain is much more important than that of dust.

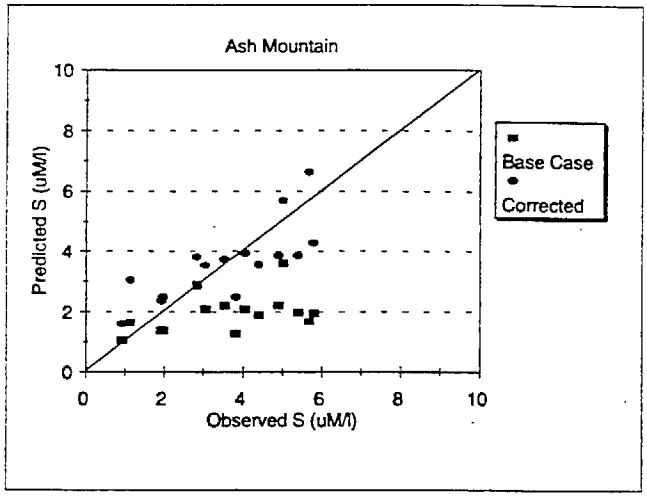
5.3 Evaluation with Dry Deposition Data

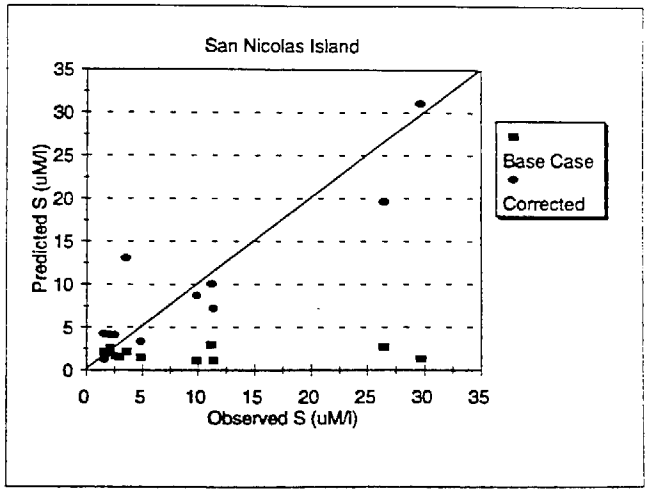
We also compared STATMOD estimates of ambient concentrations of SO_2, aerosol sulfate, and nitrates with the limited amount of CADMP dry deposition data that were available to us. Only data from May 1988 to October 1989 were used in our comparisons, because the revised database after October 1989 was not available in time for our study.

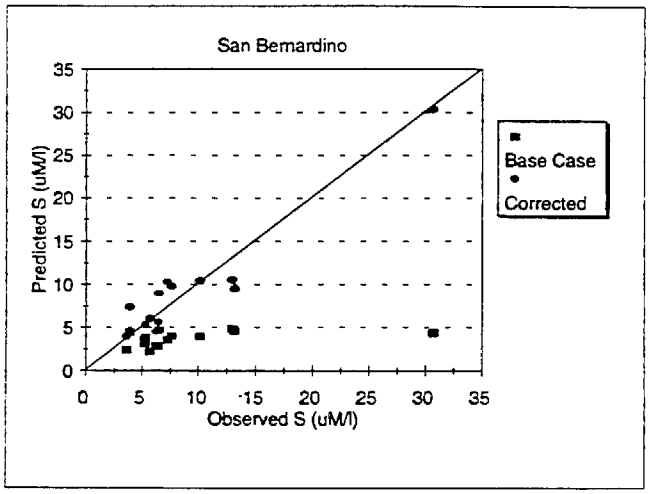
Figures 5-7 and 5-8 present the comparison between model estimates and observations for the annually averaged data for 1988 and 1989, respectively. Ambient sulfate concentrations are severely underestimated by the model, but it does a better job of estimating SO_2 and nitrate concentrations. Note that if SO_2 were not being converted to sulfate rapidly enough, then the underestimation of sulfate concentrations would be accompanied by large overestimation of SO_2 concentrations. As mentioned earlier in the discussion of the precipitation chemistry results, we suspect that local sources that are not in the emission inventory are contributing to the high sulfate levels measured at some of the receptors.

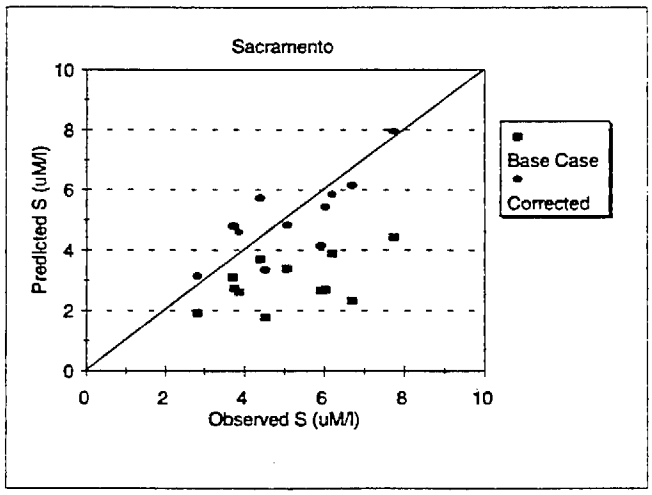
We see a similar performance for the seasonal simulations, shown in Figures 5-9 through 5-13. The model underestimates the sulfate concentrations for all five seasons for which data are available, while its performance for nitrate and SO_2 is similar to that of the annual runs.

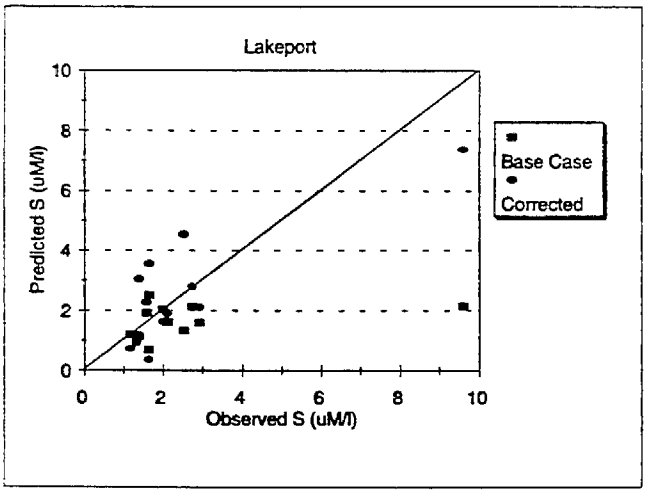
Because we only have two years of dry deposition at less than 10 sites, it is difficult to perform the kind of regression analysis that we performed with the wet deposition results to determine the role of natural or local sulfate sources that are not included in our emission inventory. This issue needs to be resolved when more data become available.











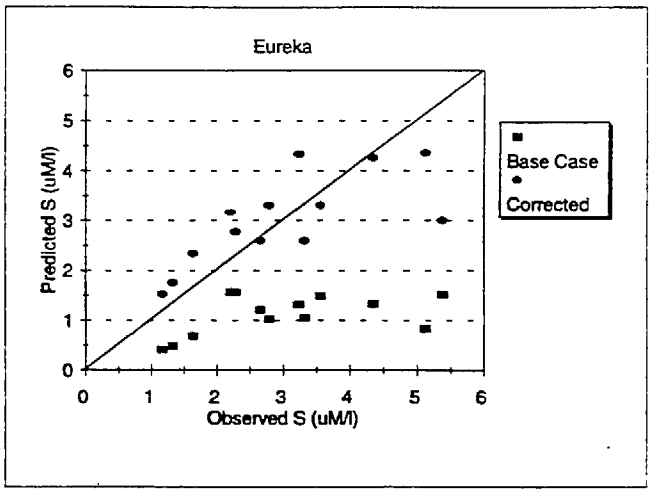


FIGURE 5-6. Comparison of Seasonally Averaged Observed Non-Sea Salt Sulfur Concentrations in Rain with Model Estimates Before and After Correction for the Soil Dust Contribution at Six Locations

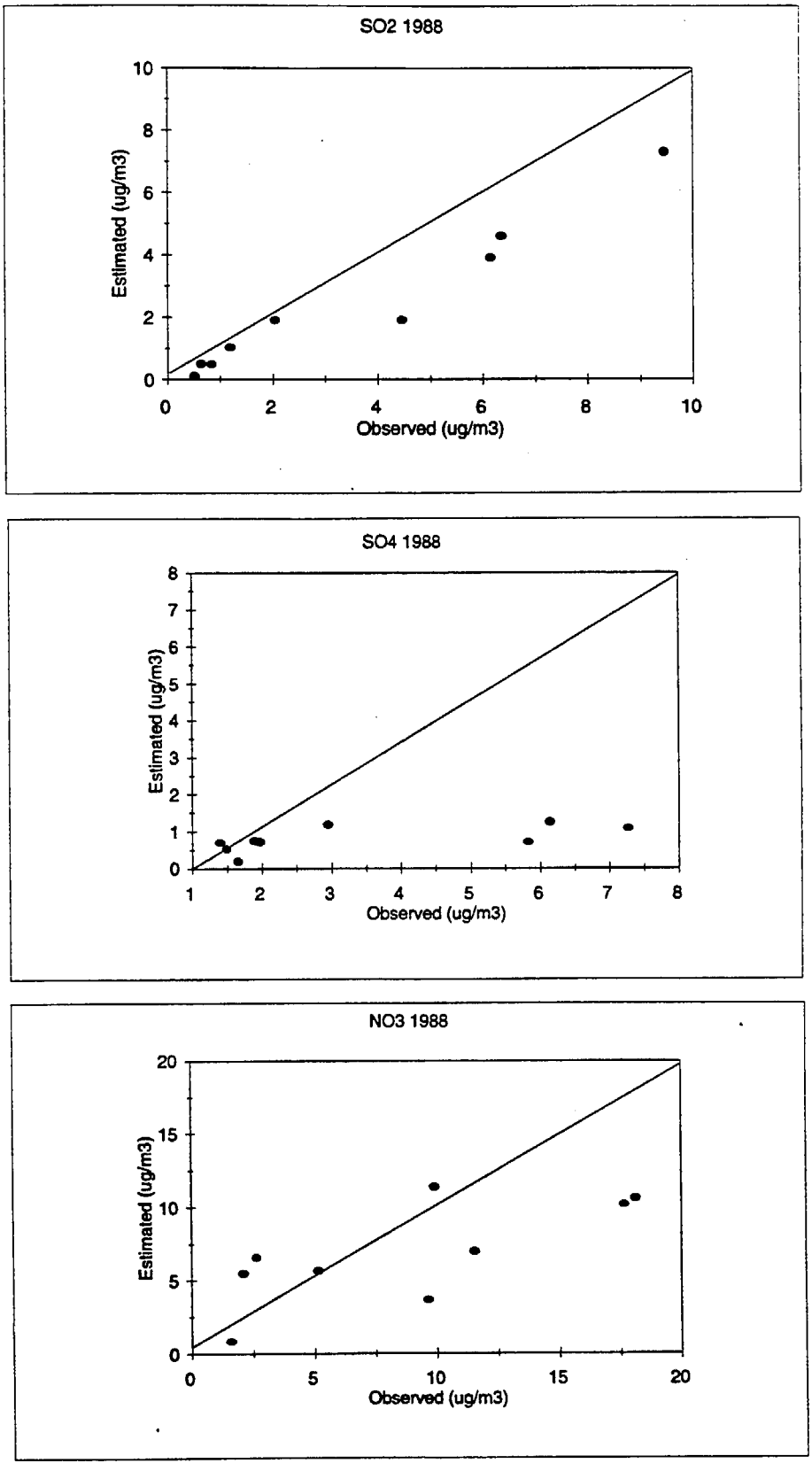
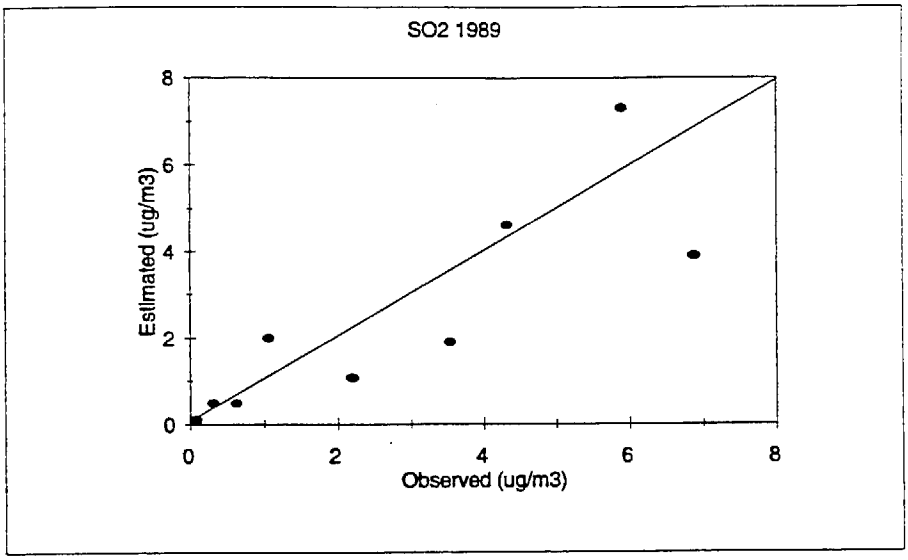
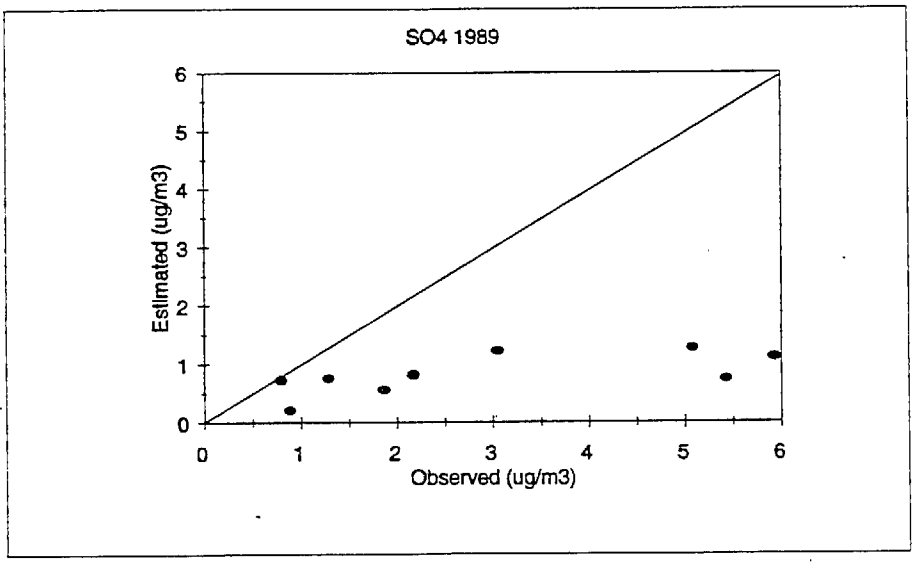


FIGURE 5-7. Comparison of Annually Averaged Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for 1988





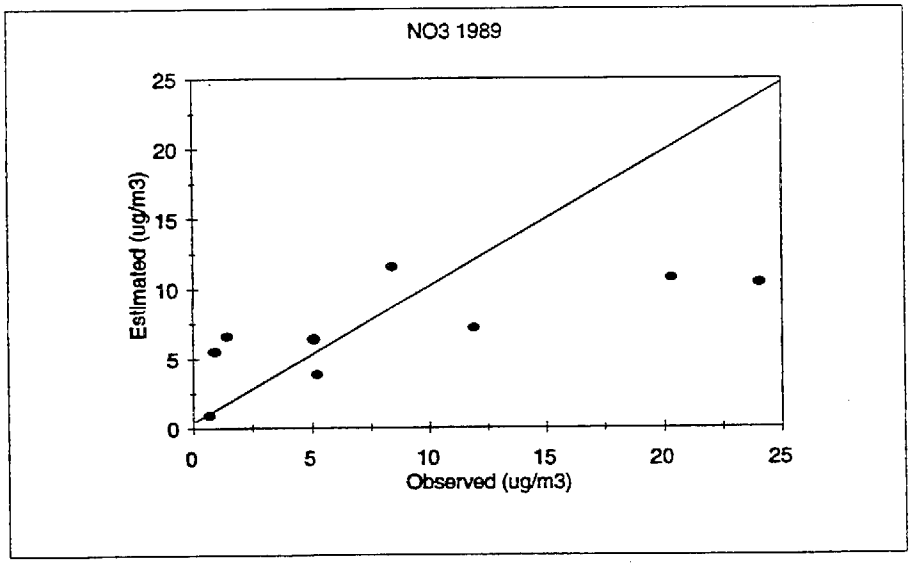


FIGURE 5-8. Comparison of Annually Averaged Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for 1989

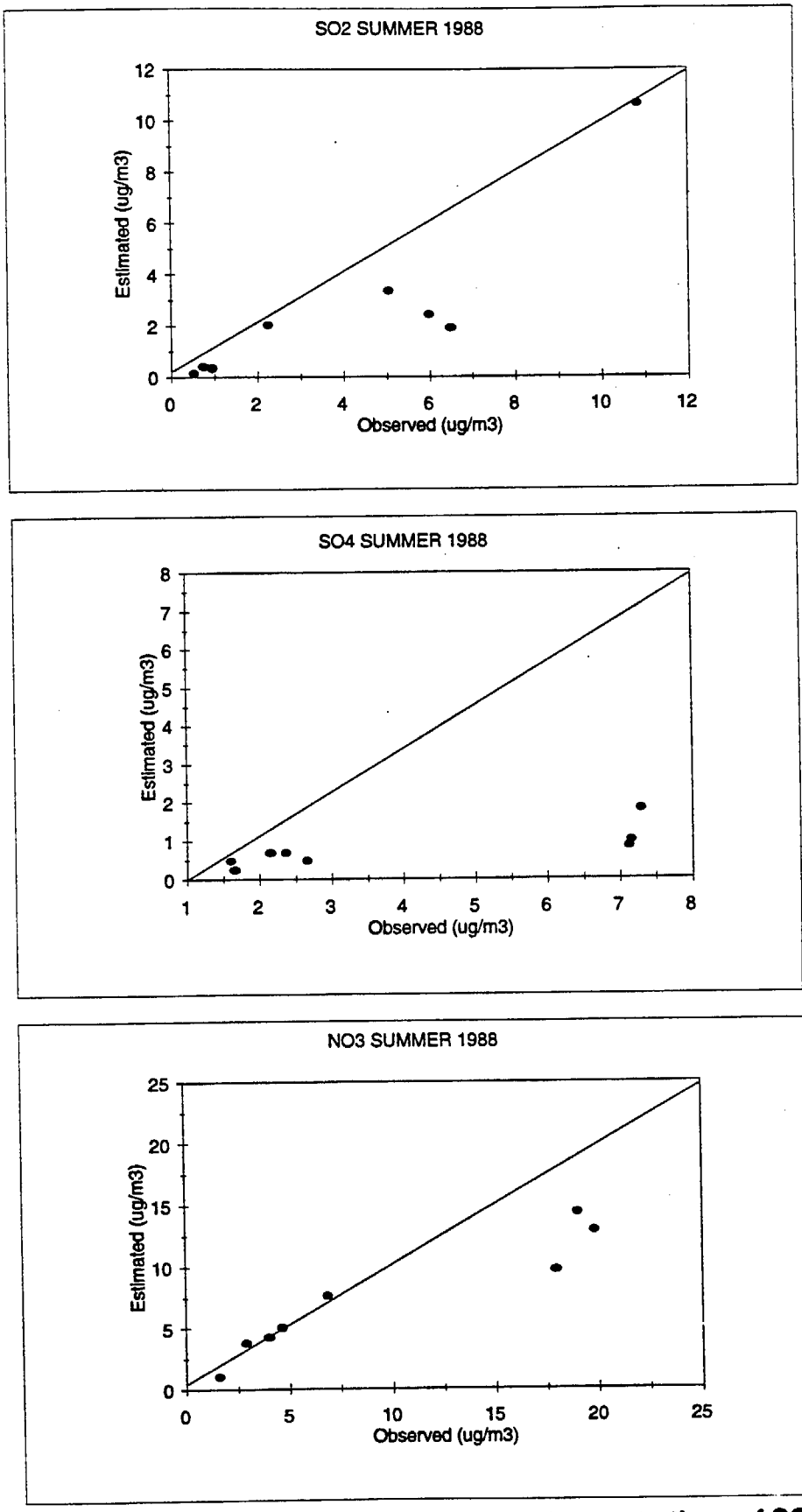
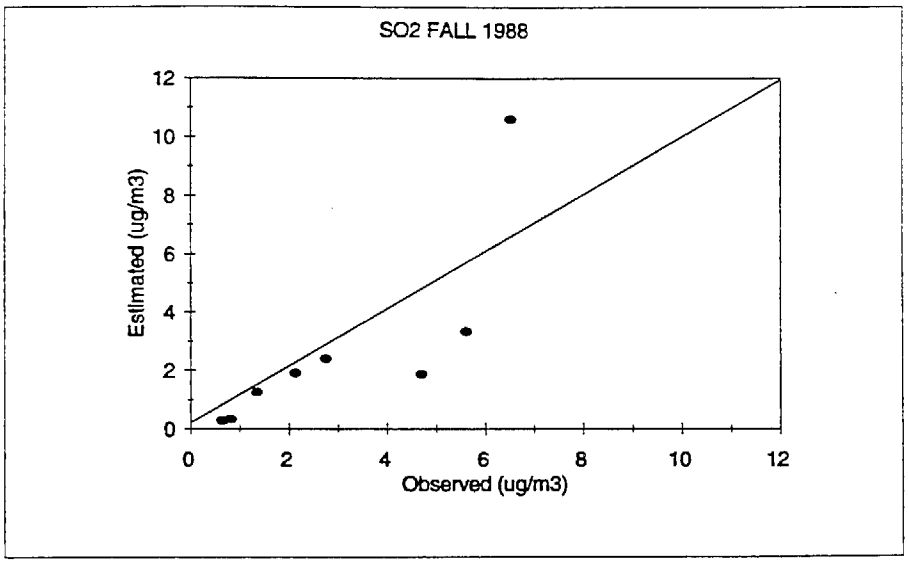
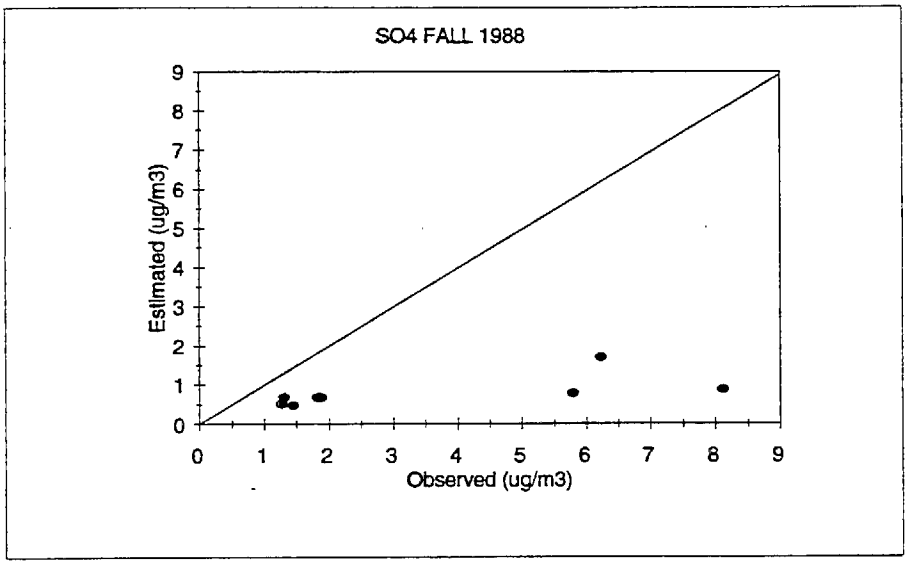


FIGURE 5-9. Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Summer 1988





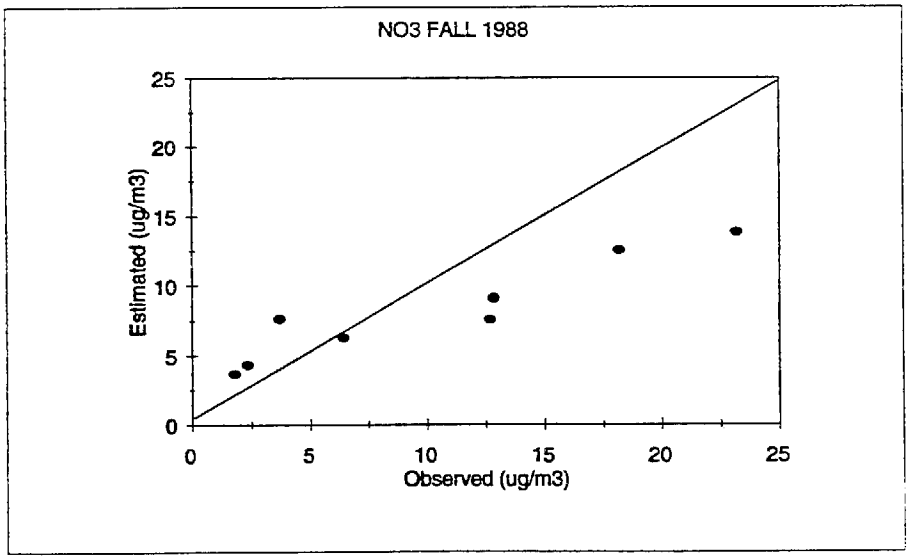


FIGURE 5-10. Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Fall 1988

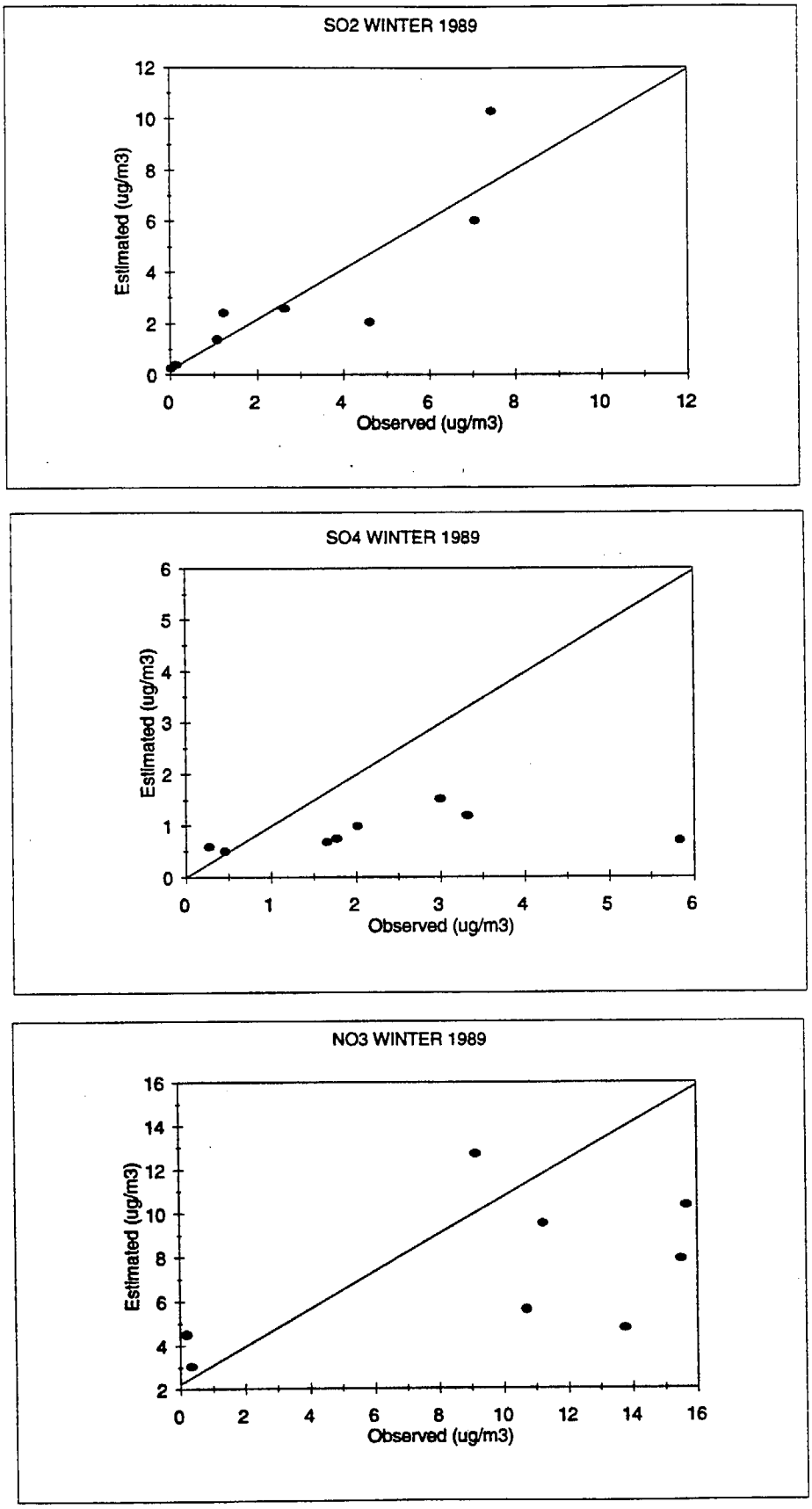
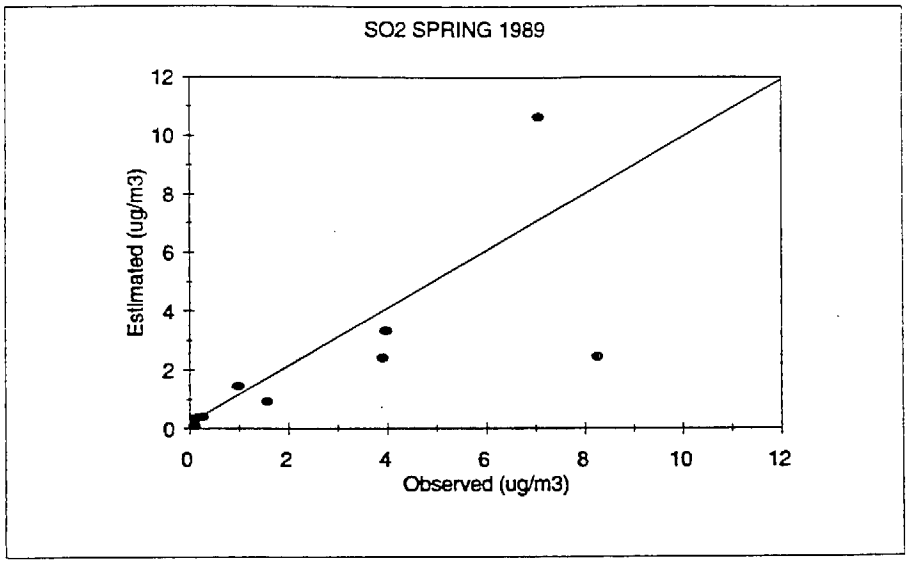
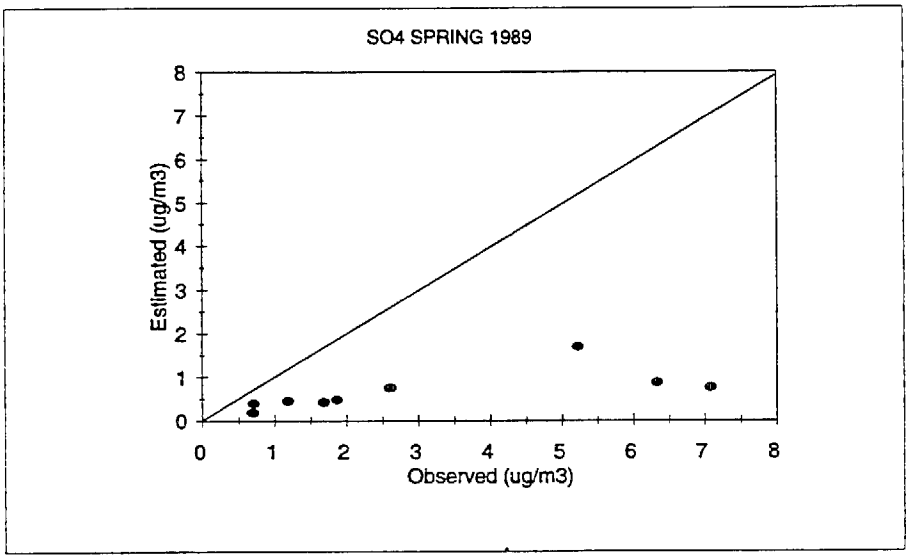


FIGURE 5-11. Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Winter 1989





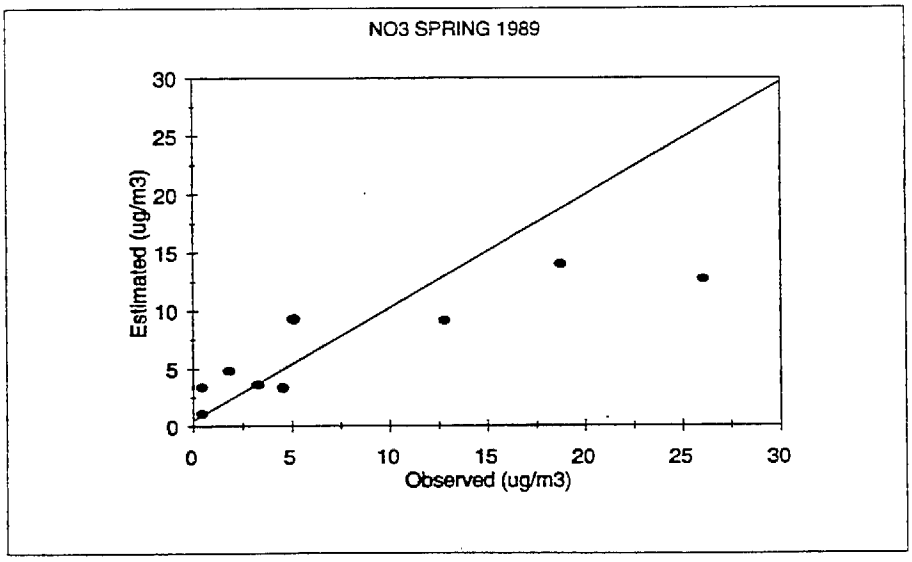
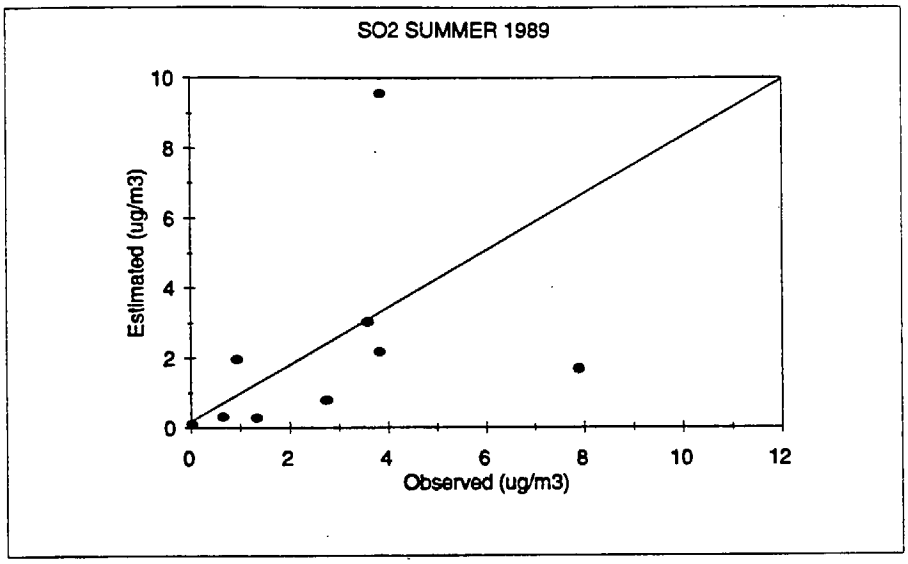
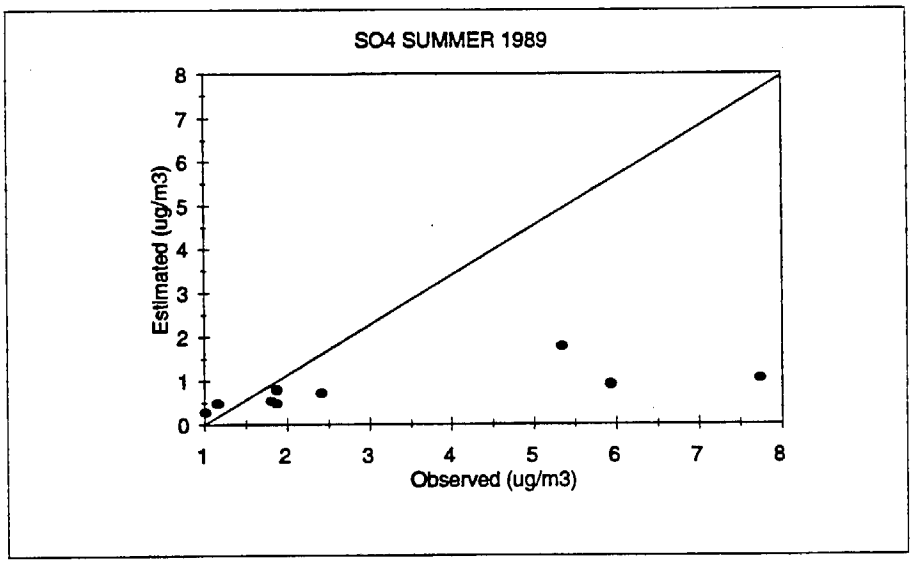


FIGURE 5-12. Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Spring 1989





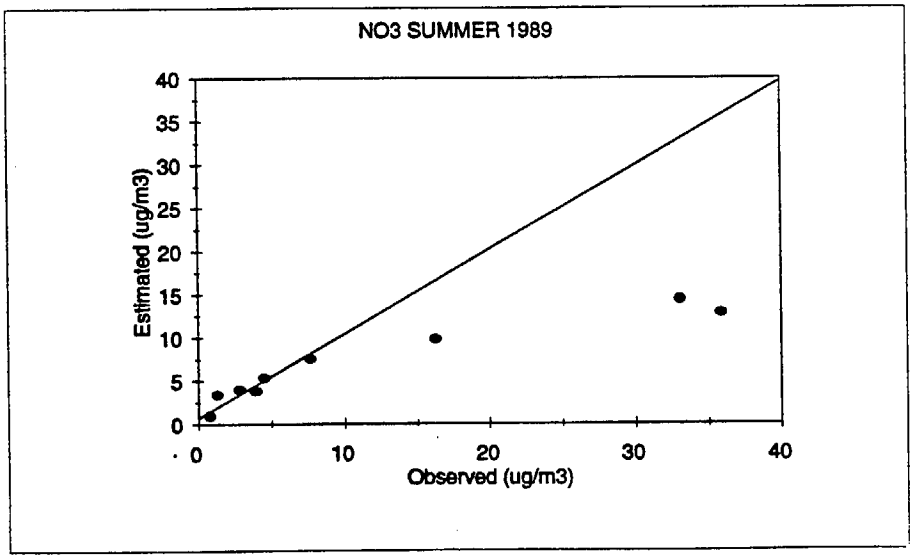


FIGURE 5-13. Comparison of Observed Ambient Concentrations of SO_2, Aerosol Sulfates, and Nitrates with Model Estimates for Summer 1989