DOE/EV/04968--5

DE82 005462

THE CHARACTERIZATION OF AMBIENT AIR POLLUTION
FOR STOCHASTIC HEALTH MODELS

MASTER

S. Batterman
M.I.T. Energy Laboratory
Cambridge, Massachusetts 02139

There is no objection from the patent
point of view to the publication or
dissemination of the document(s)
listed in this letter.

BROOKHAVEN PATENT GROUP
8/23 1981 By *UAC*

August 1981

Prepared for

THE U.S. DEPARTMENT OF ENERGY
UNDER CONTRACT NO. DE-AC02-78EV04968.A003

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
MSW

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

THE CHARACTERIZATION OF AMBIENT AIR POLLUTION
FOR STOCHASTIC HEALTH MODELS

by

STUART ARTHUR BATTERMAN

Submitted to the Department of Civil Engineering on July 16, 1981
in partial fulfillment of the requirements for the degree of
Master of Science in Technology and Policy

ABSTRACT

This research is an analysis of various measures of ambient air pollution useful in cross-sectional epidemiological investigations and risk assessments. The Chestnut Ridge area health effects investigation, which includes a cross-sectional study of respiratory symptoms in young children, is used as a case study. Four large coal-fired electric generating power plants are the dominant pollution sources in this area of western Pennsylvania. The air pollution data base includes four years of sulfur dioxide and five years of total suspended particulate concentrations at seventeen monitors.

Some 70 different characterizations of pollution are constructed and tested. These include pollutant concentrations at various percentiles and averaging times, "exceedence measures" which show the amount of time a specified threshold concentration is exceeded, and several dosage measures which transform non-linear dose-response relationships onto pollutant concentrations.

The relative ranking of geographic units in the case study change dramatically with different pollution measures. Rankings also change significantly from year to year. Pollution measures are evaluated for stability using coefficient of variation and analysis of variance tests. Measures are also evaluated for their representativeness, in terms of explaining known changes in pollution patterns.

Exceedence measures at low thresholds are found to be stable and quite insensitive to averaging time. Exceedence measures and mean annual concentrations are used in a multiple logistic regression model, which controls confounding and risk factors, to test the association between pollutant measures and health effects.

A statistically significant association is found between the prevalence of serious chest illness before age two and sulfur dioxide measures. The association is sensitive to both the pollutant measure and the choice of geographic unit. The use of four or five year means of 50th and a higher percentile, e.g., 84th or 90th, concentrations, are found to provide the most consistent and representative pollutant characterization.

Acknowledgments

I would like to thank the following people for their help with this paper: my advisor, Elisabeth Paté, for her guidance and patience; Jim Gruhl, for his useful suggestions; Marc Schenker, for his critical and insightful comments; and Dan Goulomb, for his time and recommendations. I would also like to thank Gloria Mason for her dilligent editing.

The research was preformed at the M.I.T. Energy Laboratory and was supported by project EP-78-S-02-4968, sponsored by the U.S. Department of Energy. Dr. Joseph Blair is the project monitor.

Table of Contents

	page
Title	1
Abstract	2
Acknowledgments	3
Table of Contents	4
 1. Introduction	 9
1.1 Problem Focus	13
1.2 Scope	15
1.3 Approach	17
 2. Regulatory Background	 20
2.1 Historical Perspective	21
2.2 Ambient Air Quality Standards	24
2.3 Risk Analysis and Cost-Benefit Evaluation	28
 3. Technical Background	 31
3.1 Health Effects	33
3.1.1 Acute Air Pollution Episodes	33
3.1.2 Acute Respiratory Diseases and Effects	34
3.1.3 Chronic Obstructional Respiratory Diseases	37
3.2 Causal Factors	39
3.2.1 Problems in Dose-Response Estimates	43
3.2.2 Construction of Dose-Response Relationships	46
3.3 Air Pollution Epidemiology	52
3.3.1 Methodological Challenges -- Confounding	55
3.3.2 Indoor Air Quality	59
3.3.3 Study Types	60
3.4 Stochastic Health Models	61
3.4.1 Regression Models	66
3.4.2 Cigarette Models	68
3.4.3 Infectious Epidemiological Models	69
3.4.4 Cancer Models	70
3.5 Dispersion Models	71
3.6 Exposure Estimates	74
3.6.1 Selection of Pollutant Species	76
3.6.2 Monitoring Site and Geographic Unit	83
3.6.3 Representation of Exposure	85
3.6.4 Pollutant Interaction	91
3.7 Uncertainty	95

Table of Contents (continued)

	page
4. Chestnut Ridge Investigation	98
4.1 Health Study	100
4.1.1 Confounding Factors	101
4.1.2 Statistical Analysis	102
4.1.3 Screening Models	103
4.2 Pollutant Representation	106
4.3 Pollutant Indices	112
4.3.1 NAAQS	112
4.3.2 90th Percentile Measures	114
4.3.3 Arrowhead Indices	114
4.3.4 Exceedence Measures	115
4.3.5 Non-Linear Dosage Measures	120
4.3.6 Threshold Dosage Measures	122
4.4 Regional Characterization	126
4.5 Evaluation of Pollutant Indices	129
5. Case Study Analysis	132
5.1 Analysis of Regional Pollution Patterns	132
5.1.2 Characterization of Geographic Units	141
5.2 Analysis of Indice Stability	144
5.2.1 Relative Stability	153
5.3 Consistency	160
5.3.1 Arrowhead Indices and NAAQS	160
5.3.2 Exceedence Measures	161
5.3.3 90th and 99th Percentiles	165
5.4 Inter-Index Correlation	167
5.4.1 Arrowhead Indices and NAAQS	167
5.4.2 Exceedence Measures	170
5.4.3 90th and 99th Percentiles	179
5.4.4 Dosage Measures	184
5.4.5 Running and Calendar Year Averaging	184
5.5 Health Effects	184
5.5.1 Selection of Screening Model	185
5.5.2 Correlations to Mean Pollution Scores	186
5.5.3 Chronological Effects	194
5.5.4 Logit Model Results	198
5.5.5 Uncertainty in Pollutant Characterization	201
5.5.6 Risk Assessment of Proposed Power Plant	204
6. Conclusion	207
7. Appendices	
Appendix A. A Suggested Arrowhead Approach	213
Appendix B. Derivation of Non-Linear Dosage Estimates	215
Appendix C. Derivation of Threshold Dosages	219
Appendix D. Weighting for Statistical Variation	223
8. Bibliography	227

List of Tables

	page
Table 2.1 The National Ambient Air Quality Standards	26
Table 3.1 Definition of the Pollutant Standards Index	94
Table 3.2 Cross-Sectional Estimates of Mortality	94
Table 4.1 Power Plant Capacity and Stack Height	105
Table 4.2 Number of Children and Symptom Prevalence Rates	105
Table 4.3 Averaging Times on Arrowhead Curves	110
Table 4.4 Quantile Information on Arrowhead Curves	110
Table 4.5 Definition of Arrowhead Indices	113
Table 5.1 Annual Sulfur Concentrations at Monitors	136
Table 5.2 Annual TSP Concentrations at Monitors	136
Table 5.3 Correlations Between Annual Mean and Short-Term Monitor Concentrations	142
Table 5.4 Range of Districts, Townships and Schools as Compared to Monitors	142
Table 5.5 Correlations for Weightings of Sulfur Concentrations. . .	148
Table 5.6 Correlations for Weightings of TSP Concentrations . . .	149
Table 5.7 Ranked 9 Cleanest and Dirtiest Districts	151
Table 5.8 Ranked Districts for Minimum Population Size of 1000 .	152
Table 5.9 Normalized District Scores for Annual Sulfur Values . .	154
Table 5.10 Mean Normalized School Scores for 11 Sulfur Indices .	155
Table 5.11 Mean Normalized School Scores for 11 TSP Indices . .	155
Table 5.12 Combined Normalized School Scores for TSP and Sulfur .	155
Table 5.13 Ranked District Scores for Annual Sulfur Concentrations	159
Table 5.14 Annual COVs for 11 Sulfur Arrowhead Indices.	162
Table 5.15 Annual COVs for 11 TSP Arrowhead Indices	162
Table 5.16 Non-Parametric Stability for Sulfur and TSP Indices .	162
Table 5.17 Annual COVs for Sulfur Exceedences above 30 ppm. . . .	163
Table 5.18 Annual COVs for Sulfur Exceedences above 80 ppm. . . .	163
Table 5.19 Annual COVs for TSP Exceedences above 75 u.gms	164
Table 5.20 Annual COVs for TSP Exceedences above 150 u.gms. . . .	164
Table 5.21 Annual COVs for 90th and 99th Sulfur Percentiles . . .	166
Table 5.22 Annual COVs for Logarithms of 90th and 99th Sulfur . .	166
Table 5.23 Correlation Matrix for Sulfur and TSP Arrowhead Indices and NAAQS	168
Table 5.24 Correlation Matrix for Sulfur and TSP Standards. . . .	169
Table 5.25 Clustering Partitions for Sulfur and TSP NAAQS	172
Table 5.26 Correlation Matrix for Exceedence Measures and NAAQS .	174
Table 5.27 Correlation Matrix for the 90th and 99th Percentile .	181
Table 5.28 Range and Mean of 90th and 99th Sulfur Percentiles . .	181
Table 5.29 ANOVA Results for 90th and 99th Sulfur Percentiles . .	182
Table 5.30 Correlations for Dosage Measures	183
Table 5.31 Sulfur-TSP Correlations for Calender and Running Year.	183

List of Tables (continued)

	page
Table 5.32 Correlations between Arrowhead Indices and Health Effects	189
Table 5.33 Correlations between Exceedence Measures and Health Effects	190
Table 5.34 Correlations between Dosage Measures and Health Effects	190
Table 5.35 Correlations with Normalized Sulfur and TSP Indices and Health Effects	190
Table 5.36 Correlations between Combined TSP and Sulfur Indices	193
Table 5.37 Chronological Effects of NAAQS and Health Effects . .	193
Table 5.38 Chronological Effects of Exceedence Measures	195
Table 5.39 Correlations between Dosage Measures and Health Effects	196
Table 5.40 Running Vs. Calendar Averaging and Health Effects . .	197
Table 5.41 Results of Logit Analysis	200
Table 5.42 Relative Risks for Serious Chest Illness	200
Table 5.43 Correlations with Weighted Screening Model	203
Table 5.44 Predicted Pollution from Proposed Power Plant	203

List of Figures

	page
Figure 3.1 Hypothetical Linear Toxicity	42
Figure 3.2 Hypothetical Non-Linear Toxicity	42
Figure 3.3 Hypothetical Static Dose-Response Function	49
Figure 3.4 Hypothetical Population Dose-Response Curves	49
Figure 3.5 Dose-Response Effects over an Air Pollution Episode .	51
Figure 3.6 Dose-Response Effects over Two Air Pollution Episodes	51
Figure 3.7 Epidemic Curve	65
Figure 3.8 Distributions on Population Dose-Response Curve . . .	65
Figure 4.1 Chestnut Ridge Study Region.	99
Figure 4.2 1978 TSP Arrowhead Profile for Monitor 1	108
Figure 4.3 Running versus Calendar Year Averaging	116
Figure 4.4 TSP Arrowhead Profiles for Four Consecutive Years . .	117
Figure 4.5 Dose-Response Functions for Dosage Measures.	123
Figure 4.6 Location of Districts, Monitors and Schools	127
Figure 4.7 Block Diagram showing Derivation of Pollutant Score .	128
Figure 5.1 Block Diagram of Index Analysis Process	132
Figure 5.2 1975 Annual and 3 Hour Sulfur Concentrations at Monitor Locations	137
Figure 5.3 1976 Annual and 3 Hour Sulfur Concentrations at Monitor Locations	137
Figure 5.4 1977 Annual and 3 Hour Sulfur Concentrations at Monitor Locations	138
Figure 5.5 1978 Annual and 3 Hour Sulfur Concentrations at Monitor Locations	138
Figure 5.6 1975 Annual and 24 Hour TSP Concentrations at Monitor Locations	139
Figure 5.7 1976 Annual and 24 Hour TSP Concentrations at Monitor Locations	139
Figure 5.8 1977 Annual and 24 Hour TSP Concentrations at Monitor Locations	140
Figure 5.9 1978 Annual and 24 Hour TSP Concentrations at Monitor Locations	140
Figure 5.10 Normal Probability Plots of Annual Sulfur Concentrations	145
Figure 5.11 Normal Probability Plot of 24 Hour Peak Sulfur Concentrations	145
Figure 5.12 Distribution of Ranked TSP District Scores	147
Figure 5.13 Cluster Tree of Annual Sulfur Concentrations	158
Figure 5.14 Annual TSP versus Sulfur Concentrations	171
Figure 5.15 24 Hour Peak TSP versus Sulfur Concentrations	171
Figure 5.16 Relationship of Sulfur Exceedence Measures	175
Figure 5.17 Relationship of TSP Exceedence Measures	175
Figure 5.18 Sulfur Exceedences (80ppm) versus Annual Sulfur . . .	176
Figure 5.19 Sulfur Exceedences (150ppm) versus Annual Sulfur . .	176
Figure 5.20 TSP Exceedences (75 u.gms) versus Annual Sulfur . . .	177
Figure 5.21 TSP Exceedences (150 u.gms) versus Annual Sulfur . .	177
Figure 5.22 Annual Sulfur Concentrations versus Chest Illness . .	188
Figure 5.23 TSP Exceedences versus Prevalence of Chest Illness .	188

Chapter 1. Introduction

The goal of this study is to assess the role and validity of measures of ambient air pollution used in health models which associate health effects with air pollution. At present, health models are used in the standard setting policy process for ambient air quality. There is also considerable interest in using health models in a risk-cost-benefit approach to standard setting, power plant siting, and assessment of pollution abatement technology. Potential improvements to these models include the use of better pollution measures and the introduction of uncertainty.

Like other public safety and health regulations, ambient air quality standards and policies evolve in the policy process through consideration of scientific, technological, economic and political knowledge and values. Informed scientific policy making must take into account both scientific advances and technological limitations to achieve a regulatory agenda which is desirable and feasible. However, policy making takes place in the democratic, hence political, context and thus also must respond appropriately to shifts in political values and economic priorities. A prime example is the recent emphasis that regulations be cost-effective. Out of this scientific, political and bureaucratic milieu emerge regulations that may not meet the legislative intent in the most effective manner, and, as shown by the evolution of the ambient air quality standards, may be extremely controversial.

The chief motivation for the establishment of the National Ambient Air Quality Standards (NAAQS) in the Clean Air Act as Amended is the

protection of the public's health and welfare. The Primary Standards were developed as a preventative strategy to control exposures to pollutants which have been identified as underlying causes of human disease. These Standards, in the judgement of the Administrator of the Environmental Protection Agency (EPA) are to protect the most susceptible portion of the population, though not necessarily the most sensitive individual within that subgroup, from adverse health effects with an adequate margin of safety. To date, this has been accomplished by the designation of maximum tolerable ambient concentration levels for seven pollutants.

Epidemiological evidence concerning the relationship of air pollution exposures to morbidity and mortality does not validate the assumptions of the simple, i.e., non-synergistic, non-cumulative and threshold triggered dose-response model embedded in air quality regulations. The most extensive (\$22 million) air pollution epidemiological study, the Community Health and Environmental Surveillance System, conducted by the EPA between 1967 and 1975, pointed out the many difficult problems in carrying out air pollution health research (Committee on Science and Technology, 1976; Cooper, 1979). Largely due to the dearth of quality data, the state of the art in the assessment of air pollution and pollutant-induced health effects is, in general, quite crude. Because of the difficulty and expense of collecting and analyzing health data, most epidemiological studies have relied upon standard mortality and health care facility utilization files. This data is of poor quality, often unvalidated, biased and contains numerous socio-economic, personal, and environmental factors which tend to confound results. Additionally, air quality data used in such analyses commonly suffer defects of inaccuracy,

bias, and irrelevancy. These data problems often obscure the relationship of air pollution to health. Moreover, there is no standardized procedure for the selection and treatment of pollution variables, health measures and confounding factors in health effects research concerning low level ambient pollutants. Published estimates of health impacts for various air pollutants vary widely. Consequently, there is little truly accurate information to provide guidance to policy makers who strive to make a reasonable assessment of health risks for both the selection of hazardous air pollutant species to be regulated and the determination of the threshold level, that is, the level considered to be hazardous.

Additionally, there exists considerable debate on the appropriateness of the present regulatory mechanisms for the control of air pollution. The enforcement of the Ambient Air Quality Standards demands a level of precision beyond the capabilities of current technology. Air quality dispersion models are recognized by some to be the most equitable means available to support the Nation's air quality goals (American Meteorological Society, 1980). Such models play a determinant role in the present environmental assessment process. However, critical input data, particularly meteorological and topographical information, are often lacking or inadequately accounted for. Thus, these models provide rather inaccurate forecasts of ambient pollutant levels, and quite poor predictions of the peak measures required in the NAAQS. Frequently used models are seriously constrained, both temporally and spatially, and thus can only model local (50-80km) impacts. The immigration of air pollutants from outside the modelled region contributes to an often important background level. Exports of pollutants to other regions, perhaps already

significantly affected, are similarly unaccounted for. Uncertainty in characterizing point source emissions, due to operations and technology of the pollution source, further exacerbates the difficulties of ascertaining ambient levels of air pollution. As a response to some of these problems, the EPA recently established the New Source Performance Standards for steam electric generating facilities. These Standards limit pollution outfall per unit of heat input. Other regulatory mechanisms, yet untried in this country, are also available. For example, power plant siting banks and capacity planning, or real time pollution abatement strategies, potentially might provide, in conjunction with certain existing provisions of the Clean Air Act, more effective and realistic control of ambient air pollution.

Finally, it must be recognized that although the intent of the Clean Air Act is to prevent adverse health impacts, in practice, the existing regulatory framework is not health based. Assessments of demographically and exposure adjusted doses, and calculations of estimated health impacts are not conducted at the present time. The Prevention of Significant Deterioration (PSD) regulations could be regarded, in part, as a step in this direction, recognizing that regional differences demand regulatory acknowledgement. Because the margins of safety originally set in the CAA have turned out to be very modest or nonexistent, according to the Committee on Interstate and Foreign Commerce (1979), they may have resulted in health benefits to inhabitants of the region. However, parity of risk among PSD regions would be coincidental. Rather, if a cost-benefit policy is adapted, risk-cost-benefit assessments should be used for the evaluation of pollutant standard efficacy. This permits a

balancing of costs of regulatory compliance with health and other benefits resulting from lower pollution levels in and among particular regions.

As stated at the outset, the determination of standards takes place in a politicized environment. The resolution of this complex, value-laden issue must consider views from industry, as represented by the Electric Power Research Institute's estimate that the New Source Performance Standards will cost utilities \$200 billion by the year 2000 (Cooper, 1979). Views of "health advocates," represented here by Meyer's and Cederwell's estimate that by the year 2000 over 26,000 deaths will occur annually from the projected increased coal utilization with 80% sulfur oxide control (Viren, 1978), also demand legislative response. Implications for the nation's energy future, economy and health status resulting from ambient air quality policies are large. In the policy debate, impact assessments can influence substantive policy outcomes, but only to the extent that bureaucracies and political processes respond to the dissemination of such information. (Bacow, 1981) In turn, these assessments are based on exceedingly difficult and intricate health models, which strive to represent a reasoned judgement on a border area of scientific knowledge and policy choice.

1.1 Research Focus

To be useful in the scientific policy process, health models must be consistent, internally valid, and make the best use of all available information, including probabilistic air quality and health information

that may suffer from both scientific and statistical uncertainties. Ideally health models would be robust and powerful. For predictive purposes, given imperfect information, they will have probabilistic output. Furthermore, their design should be amenable to the policy context, accommodating various functions, yet also displaying limitations on their uses and predictive validity. With this tool, the policy maker can assess the desirability of existing air regulations and suggest or support modifications leading to better standards.

A critical element in both air quality assessments and health effects research is the representation of ambient pollution. The best dosage measure for health effects from common ambient air pollutants is unknown. At present, maximum levels at various averaging times are set for the several regulated pollutants. The same measures are generally used in air pollution epidemiological studies. This choice of metric neither fits physiological hypotheses or models for health impacts from air pollutants nor is the result of empirical data extrapolation from the epidemiological studies. The focus of this research is to develop and test, using the case study described below, pollutant parameter representations which best bridge physiological hypotheses with observed impacts of several specific health effects in the study population. Such ideal formats for human dosage characterization may help optimize predictive dosage modeling and pollution related health assessments.

The development of such clinically meaningful pollutant measures may reduce the dimensionality of characterizing ambient air pollution exposures to a few, or possibly a single, index which can best represent health impacts. At present, ambient air pollution measures potentially

1

might include any number of averaging times, cumulative frequency distributions, patterns of occurrence, multiple pollutants, and pollutant interactions, as well as being cohort specific, i.e., adjusting for attributes of the population at risk, for example, whether or not individuals smoke cigarettes, or disease specific, i.e., treating diseases with a long induction period differently from those which have more acute origins. The reduction of dimensionality down to one or several indices necessarily will lose some information, but hopefully gain relevance on physiological terms, which is of paramount interest in health assessments.

1.2 Scope

Ambient air pollution policies and health effects research strive for the purpose of protecting the public's health. Today, the protection of public health means the attainment of desirable positive health values (a state of complete physical, mental, and social well being, according to the World Health Organization), in addition to the prevention of undesirable disease or infirmity (Purdon, 1980). Public welfare is defined to include effects on soil, water, crops, vegetation, animals, wildlife, weather, visibility, climate, manmade materials, economic values, personal comfort and well being (Gage, 1979). To help determine policy posture, health and environmental consequences as well as costs and impacts of regulatory compliance must be investigated.

The scope of this research is an analysis of commonly used pollution measures and some alternative measures which may be used in health models

and risk-cost-benefit assessments. The stability of these measures and their usefulness as health indicators is assessed by the use of a case study, described below.

The present research is limited in many aspects. First, only two of the seven original "criteria" air pollutants -- sulfur dioxide and total suspended particulates -- are examined in this study. Carbon monoxide, hydrocarbons, oxides of nitrogen, and ozone are omitted. Pollutants regulated under the National Emission Standards for Hazardous Air Pollutants (Section 112 of the CAA), which include benzene, cadmium, lead and vinyl chloride also are omitted 1/. In addition, other pollutant residues may be particularly important in both ambient and occupational exposures. The emphasis of this study, however, is on the sulfur oxide and particulate complex commonly found in "ordinary" ambient air pollution. Secondly, the pollution mix or "cocktail" found in the study region represents an area heavily impacted by coal use: four large mine fed power plants are the major sources in the study area. Thus, this study may be relevant to impact assessments of this common fuel.

The third major limitation of this study is that the examination includes only respiratory and pulmonary system health effects. Other types of health effects are linked to elevated air pollution exposures. Studies have shown associations of pollutants levels with increased mortality and shortened life expectancy (Lave and Seskin, 1978), birth

1/ Radionuclides as well as 43 other pollutants are proposed additions to these standards pending review of the CAA in late 1981 (Hamilton, 1981; National Commission on Air Quality, 1981).

weight (Williams, 1977), sensory irritation, central nervous system and human performance effects, cardiovascular disease, cancer of the lung (American Thoracic Society, 1978), cancer initiation in other sites (from benzene, U.S. Supreme Court, 1980), and decreased survival rates after (stressful) cancer operations (Wilkinson, 1979). Additionally, studies have related pollution to indirect indicators of health, using hospital admissions and other measures of utilization of health care services.

This research does not include an assessment of the technical feasibility or pollution control costs potentially associated with hypothetical regulations founded on such clinically meaningful measures. Alternative pollutant data collection schemes which potentially could result in far greater exposure information also are not assessed.

Increased coal utilization and the resulting air pollution raise several other serious questions which concern indirect health effects. These include long range climatic change from the "greenhouse effect" due to increased carbon dioxide levels. They also include effects of acid precipitation, attributable to sulfur and nitrogen oxides which adversely affect aquatic plants and animals, and increase the leaching of toxic metals, possibly into potable water supplies. These questions are unresolved, and beyond the scope of this report.

1.3 Approach

The regulatory experience with ambient pollution standards is described in Chapter 2. This information is important since existing regulatory requirements largely have determined available pollution

measures used in recent epidemiological research and risk-cost-benefit assessments. It is also desirable to avoid pollution measures which might present difficulties in terms of technical feasibility or implementability. Then, a survey of the physiological and epidemiological evidence and hypotheses used in the identification and quantification of dose response dependencies is presented in Chapter 3. This information, with the brief background on air quality forecasting presented, illuminates contemporary practice, gaps in knowledge, and potentially profitable directions in which to proceed with respect to clinically relevant pollution measures.

The Chestnut Ridge air pollution and health study in western Pennsylvania (funded by the U.S. Department of Energy, contracted to the M.I.T. Energy Laboratory, with a subcontract to the Harvard Medical School, and with the participation of area electric utilities) is used as a well controlled case study. It is described in Chapter 4. This project provides opportunities to test various pollution measures and regional characterizations (for the cross-sectional health studies) which correlate to respiratory and pulmonary system health effects found in the inhabitants of this area. Considerable work has gone into health effects research in the Chestnut Ridge area. Health surveys using several distinct epidemiological approaches have collected resident morbidity and mortality data. Confounding and risk factors have been isolated and corrected by using multiple logistic regression techniques. Selected pollution measures, and weightings of the certainty and temporal effects of the characterizations are used to explore the variance which is not explained by risk and confounding variables in the health surveys.

A second major task, which precedes the health effects research, is the development of pollutant measures which are stable over many years of apparently conflicting air quality measures. This task is facilitated by the existence of an extensive network of air quality monitors and well developed analytic dispersion models in this region. Pollutants monitored in this network include total gaseous sulfur, sulfate, total suspended particulates, total oxides of nitrogen, nitric oxide, total oxidants, ozone and coefficient of haze.

Lastly, an assessment of the health risks posed by a proposed coal power plant in the Chestnut Ridge area is made using the regression health model and an air pollution dispersion model.

Many qualifications limit the applicability of the evidence and techniques in the assessment. A major task is simply the development of data validation and interpretation techniques for air pollution measures. These will permit a region to be accurately and consistently characterized as clean, dirty, or any intermediate level of relative pollution impact -- before the health data is examined. The separation of issues of stability from pollutant measures of clinical relevancy may severely restrict the range of measures possible. Moreover, a significant caveat concerns the extent to which this research, which tests hypotheses on a restricted data set, can be generalized. Despite the several study populations and approaches used in the epidemiological study, other assessments of these measures and techniques will be necessary to validate the results.

Chapter 2. Regulatory Background

Arguments concerning public and private decisions relating to air pollution standards and policies surface in many forms, manifesting various ideologies and traditions. Some degree of pollution, which may or may not be detrimental to human health, environment and property, may be economically unavoidable to society as we know it. Perhaps the underlying rationale for government activity in the area of pollution control and management may be found in theories of distributive justice. It is clear that the impact of air pollution is not equally distributed among individuals: only some will suffer adverse health effects and property damage, which, for all practical purposes, is involuntary.

These external costs, that is, the difference between social and private costs, arise when the polluter does not bear the full cost of his or her actions, thus saving money. Polluters may fail to voluntarily clean up their emissions because their individual actions may not manifest any discernible improvement in air quality. Without adequate legislation designed to protect the rights of individuals to a healthy environment, air quality would tend to get worse. Most governments have recognized these consequences and have adopted a protective attitude towards public policy decision making. Enabling legislation, based on the police power of the state to protect the public's health, property and welfare, has established the regulatory agencies that are authorized to enforce pollution abatement laws. Today, they are the primary actors in environmental litigation. Private actions, based on the common law theories of tort liability including nuisance, negligence, and strict

liability, also occur but usually are of secondary importance. Individuals may have difficulty in obtaining standing to sue, adequate scientific and economic expertise, and, in the event of a favorable decision, may find that the remedies are not altogether satisfactory. However, the interpretation of the law by the courts is open and somewhat flexible. Although no constitutional guarantee setting a clear goal like "no harm is done to anyone" is likely to exist in the near future, the law does seem to be evolving towards a doctrine of public stewardship of the public trust, including clean air.

2.1 Historical Perspective

The history of air pollution regulation goes back many centuries. In fourteenth century England, a royal proclamation by King Edward I prohibited the use of "sea coal" (bituminous coal) in furnaces due to an overabundance of smoke. Violators were not treated lightly: in 1307 one offender was executed [Murdoch, 1975]. About two centuries later, in 1611, an English court decided Alfred's Case, which concerned objectionable odors emanating from a pig farm. This case is generally regarded as the first instance of air pollution litigation. Still, in those days pollution control was technologically difficult and a 1661 report stated that one-half the deaths in London were "phthisical and pulmonic distempers" resulting from polluted air. Later, the control of air pollution was considered a health agency responsibility in the Public Health Act of 1848 (Purdom, 1980).

The earliest governmental activities in the U.S. were local smoke

ordinances that declared black smoke to be a public nuisance. Chicago and Cincinnati first enacted such smoke ordinances in 1881, and, by 1912, most cities had smoke control laws. Many of these laws stipulated that emitting smoke darker than a particular density, as indicated by the Ringelmann scale, constituted a punishable offense (Lynn, 1976). Generally these laws were enforced by some unit of the building inspection or boiler safety agency. By 1930 it was apparent that smoke ordinances were sometimes a very inadequate pollution control measure, because small consumers had no really effective methods to control smoke emissions from their small and inefficient appliances. As a result, several cities such as St. Louis and Pittsburgh prohibited the sale of bituminous coal with a volatile content of more than 20% and provided alternative smokeless coals. At the same time, a widespread conversion to oil or natural gas dramatically improved the pollution situation in many parts of the U.S. (Halliday, in Heimann, 1961).

The air pollution situation in Los Angeles was getting worse, however, due to the well known "L.A. smog." Thus, in 1945, the City Air Pollution Control Office was established. A year later, with the realization that air pollution is a regional problem rather than a city problem, it was reorganized as a county office. In 1947 the California legislature enacted the first state law dealing with air pollution, beginning a general trend towards state involvement in what had been primarily a local concern.

Federal air pollution programs began in 1955 with research and technical assistance to states. The original Clean Air Act (CAA), established in 1963, effectively required each state to have a state

level program. This Act was amended in 1967, 1970, and 1977. The 1970 revision established the National Ambient Air Quality Standards. Because the federal government has no formal police powers, the primary responsibility for the prevention and control of air pollution still rests with state and local governments. However, the federal Environmental Protection Agency (EPA) currently is the most prominent actor in policy formulation and implementation.

In overview, the EPA conducts research, sets ambient and certain emission standards, adopts regulations for achieving and maintaining standards, disburses funds, approves and oversees the carrying out of state implementation plans, and enforces the CAA requirements. States develop and enforce strategies for achieving and maintaining ambient standards and incorporate those measures into their state implementation plans. Local governments still maintain the options to develop and enforce control strategies for stationary and mobile sources within their jurisdiction (National Commission on Air Quality, 1981).

The complexity, cost and time delays inherent in the enforcement of the CAA are criticized frequently. For example, there are seven permit programs; it takes an average of three years to obtain a permit; and the average application costs \$250,000 to \$300,000 (Orloff, 1981).

There are five general classes of potential solutions suggested to reduce the "drag" on the economy imposed by environmental legislation such as the Clean Air Act. These include the 1) de-enactment of portions of the CAA legislation, 2) streamlining and facilitation of the permit process (i.e., the ill-fated Energy Mobilization Board), 3) establishment of self-executing requirements by private compliance with prescriptive

regulations on the fear of a federal audit and stiff fines, 4) economic incentive approaches, using pollution taxes, credits, etc., which are bought and sold among polluters and/or a public trustee, and 5) risk-cost-benefit determination of standards, on a national and regional level, using any of the several possible mechanisms for enforcement. Several economic incentive approaches recently have been adopted for use in the U.S. These include 1) emission offsets, 2) the bubble policy, and 3) emission banking. Being studied are 4) transferable emission reduction assessments, 5) emission fees, and 6) subsidies [National Commission on Air Quality, 1981]. Lastly, site banking, which reserves sites with good dispersion potential for major emitters, may be considered an economic approach with the appropriate tax incentives, etc. It is important to realize that except for the risk-cost-benefit evaluations, these are mechanisms for the implementation and compliance of regulations and are not a social or policy evaluation of risk to health and property.

2.2 National Ambient Air Quality Standards

The Clean Air Act (CAA) is one of the earliest major environmental laws enacted by Congress. It is based on the principle that government must act to control potentially harmful pollutants despite scientific uncertainty about the precise nature and extent of the harm. The statutory basis for setting the National Ambient Air Quality Standards (NAAQS) does not take economic factors into account. Congress recognized that while the costs of meeting a specific standard would be substantially higher in an area with severe pollution, levels of

pollution at which public health is affected generally do not vary among different locations. Therefore, the rationale for uniform national standards like the NAAQS is the equal protection of public health and welfare [National Commission on Air Quality, 1981]. Congress implied that the economic effects of achieving the Standards, would be acceptable. It is generally acknowledged that the NAAQS do reduce serious and fatal health effects caused by air pollution "episodes," that is, very high levels of pollution which last several days. However, industry and others contend that the health benefits of achieving specific levels of certain pollutants have been oversold [APCA, 1974]. Some also contend that the costs of compliance with the NAAQS do not bear a reasonable relationship to the health benefits of meeting the standards [Cooper, 1979].

Much of the evidence which was used to set the NAAQS is based on epidemiologic studies. Early studies tended to indentify a threshold level of concentration for several pollutants, below which no adverse health effects were noted 1/. To incorporate a margin of safety necessary to protect sensitive individuals in the population, these threshold levels were lowered and embodied into law as the NAAQS. They are shown in Table 2.1. For several pollutants, e.g., particulates, the U.S. makes a distinction between health and welfare standards, contrary to practice in several other countries. Secondary Standards were established to provide for welfare in situations of extremely high pollution but of short duration, which were not thought to affect health.

1/ The definition of "no adverse effects" is itself difficult to define, and varies with medical judgements, the particular individual, etc.

Pollutant	Maximum Concentration at Averaging Time		
	3 hour	24 hour	annual
Sulfur Oxides			
primary		0.014 ppm (80 ug/cu.m)	0.03 ppm (365 ug/cu.m)
secondary	0.5 ppm (1300 ug/cu.m)		
Suspended Particulate Matter (TSP)+			
primary		75 ug/cu.m.	260 ug/cu.m
secondary		60 ug/cu.m.	160 ug/cu.m.
Carbon Monoxide	35 ppm *	9 ppm	
averaging time	(1 hour)	(8 hour)	
Ozone	0.12 ppm (1 hour)		
Reactive Hydrocarbons (non-methane)	0.24 ppm		
Nitrogen Dioxide			0.05 ppm
Lead		1.5 u.gm/cu.m	
averaging time		(3 months)	

+ geometric average

* proposed reduction to 25 ppm.

Table 2.1. Definition of the Primary and Secondary National Ambient Air Quality Standards. Concentrations for averaging time of 24 hours or less are not to be exceeded more than once a year.

The ambient standards have many shortcomings. For example, synergisms between pollutants which produce a more dangerous substance than either alone, e.g., the sulfur and particulate complex, are ignored. Also, total suspended particulate (TSP) measures do not necessarily indicate levels of more harmful respirable particulates. These issues are discussed in the Technical Background, Chapter 3.

Since 1970, a number of epidemiologic studies have indicated that the margins of safety in the NAAQS are very modest -- much smaller than those maintained with other environmental hazards, such as radiation and pesticides. For example, for nitrogen dioxide and ozone, the NAAQS provide less than a twofold margin of safety for the healthiest and best conditioned segment of the population. Vulnerable groups, such as infants, pregnant women, persons with heart or lung disease or under stressful conditions, including heat and occupational exposures, have shown adverse effects at levels 25 to 40% more stringent than the existing standards. Levels of sulfur dioxide and TSP slightly lower than the NAAQS have appeared to aggravate asthma, increase symptoms of respiratory disease in children and the elderly, as well as lower the lung function of children [Committee on Interstate and Foreign Commerce, 1978].

The National Academy of Science has stated that there is no evidence that threshold levels have a clear physiological meaning, in the sense that there are genuine health effects above some level of pollution, but no effects at all below that level [Committee on Interstate and Foreign Commerce, 1978]. The severity and frequency of health effects may accumulate over time with low level exposures, or be a continuous phenomenon, increasing with even small pollution exposures. These

hypotheses are developed in Chapters 3 and 5. The threshold concept must be recognized as meaning a level below which no observable effects occur [Lynn, 1976]. There are alternatives to threshold standards, which rather imperfectly attempt to protect human health on an absolute basis. Air quality policies should be based, perhaps, on allowing a tolerable level of adverse health effects. The evolving regulatory strategies which deal with the health risk posed by air pollution and the economic, social, energy and other costs, are the subject of the next section.

2.3 Risk Assessment and Cost-Benefit Evaluation

Congress implicitly determined that the economic effects of several air quality programs would be acceptable. These include the establishment of the NAAQS, achievements of these standards by established deadlines, emission standards for hazardous pollutants, and several aspects of the federal motor vehicle control program. On the other hand, Congress explicitly ordered the E.P.A. to consider the economic effects of the New Source Performance Standards, portions of State Implementation Plans, determination of best available control technology, and the remaining aspects of the motor vehicle program [National Commission on Air Quality, 1981]. There is strong impetus at present, however, to consider and balance the costs and benefits of all regulatory policies. This must be determined by legislative mandate: it is not normally at the discretion of the executive (agency level) branch of government.

The economically optimal level of pollution is reached when the sum of societal costs due to pollution and costs of pollution control is

minimized. Costs and benefits which result specifically from pollution control are extremely difficult to define and measure accurately, and will not be examined here. Instead, a brief description of the process of risk-cost-benefit analysis is given. Risk assessment first involves the recognition of an environmental hazard, i.e., air pollution. Then, an estimate of the risk in terms of the likelihood and magnitude of the hazard is developed, and usually quantified. Lastly, to evaluate the environmental condition, a judgement is made of the meaning attributable to the risk estimate [Kates, 1979]. Risk-cost-benefit evaluation uses economic criteria, i.e., cost-effectiveness, to judge the seriousness of the risk and/or the appropriateness of the proposed policy.

Applied to setting standards, a risk-cost-benefit approach should elucidate all the necessary facts and uncertainties of the costs and benefits of pollution control. It should take all relevant information into account in a consistent manner [Paté, 1980]. The formal analysis makes the balancing of risks, costs, and benefits, now done internally in Congress or the E.P.A., explicit. Moreover, it may result in better regulatory decisions based on health and economic impacts, rather than incremental increases in emissions or ambient standards [Gruhl, 1976].

The task is greatly complicated by political and scientific disagreement and uncertainty. Areas of particularly large uncertainty or dispute include 1) the estimation of the appropriate dosage measure of pollution exposure, 2) the modeling of the relationship between ambient air pollution and health effects, and 3) the evaluation of somewhat disparate effects or measures, i.e., quality of life indicators, or dollar equivalents of various health effects. These uncertainties must be

incorporated, usually probabilistically, in risk-cost-benefit assessments of air pollution. The benefits of this approach depend on whether 1) the results can discriminate among policy alternatives, 2) the public accepts the optimality of policy decisions based on such methods, 3) other policy approaches are not more acceptable, e.g., no adverse effects, and 4) assessments can be carried out in a reasonably efficient manner [Evans, 1981]. It should be noted that the federal government is presently conducting a annual total of only \$90 million dollars worth of research in the air pollution area, a small amount compared to air pollution control expenditures of some \$1.6 billion.

This research is primarily concerned with measures of air pollution that are able to characterize the actual health impacts felt by a population. Such idealized dosage measures may be derived from actual monitor concentrations and/or pollutant dispersion models. The latter approach has been used to evaluate health risks of proposed energy facilities [Morgan, 1978; Henrion, 1981]. There is considerable uncertainty in dosage estimates due to the 1) air pollution modeling, which estimates representative concentrations for sectors of the population in different geographic regions, 2) dose-response uncertainties in the construction of the measure from the concentration estimates, and 3) temporal effects, including year to year variation and cumulative effects of exposure. With respect to risk assessment and evaluation, the major contribution of this research is an analysis of the effect and magnitude of these uncertainties, and an estimate of the minimum amount of information necessary in order to provide valid predictions of risk.

Chapter 3. Technical Background

Research in air pollution related health effects is aimed at understanding disease causation by an evaluation of the relationships between ambient air pollution exposures and disease rates. Standardization for confounding or intervening factors due to peculiarities of the test subjects, exposure characterization, or study type is usually required to give meaningful estimates of risk to human populations.

There are two general and complimentary approaches with which to obtain evidence regarding the health effects of pollutants; namely, the sciences of toxicology and epidemiology. The toxicological approach, which is primarily experimental, involves the testing of individual pollutants or mass emissions, that is, representative exhaust stack mixtures, for the poisonous properties of compounds and their mechanisms of action. Animal tests frequently are used to observe physiological responses and development of pathologies. Human subjects also may be exposed to artificial atmospheres in clinical studies. These experimental studies are performed in a clinical environment where great control may be exercised in the selection of subjects and the exposure dose. Such studies may yield precise dose-response dependencies between pollutants and health effects of the test subjects, albeit in a rather artificial pattern of exposure 1/ and with limited representativeness of the

1/ Exposure levels may be relatively high in terms of annual concentrations. Clinical exposures periods are much shorter.

population and its behavior. Studies involving human subjects are generally limited to acute effects at relatively low exposure levels, for obvious ethical and practical reasons. By contrast, epidemiological investigations are principally observational studies on human subjects. They may take several forms, typically occupational (e.g., industry specific), cross-sectional (e.g., between cities), or longitudinal (e.g., within cities over a period of time). They are described at length in the following sections.

Adverse health effects have been associated with four major types of air pollution related to three major source categories. These include the sulfur oxide and particulate complex which arise from the combustion of fossil fuels containing sulfur. This reducing type of pollutant is the focus of this research. The combustion of fossil fuels also produces various organic carbon compounds, forming the second type of air pollution, which includes soot, particulates, and polycyclic aromatic hydrocarbons. Many of these polycyclic organics have been proven to be mutagenic in bacteria and carcinogenic in animals. There is strong direct evidence of their carcinogenicity in man from occupational studies (Hamilton, 1981). Photochemical oxidants formed in the atmosphere from a complex chemical reaction between hydrocarbon compounds, nitrogen oxides and carbon monoxide are the third type of air pollutants. These are related to motor vehicle emissions and fossil fuel combustion. The fourth type is a miscellaneous class of air pollutants, such as arsenic, asbestos, beryllium, cadmium, hydrogen sulfide, lead, mercury, vinyl chloride and polyvinyl chloride. These pollutants are mostly related to localized point sources, such as smelters, refineries, and manufacturing

plants (American Thoracic Society, 1978) and may result in occupational and environmental exposures. They are also present as trace constituents in fossil fuels, especially coal, and appear adsorbed on inert particulate emissions.

3.1 Health Effects

There are five general categories of adverse health effects which have been related to exposures from ambient air pollution. These include (1) air pollution associated mortality, (2) respiratory and pulmonary symptoms and disease, and ventilatory disfunctions, (3) cancer of the lung, (4) cardiovascular disease and (5) sensory, neurological and behavioral effects.

3.1.1 Acute Air Pollution Episodes

Acute air pollution episodes first prompted health care professionals to show genuine interest in the fact that the air we breath can cause disease, that is, be an etiologic factor itself. Just several decades ago, no firm proof existed that the inhalation of gases and aerosols was related to ill health except to the extent that air might act as a medium that permits and even aids the conveyance of certain airborne bacteria and viral infections.

The "classical smog" (popularly thought to be due to smoke and fog, or at least moisture) is associated with the well know air pollution disasters in the Meuse Valley of Belgium in 1930; in Donora, Pennsylvania

in 1948; in London in 1952 and again in 1962; and in New York City in 1953. These episodes are well documented (Heimann, 1961; Lynn, 1976). In Donora, for example, some 40% of the population reported some degree of illness. Symptoms produced, in decreasing order of occurrence, included cough, sore throat, constriction of the chest, headache, dyspnoea (shortness of breath), a burning sensation of the eyes, onothopnoea, lacrimation (excessive secretion of tears), vomiting, nausea, and excessive nasal discharge. Of the 17 deaths that occurred in Donora, 15 occurred on the third day of the episode. The single significant factor among the fatally ill appeared as a pre-existing disease of the cardio-respiratory system (Heimann, 1961).

These acute air pollution episodes most often occur in the winter months when much fuel is being consumed and in a period of air stagnation or inversion. Several days of the severe air pollution present in such episodes may be related to sometimes dramatic increases in mortality, that is, the excess mortality which is equal to the difference between the number of deaths recorded during the episode and the number expected for the same time period. These episodes are not usually helpful in identifying the pollutant species responsible for the excess mortality because the stagnant conditions diminish the rate of dispersion for all emission sources.

3.1.2 Acute Respiratory Diseases and Effects

Acute diseases and health effects mostly have been related to the reducing type of pollution. There is less evidence which associates the

oxidant type of air pollution with health effects, though there is no question of the ability of this type of air pollution to produce severe eye irritation (US Department of Health, Education and Welfare, 1967).

Laboratory animals frequently have been used to develop experimental dose response dependencies for ambient air pollutants. These studies include lung tissue injury to various energy related pollutants (Musafa, 1981), and impaired resistance to respiratory infection and disease (Thomas, 1981). Numerous other toxicological, clinical and occupational investigations have been conducted as well, but usually test responses to a single pollutant, rather than a more representative mix of ambient air pollution.

Lower tract respiratory diseases and effects: Several epidemiological studies have supported associations of acute lower tract respiratory diseases and effects fairly consistently. These diseases include acute bronchitis and pneumonia. Acute effects include bronchoconstriction, increased airways resistance, respiration and pulse rate. The resultant stress from these effects is a reasonable explanation for the increased morbidity and mortality found among individuals suffering from cardio-respiratory disease in acute air pollution episodes.

Upper tract respiratory diseases: Epidemiological studies have not supported associations of upper tract respiratory diseases as consistently as that of lower tract diseases. Such upper respiratory tract diseases include the common cold, pharyngitis and laryngitis.

Aggravation of existing disease: Air pollution has been associated with an aggravation of bronchial asthma, which is a heterogeneous syndrome in which various diseases entities are lumped together. The

common feature is an increased responsiveness of the trachea and bronchi to various stimuli, including various aeroallergens of natural origin. The resulting generalized narrowing of the airways results in impaired breathing, wheezing and irritating cough. This narrowing is dynamic and changes in degree either spontaneously or because of therapy (Goldstein, 1981). Attack rates have been associated with sulfate levels. The effect of sulfates is greater as ambient temperatures decrease.

Ventilatory Function: Studies of ventilatory function provide valuable information on human responses to individual pollutants and interactions between pollutants. Although slightly diminished lung function may not be important in itself, it may have long term consequences, leading to obstructive airway diseases (American Thoracic Society, 1978). The link between childhood disease and subsequent development of chronic bronchitis is weak, however (Hamilton, 1981). Children may provide a particularly useful measure of these air pollution effects. Confounding factors which greatly complicate most epidemiologic investigations, such as smoking and occupational exposures, can be greatly minimized in studies involving child subjects.

Ventilatory function effects have been demonstrated experimentally in persons exposed to sulfur dioxide, ozone, and nitrogen dioxide at relatively high levels. The major effect of an increase in airway resistance is generally rapidly reversible. Severity of effect is dependent on the pollutant's chemical composition and its physical form. For example, larger particulates and more soluble gases impact the nasal cavities while finer particulates and less soluble gases proceed much farther, perhaps to the alveoli in the lower lung. Measures of pulmonary

function which may be tested include airway resistance, expiratory flow, vital capacity, and diffusing capacity.

3.1.3 Chronic Obstructional Respiratory Diseases

The chronic diseases and symptoms which follow are often investigated in air pollution epidemiologic studies.

Bronchitis: The main feature of early chronic bronchitis is hypertrophy of the mucus secreting glands and goblet cells, representing various nonspecific reactions to the bronchial mucosa (lining) from a number of different respiratory tract irritants. Excessive amounts of mucus may provide a mechanism for the spread of infection towards the periphery of the lung and therefore lead to acute exacerbations. Other pathological changes in the bronchioles include the formation of abscess cavities, the obliteration of the lumen and dilation. More severe histological changes in the epithelium of the trachial bronchial tract found in cigarette smokers include hyperplasia (abnormal increases) and disorderly arrangement of the basal cells, increases in the number of abnormal cell rows, a loss of cilia, and the presence of atypical cells (Holland, 1972).

Various acute inflammatory changes in the course of chronic bronchitis may result in acute bronchiolitis and later, broncho-pneumonia. Structural changes include fibrosis centered around bronchioles, the last stage of fibroproliferative inflammation (Lee, 1977). This inflammation leads to nodules of collapsed alveoli, resulting scar tissue, and the development of emphysema (Lynn, 1976).

Chronic bronchitis, though not generally fatal, is a significantcrippler of the middle-aged (Lynn, 1976). It is characterized by chronic cough, which is an attempt to rid the respiratory tract of excess mucus resulting from an over-secretion, and reduced cleansing due to inhibited ciliary activity. Bronchitis is diagnosed as "chronic" if the cough is present on most days for a minimum period of three months of the year and for not less than two consecutive years (Doulls,1980)). Symptoms are exacerbated by cold and damp weather which reversibly decreases mucus transport.

Emphysema: This disease is usually defined in terms of pathological changes within the lung. Pulmonary emphysema is a progressive destructive change in the alveoli, the pinhead-sized air sacs where gas exchange between the blood and the air takes place. Alveoli are thus combined into one larger volume. The resultant decrease of alveolian surface area causes the inefficient use of the lung and increases the risk of pulmonary hypertension and ultimately, cor pulmonale, "cardiac failure" (Holland, 1972). Chemical reactions with the lung elastic tissue, elastin, cause lung tissue to become less compliant and lower the elasticity and recoil of the lung. This tends to inhibit exhalation and increases the use of chest muscles (Lee, 1977). The earliest symptom of emphysema is a shortness of breath, first noticed on exertion or morning rising, which progressively worsens (Lynn, 1976).

Reduced resistance to infectious disease: An apparent breakdown of the defense mechanisms of the lower respiratory tract reduces the ability to eliminate inhaled or aspirated infectious agents, thus increasing the susceptibility to infectious diseases. The most agents isolated from the

sputum include the following organisms:

Mycoplasmas: *Mycoplasma pneumoniae* frequently has been associated with chronic bronchitis. In animals, these agents cause damage similar to human chronic bronchitis: in some of these studies, the course of the disease is accelerated by chlorine gas. From rather circumstantial observations, questions arise as to whether some cases of human bronchitis may be related to such persistent mycoplasmas with air pollutants acting as adjuvants (Holland, 1972).

Bacteria: Many studies have shown that patients with chronic bronchitis often have a permanent bacterial colonization of the lower respiratory tract. Species most often found are *Haemophilus influenzae* and *Streptococcus pneumoniae* (Holland, 1972).

Respiratory diseases associated with air pollution have been described in pathological and structural terms. A description of the physical, chemical, and biological mechanisms which cause the various diseases is beyond the scope of this paper. However, we now turn to the relationship between causal factors and these diseases, and outline a framework for analysis.

3.2 Causal Factors

The association of the diseases previously described with high pollutant levels present in acute air pollution episodes has been firmly demonstrated. There is considerable disagreement as to which chemical specie(s) is the causative agent or whether the correlation in such episodes is indicative of other processes yet undiscovered (Hamilton,

1981). At lower exposure levels -- at or near the ambient air quality standards for the criterion air pollutants -- there remains considerable controversy as to the form or existence of dose-response dependencies. The relatively low and continuous doses typical of ambient air pollution exposures exacerbate the difficulty of describing associations, let alone cause and effect, with limited data (MacMahon, 1979).

Causes may be defined as an act, event or state of nature which initiates or permits, alone or in conjunction with other causes, a sequence of events resulting in a health effect (Rothman, 1976). Respiratory diseases probably are caused by many factors, only one of which is air pollution. A cause which inevitably produces the effect is a sufficient cause. Unusually high levels of air pollution which result in acute health effects, e.g., bronchoconstriction, may be considered a sufficient cause. At lower levels, however, air pollution is more likely to constitute a component cause, that is, a cause which either aids or facilitates the development of disease. Thus, one possible effect of pollutant exposure is a synergism with other causes, where joint effects of the causal factors exceed the sum of the separate effects. All causes may vary in strength. Component causes may also be necessary causes that supply a required agent, e.g., an infection, which with other component causes produces the effect.

Establishing proof of causality is complicated by the likely multiple causality of diseases associated with air pollution. Personal lifestyle habits, such as smoking and exercise, may be stronger causal factors than air pollutants. In addition, the presence of environmental pollutants, or, more precisely, physical or chemical agents capable of adversely

affecting man or other organisms, is a very pervasive phenomenon (Murdoch, 1975). There are literally thousands of types of air pollutants, composed of both man-made and natural substances to which all people are exposed. Because many types of pollutants are inhaled, isolating the effect of a single pollutant is difficult. In some instances, other pathways of exposure to hazardous substances may be important. These include skin or mucus membrane contact, ingestion, and parenteral means (intravenous or intramuscular injection).

Of course, the major interest of this paper concerns the ability of air pollution to cause disease as a result of its toxicity. Toxicity may be defined as the net result of two competing processes: the effects of the toxic agent in the body; and the body's adaptive or homeostatic reaction which attempts to counter the toxin's effect (Ashford, 1976). Thus, effects may not be observed until a certain level -- the "threshold level" -- of dosage or exposure has been reached ^{1/}. Figure 3.1 illustrates the threshold effect. No symptoms occur (i.e., minimum net effect on curve C) until dosage exceeds the threshold level d. At lower levels, the body's increased homeostatic reactions counter the stress or the damage caused by the pollutant. Figure 3.2 demonstrates some of the complexities present in the dose-response model. In this case, the toxicity of the dosage (curve B) results from two or more component causes in ambient air pollution, e.g., sulfate and particulates, whose

^{1/} This discussion does not apply to radiation damage and carcinogenesis. Cancer initiation apparently occurs without regard for body defenses; thus there is no threshold effect. Cancer development, however, might in other ways be analogous to toxicity dose-response relationships.

Figure 3.1. Hypothetical linear toxicity, defense and net effects vs. dosage.

Figure 3.2 Hypothetical nonlinear toxicity, defense and net effects vs. dosage.

probability of resulting in a damage event is directly proportional to the dosage of each substance. The body's defenses (curve A) effectively counter the pollution exposure until point f. At this dosage level, defenses become less effective, due perhaps to fatigue or saturation of ciliary cleansing action. Effects are not manifested until threshold f, and increase rapidly with higher dosages.

The determination of "dose-response curves" and "dose-response relationships" between pollutant species and health effects is one of the major roles of toxicology and epidemiology (Ashford, 1977). Dose-response curves are developed by a mechanistic and/or mathematical description of the biological processes relevant to the health effects (Ricci, 1981). Toxicology seeks to establish the dose-response relationships, that is, a description of how biological organisms' respond to a toxic substance as its exposure to the substance increases (Ashford, 1976). This "response" may be a measure of a specific body function, e.g., lung capacity, or a gestalt, e.g., frequency of asthma attacks. By contrast, epidemiology deals with "population dose-response relationships" which estimate the excess risk factor for an individual, or the number of individuals in a given population who are expected to be affected. These investigations may quantify or confirm the dose-response relationship established in a clinical fashion by toxicology. These epidemiological "public health models" are discussed in Section 3.4.

3.2.1 Problems in Dose-Response Estimates

There are several difficulties in the identification and

quantification of dose-response dependencies of harmful environmental agents. The first problem is the great amount of uncertainty present in both health effects diagnosis and exposure estimates. One form of uncertainty is due to an information problem in obtaining both accurate diagnosis of health effects and relevant records of pollutant exposure and other factors which may have contributed to the disease. In general, verification of morbidity data is particularly troublesome and difficult. Respiratory diseases and ailments are commonplace. Moreover, symptoms for many diseases vary in severity and effect among individuals. Thus, diagnoses are frequently inaccurate and unconfirmed. As a result, epidemiological investigations often use death as an indicator of health effects due to the greater certainty in the diagnosis of mortality as compared to the diagnosis of morbidity. A second source of uncertainty is a lack of scientific knowledge of the factors which cause disease, as described previously.

The second difficulty in establishing the dose-response relationship is the presence of risk and confounding factors. The control of confounding factors might be regarded as the central methodologic issue in epidemiological research (Rothman, 1975). Confounding is defined as a distortion of effect introduced by an extraneous variable, which occurs only in the context of a particular study. The variable must be associated with the outcome or disease and be associated with the exposure. For example, if ignored, the socio-economic status of the individuals in the treatment groups may tend to confound associations between respiratory effects and pollutant levels. Individuals in higher socio-economic classes tend to live in less polluted environments and be

nonsmokers, as well as have better diets, less physically strenuous jobs, better medical care, etc. Thus, they may experience fewer severe respiratory effects than an individual in a lower socio-economic class. Strictly speaking, factors which are not necessarily or randomly associated with exposure, e.g., sex and sometimes age, are not confounding factors, but risk factors. These factors require standardization. Well known, for example, are age adjusted mortality tables.

The third difficulty in establishing the dose-response relationship is paradigmatic blindness and conservativeness. The relatively new science of epidemiology has developed its own traditions which tend to reflect the generally conservative practices of the larger medical establishment. In large part, this is due to a misunderstanding and suspicion of epidemiological methods. Epidemiologic investigations are generally statistical in nature: results may be challenged on either their statistical basis, generalizability, and (of course) policy implication. Two effects manifest from this environment. Blindness, that is, the use of inappropriate hypothesis on which studies may be formulated and tested (certainly not unique to epidemiology), results when investigators study what medicine may already "know," and thus miss important new relationships. The second result, methodological conservativeness, is demonstrated by the prevalent use of "tried and true" techniques, ranging from data gathering tools like surveys, to standard statistical data analysis techniques. "Heretical epidemiology" might, for example, account for the precision, validity and certainty of all data used in the study in a more analytic fashion than what is

currently done. Presently, investigators do not weight the data or develop confidence intervals or distributions for input variables (as distinguished from that done for results, which is standard practice) as would a stochastic analysis. Rather, investigators make an analysis of the data, for example, making a diagnosis from available data, and then proceeding deterministically. Of course, epidemiology, as all other sciences, attempts to incorporate new techniques into its methodologies. For acceptance in the larger medical community, which generally does not have an analytic or engineering background, the techniques must be communicable and understandable. Moreover, they should be generalizable as verified through use in a number of different settings.

3.2.2 Construction of Dose-Response Relationships

Most effects of ambient air pollution are far more subtle than those of substances usually considered in toxicology. However, health effects from exposures to air pollutants may be divided into three general toxicological classes: acute, chronic and latent effects. Acute effects appear immediately following or shortly after exposure, for example, bronchoconstriction and other ventilatory function effects. In these cases, the determination of contributing causal factors is generally straight-forward although the identification of the causal agent may be difficult.

Chronic effects develop gradually over time as the result of two consequences: "classical" chronic toxicity, and cumulative chronic functional dispairment. Classical chronic toxicity is the accumulation of

toxins in the body over time that results in injury when the effective threshold concentration is reached. This category is relevant for respiratory effects whose severity is contingent upon the rates of several competing or antagonistic processes. For example, the early stages of bronchitis, characterized by chronic cough, may result when the rate of mucus transport is less than the rate of mucus secretion. It is also relevant for the inhalation of hazardous metals, e.g., mercury and lead. Chronic cumulative functional disimpairment, that is, the accumulation of irreversible damage events, for example, the destruction of alveoli, appears the more relevant type of chronic toxicity for air pollution related emphysema. Both types of chronic effects often get worse with increasing age. Thus, the medical definition draws an arbitrary distinction between what is considered to be normal and healthy, and unhealthy.

A third category of latent effects, that is, diseases which appear considerably after the damage or initiating event, are related to molecular biological diseases, e.g., cancers, and are beyond the scope of this report. However, the term "latent" is used in several other ways which may be confusing. Latent effects and latent periods, that is, when a significant period of time elapses after damage events but before symptoms are observed, are called empirical induction periods (EIPs) here, following Rothman's terminology (1976).

Dose-response curves model the relationship between exposures and the incidence of disease in a population. Several general types of population dose-response curves may be observed, and depend primarily upon the mechanism of the toxic action and the type of defenses available to the

body for removing the agent and repairing the damage (Center for Policy Alternatives, 1977). In all cases, the uncertainties (and controversy) increase at low levels of exposure. Figure 3.3 illustrates the several types of dose-response functions. Broken lines indicate the regions of uncertainty. Curve A may represent a linear dose-response relationship, where the probability of symptoms for an individual is directly proportional to dosage. This curve illustrates a response which, if extrapolated, appears to have a "zero threshold" at the origin, although a threshold e' may exist, below which no effects are observed. A "natural" rate of incidence may exist, that is, there is a (low) prevalence of the symptom without any exposure, as indicated by point e . For a dose-response relationship which clearly shows a threshold effect, i.e., curve C, the level of the threshold may vary a considerable amount, between f and f' , depending on the individual, stress, previous exposure history, personal habits (smoking), etc. Curve B illustrates a non-linear toxicity which might be dependent upon the joint probability of two independent agents. Damage increases slowly until both agents are at relatively high levels. This type of function is portrayed in "multihit" cancer models (Section 3.4.4).

Figure 3.4 illustrates how the family of curves in Figure 3.3 may be modified by socio-economic, environmental and personal factors. For any given degree of incidence or risk factor, (or conversely, at any given pollutant exposure), an older or more susceptible population, with repeated exposure, or in the winter months may tend to increase the frequency of observed health effects. In this instance, the threshold for curve C has lowered to f'' .

Figure 3.3. Hypothetical static dose-response functions [developed from Ashford, 1977].

Figure 3.4. Hypothetical population dose-response curves [American Thoracic Society, 1978].

Dose-response relationships also appear to be dependent upon the frequency and pattern of air pollution exposures. Exposure to air pollution is a continuous phenomenon. We may, quite arbitrarily, split exposures into two components: low level background exposures, and episoidal exposures at greatly elevated pollutant levels. Figure 3.5 shows the degree of impairment an individual might suffer over a pollution episode which lasts three days. Acute effects, represented by curve A, quickly arise and disappear totally a week or so after the episode. Curves B and C, representing chronic functional disimpairment, reversible and irreversible, respectively, begin to peak near the end of the episode, due to latency or a different threshold phenomenon, e.g., a minimum product of exposure and time requirement, than the acute effects.

Figure 3.6 displays health effects over a period of two similar pollution episodes. Dose-response relationships are similar to Figure 3.5. However, acute effects, curve A, are greatly diminished in the second episode, due perhaps to some acclimitization. Reversible chronic effects, curve B, increase in severity, possibly due to the increased toxicity of the pollutant vis-a-vis a weakened defense system. Irreversible chronic effects appear additive, with the same increment of permanent damage as occurred in the first episode.

Of course, respiratory diseases and symptoms are comprised of many acute and chronic effects, each different in severity, dose-response relationship, and time factors. Thus, the pathological components of health effects, which have been artificially separated in this section, can only describe in general terms a model of disease and exposure. In summary, dose-response curves capture only a portion of the information

Figure 3.5. Hypothetical acute and chronic dose-response effects over an air pollution episode.

Figure 3.6. Hypothetical acute and chronic dose-response effects over two closely spaced air pollution episodes.

necessary to predict health effects. The experimental evidence established in toxicological investigations must be supplemented with epidemiological evidence in order to estimate the parameters of the dose-response relationship. Only then, and with many caveats, may predictions of the health effects or excess risk which is caused by air pollution exposure be made with scientific justification. Probabilistic and statistical health models which are used for such predictions are described in Section 3.4.

3.3 Air Pollution Epidemiology

The definition of epidemiology usually involves the concept of the group or population experience as the object of study, with the implication that biological phenomenon are predictable for a group, but not necessarily for an individual. Hippocrates recognized the fundamental objective of epidemiology 2400 years ago -- understanding the relationship of a population's disease experience to its environment (MacMahon, 1977). However, the relatively recent science of epidemiology aims to go beyond a description of the distribution of disease, that is, the health status of the population in relation to age, sex, race, geography and other factors. It strives for an analytic determination of causation, that is, an explanation of the patterns of disease in terms of (environmental) causal factors (MacMahon, 1968).

The strengths and limitations of epidemiology (as opposed to toxicology) are determined in large part by its two principal characteristics: its predominantly observational nature, and its concern

with human studies (as opposed to animals). Thus, there is no need to extrapolate inferences across species (MacMahon, 1979). However, because of differences in lifestyle, environment, diet, genetic composition, etc., epidemiologic studies almost always require repetition (as do other sciences) and different means of study under different circumstances unless the health impact is unusual (like the occupational exposure of vinyl chloride and its associated angiosarcoma of the liver).

A second major limitation in epidemiological studies is time limits. These impose constraints when the object of study, man, has a lifetime approximately equal to that of the investigator. It is extremely difficult to identify the causal factors for diseases that develop only after periods of time that represent substantial fractions of an individual's lifetime 1/. Additionally, inappropriate time frames for pollutant exposure estimates, relative to the diagnosis of health conditions, generally tend to dilute the strength of the association.

Epidemiological research has two general goals: precision and validity (Morrison, 1980). Precision concerns statistical error, that is the confidence interval which surrounds estimates of health effects. Large sample sizes are necessary to gain statistical significance or a lack of random error when a meaningful effect or health impact occurs infrequently. Because ambient air pollution exposures present a relatively small risk factor, that is, a minor risk among (other) component causes, to any individual, most epidemiological investigations

1/ Retrospective cohort studies may overcome some data collection problems relating to time and duration of exposure.

use very large sample sizes, often many millions. This results in very aggregated data which may tend to evenly distribute risk factors which vary among individuals and groups, mitigating the need to standardize the risk factors. However, large sample sizes do not guarantee the elimination or adequate control of confounding factors. Moreover, large investigations can usually only assess nonspecific and unconfirmed health impacts for which standard medical files exist. Data validity checks in such studies are, for all practical purposes, impossible 1/. An alternative, a smaller, more detailed study, permits reinterview and validity checks on the data. Standardization may compensate for disparity among groups if the risk factors are known. In any study, costs of data acquisition and analysis mount rapidly as the scale increases.

Validity relates to systematic errors or biases of both subject group comparability and information quality. A valid result describes the exposure disease relationship correctly. Comparability of groups means that in the absence of the true exposure experience disease relationship, no manifestations of effects are expected to be observed. Information quality means that little or no misclassification of groups, exposure status, or incorrect diagnosis of outcome or disease occurs (Morrison, 1980). Bias may weaken associations when the true symptom or disease is obscured by other diseases. In addition, Hawthorne or interviewer bias may mis-estimate the effects experience.

Bias also may occur in estimating pollutant exposure. Existing air

1/ A central data source for morbidity and mortality information would be an invaluable resource for epidemiological investigations.

quality monitors are placed in locations that are not representative of dosages received by the population. Similarly, indoor air quality may differ significantly from outdoor air quality. In general, measures of air pollution used in epidemiological investigations only recently have been treated as other parameters in the study. In the past, most measures of air pollution were neither statistically precise nor scientifically valid. They were used simply because no better measures existed. These issues are discussed in the following section.

3.3.1 Methodological Challenges -- Confounding Factors

The control of confounding, that is, accounting for factors which would tend to invalidate results, is a general problem in using statistics, but is especially important in air pollution epidemiology. The evaluation of a fluctuation in common respiratory diseases is very difficult. Most diseases associated with air pollution are likely to have many component causes, which are not necessarily associated with ambient pollution. Thus, a background of "natural" incidence, which itself varies with time, geography, and other factors, tends to confound epidemiological studies. A major failing in many air pollution epidemiological investigations is the inadequate control of confounding factors. Confounding may either strengthen or weaken the association between air pollution and health effects. This problem has helped prevent the establishment of strong associations between air pollutants and health effects.

Confounding is generally controlled by either of two methods. Stratification restricts comparisons to more or less homogeneous categories of risk indicators and measures the (treatment) effect. For example, in a longitudinal study within one given city, observations of health effects in nonasthmatic, married, nonworking white women between the ages of 35 and 55 for a period of several days or weeks during an air pollution episode, may provide a relatively unconfounded measure of air pollution related mortality or morbidity 1/. Stratification may reduce the generalizability of the study population, be inconvenient to the investigator or study design, and not provide the desired exposure ranges. Still, when possible, stratification is preferred over multivariable regression techniques because it permits the investigator to retain a familiarity with the data and usually allows a clearer interpretation of results (Rothman, 1978). Since confounding frequently is pervasive in air pollution epidemiological investigations, a combined approach of both stratification and multivariable regressions is often desirable.

Multivariable analysis is the second method of controlling confounding factors. It is commonly employed in cross-sectional studies, and may efficiently cope with the effects of several confounding variables while retaining sufficient information about the effects of pollutant exposure to obtain reasonable estimates. The disadvantages of

1/ Shorter periods may manifest idiosyncracies in reported effect due to the day of week, holiday occasion, etc. Longer periods may experience weather changes which also tend to confound the association of the air pollution exposure variate.

multivariable analysis are 1) the need to construct an adequate functional model for the relationships of the treatment (pollutant measure) and the confounding factors to the outcome; 2) the theoretical need to fulfill certain assumptions about the evidence which are rarely achieved; and 3) difficulty in interpreting the multivariable model to learn about the treatment effect (Rothman, 1977).

It is necessary to control three classes of confounding factors in cross-sectional air pollution epidemiological studies: personal, socio-economic, and environmental factors. In general, quality information is not available and surrogates or substitutes must be used. Personal factors include race, age, sex, and smoking history or cigarette sales. Often, particularly for socio-economic and environmental factors, there are few ways to measure parameters such as job stress at work and at home, or the health risk of the environment. Socio-economic indicators have included income status (poverty or other classification), education, marriage status, birth rate, population density, housing status (age), and mass transit use (Lipfert, 1980, Viren, 1978). Environmental factors have included water hardness, geography, and climate. These factors may be combined into indices to simplify analysis, like the socio-economic class mentioned previously. The construction and physical meaning of such indices is often unclear. Additionally, it is rarely possible (especially a priori) to verify that these factors or indices are valid and precise measures of the intended effect. At present, the selection of the factors used in any study depends on the wisdom and experience of the investigators, though general patterns are beginning to emerge. At a minimum, personal factors of age, sex and smoking history need to be

controlled in air pollution health effects research. Also, to determine at least a crude indication of worker exposure and stress, some measure of socio-economic status is necessary. Longitudinal or stratified studies generally do not require the control of environmental factors because these variable are assumed to be constant for the entire study.

A serious problem in the control of confounding is multicollinearity, or inter-relation between variables. Introducing more variables to control confounding may in fact exacerbate problems due to the seriousness of this problem. For example, an individual's air pollution exposure estimate may be related to his or her socio-economic class. This problem arises because air pollution exists in virtually every industrialized part of the country. Therefore, mortality and morbidity effects may not be reliably separated from the characteristics of the area or population, despite very satisfactory fits and reasonable explanations, in terms of explained variance and sign and magnitude of coefficients in regression equations (Lipfert, 1980). Principal component analysis and ridge regression are designed to mitigate effects of variable dependence that may plague multivariable regression analysis but present other problems that limit their use.

Certain decision analysis techniques might be valuable for epidemiological and risk-cost-benefit purposes if both the toxicologic action of an agent are well known and a great level of detail has been assembled. It is unknown whether this has been attempted in this field. These might provide a hierarchal description of states, links and causal relations (Szolovits, 1980). Then, the influence of confounding factors might be minimized by a selection of the appropriate variables or parameters for

each individual or cohort in the study, using expected value criteria.

3.3.2 Indoor Air Quality

Since Americans spend most of their time indoors, indoor air pollution may comprise a major portion of their pollutant exposure. In homes, for example, indoor air quality is affected by the presence of smoking, the types and condition of heating, cooling, ventilation and cooking systems, as well as the outdoor air quality. Without significant indoor pollution sources, pollutant levels are generally lower indoors than outdoors 1/. However, much higher indoor concentrations of several gases and respirable particulates are sometimes found indoors. For example, levels of carbon monoxide and nitrogen dioxide frequently exceed the NAAQS for short-term concentrations (Hollowell, 1980; Winkelstein, 1979).

If not controlled, indoor air quality may seriously confound an epidemiological investigation because these indoor/outdoor air quality relationships are not randomly distributed across time and/or geographic units. Broad geographic patterns in cigarette consumption exist, and seem associated with a pollution gradient, as well as the degree of urbanization. Fuel choice, system type, and degree of usage of heating, cooling and cooking systems exist in regional and temporal patterns.

1/ This applies to the criteria pollutants. Building materials may emit levels of aldehydes, radon, etc., which may be found at higher levels indoors than outdoors.

Heating and cooling systems are important in that they alter the air exchange rate, provide some filtering (as well as sometimes having combustion problems). Thus, they may serve to contain indoor air pollution if the concentration is higher indoors than outdoors, or in the opposite situation, may tend to improve indoor air quality. Gas and propane stoves may contribute major sources of nitrogen dioxide to the indoor environment, especially, to the kitchen area, if the air exchange rate is low, as is likely in winter time. In smaller epidemiological studies which provide information on gas stove use as compared to non-emitting electric stoves, stove use has been shown to significantly effect the frequency of childhood respiratory diseases (Speizer, 1980). Stove use is also associated with socio-economic status, which tends to decrease the significance of the air pollution variable in such studies.

3.3.3 Study Types

Numerous epidemiological studies have examined the association between air pollution and the prevalence of chronic respiratory disease, that is, the proportion of the population reported suffering from such disease symptoms at a particular time. These studies have used prevalence surveys, or questionnaires of the sample population, work absence and sick leave comparisons, incidence surveys, that is, the probability over time, of worker incapacity, various diagnostic tests for pulmonary and respiratory function, and autopsyd group comparisons. Most studies of chronic disease and air pollution have indicated a relationship between the prevalence of chronic respiratory symptoms and the sulfur oxide and

particulate complex (American Thoracic Society, 1978).

Variations in mortality rates have been observed with short term or daily differences in pollution concentrations. Such longitudinal studies potentially are valuable to detect threshold concentrations above which excess mortality becomes likely. These studies are particularly convenient to conduct because confounding is unlikely. On the whole, they tend to agree in finding greater mortality (for all sources combined) on days of greater pollution. No clear threshold for this effect has been shown (American Thoracic Society, 1978).

Studies have also attempted to associate mortality variations to chronic exposures to air pollution. These studies, which must be interpreted cautiously due to the serious confounding often prevalent, imply a relatively consistent association between long term residence in more polluted communities and long term mortality rates (American Thoracic Society, 1978).

3.4 Stochastic Health Models

Health models are constructed to provide (hopefully) both consistent and valid estimates of pollution related health effects. Stochastic models couple exposure data to demographic distributions and other information to produce probabilistic estimates of incidence or risk of disease in the exposed population. This is in contrast to deterministic models which might, for example, predict the actual number of new cases in a given time period. Many types of models have been developed to match the different toxicological actions of various environmental

contaminants. Mathematical treatments generally have not been very sophisticated in the air pollution field due to the likely multiple causality of respiratory diseases and the great uncertainty in the data. Most air pollution epidemiological models use multivariable regression analysis, a mostly descriptive framework rather than a more physically based model, to provide an assessment of harm to the population at risk.

Health models map health effects onto either a general population or to specific groups, for example, asthmatics. Most models are quite crude and may present serious shortcomings for policy applications. Generally no information concerning the temporal and spatial variability and distribution as well as the probability of the projected health effects are obtainable from any single model. Temporal concerns include lag effects, due to induction or latent periods, which are important with many chronic respiratory diseases. They also include changes in the age distribution of the population which occur over long periods of time. Spatial distribution might include the location and types of people (in terms of susceptibilities) who are likely to be affected. In all cases a homogeneous population or a uniform distribution of susceptibles is assumed. Lastly, although most health models are probabilistic, they say nothing about the likelihood of their projections, that is, the probability distribution of projected health effects. This would give, for example, the most likely excess mortality estimate, as well as 100% error bands above and below the estimate. Such distributions are just beginning to emerge although they are derived from an examination of alternative health models, each of which uses fixed health effects coefficients (rather than a probabilistic distribution) and which, to a

varying extent, are incompatible with one another.

The mathematical handling of disease incidence and prevalence as a stochastic process with continuous functions is both convenient and suitable for investigating such large scale phenomenon as air pollution related diseases. For many types of health effects, the behavior predicted by a stochastic model is likely to be similar to that described by the corresponding deterministic model. Large sample sizes, often many millions, tend to greatly reduce the statistical fluctuations. Additionally, investigations of chronic air pollution health effects require long periods of time and repeated observations which may further reduce the random element. However, acute effects predicted by a stochastic model, especially for a small population, are likely to be quite different from that predicted by the deterministic model. In general, stochastic models have broader distributions, lower maxima (or minima) and greater asymmetry. Deterministic models might assume, for example, that the risk factor for the population, on the average, is related only to the air pollution exposure. On the other hand, models which have been constructed in a stochastic fashion might account for the lower joint probability of an individual responding to air pollution as well as that individual experiencing the minimum dosage for health effects. Figure 3.7 displays the "epidemic curve," that is, the rate at which new cases are recognized over the duration of an air pollution episode. Maximum rates are lower for the stochastic model. However, the total number of cases remains the same, although they occur over a longer period of time. Figure 3.8 shows the distribution of expected incidence at several exposure levels on a population dose-response curve. In this

case, a "surface" is developed from distributions on the median dose-response curve, line A. This shows that there is a probability that some fraction of the population will respond to pollutant levels significantly lower than the levels to which the majority of the population responds. The broad distributions, particularly at low levels, make model fitting difficult. The functional form of the dose-response relationship, that is, whether it is linear, greater than linear, with a threshold, etc., has important policy implications, as discussed previously in Section 2.3.

Figure 3.8 also demonstrates some of the policy concerns in the selection of statutory standards, or the use of risk-cost-benefit analysis for air pollution regulation. If the dose-response line A has been developed for the "average" population, then the distributions above this line represents the effects suffered by the more sensitive half of the population. If exposure level c is the "threshold" standard, which incorporates a safety factor of two over observable health effects related to pollutant exposure, then the most susceptible of the sensitive individuals, that is, those in the upper tails of the distributions, will still suffer adverse health effects, even though the exposure level is below c. Thus, these population dose-response surfaces may be useful for risk assessments because they incorporate responses of all portions of the population.

There are at least three major classes of stochastic health models. These include: 1) regression models, including static (descriptive) and dynamic (cigarette) models, 2) infectious epidemiological models, and 3) multistage multihit (cancer) models. Such models offer some complementary

Figure 3.7. Hypothetical "epidemic curve" [Bailey, 1975].

Figure 3.8. Hypothetical distributions on a population dose-response curve [E.P.A., 1980].

mentary concepts to the study of air pollution health effects, including a more sophisticated and biologically based treatment in comparison to the regression models. A brief survey of several model types follows.

3.4.1 Regression Models

Most air pollution epidemiologic investigations use multivariable regression models and ordinary least squares approximation (OLS) techniques. In its simplest form, the functional relationship between pollutant exposure and health effects is a simple linear model (pollution measures, confounding factors, etc., may be described in matrix form):

$$\underline{M} = \underline{A} \times \underline{C} + B \quad (1)$$

where $\underline{M}$ = dependent variable, e.g., morbidity or mortality
(or, in matrix form, a vector of observations on M).

A = proportionality constant (or vector of coefficients to be estimated).

$\underline{C}$ = pollutant exposure estimate (or matrix of observations of explanatory independent variables, e.g., confounding factors).

B = residual.

Such descriptive models are developed to explore the relative significance of various risk factors, including, but not limited to air pollution, in causing or contributing to adverse health effects. More advanced treatments which reflect physiological and toxicological understanding are generally neither used nor formally incorporated into most regression models. In large part, this is due to the very poor data quality typical of most studies. Although sometimes contested, strictly speaking, such models explore more than the correlation between health effects and air quality. Regression models are causal models: health outcomes are determined by the independent variables, which include air pollution. This confusion arises because air pollution is often strongly associated with other stressors, like income level or population density, in health effects investigations.

Ricci (1980) has outlined statistical points of contention in regression models which may lead to inconsistent, inefficient or biased estimates of coefficient magnitude and statistical significance. In roughly decreasing order of significance they are 1) variable dependence, 2) specification error, 3) misclassification (observation error), 4) functional mis-specification, 5) autocorrelation, 6) missing observations, 7) misinterpretation of significance level with several hypotheses, 8) instrument variability (errors due to surrogates), and 9) violation of normality assumption. Remedies to some of these problems that affect OLS approximations include data transformations (especially logit and logarithmic), ridge regression, robust estimation techniques, weighted least squares and principal component analysis. Multi-estimation procedures also may be used, including two and three stage least squares

and full information maximum likelihood methods (Ricci, 1980). Other concerns relate to the confidence level selected. At the typical 95% confidence level, the air pollution variable(s) is often insignificant, and on the basis of the "p" value alone, is eliminated in step-wise procedures. Investigators, of course, have a substantial interest in keeping this variable in the regression. As Rothman (1978) pointed out, a better indication of the meaning of a variable may be found by an examination of the relative position of the confidence interval which surrounds the estimated coefficient value.

Typical problems due to multicollinearity have been described in 3.3.1 and 3.5.7. Misclassification and errors due to surrogate pollutant measures are described in 3.5.1. We turn now to a brief examination of some other functional forms of the regression model.

3.4.2 Cigarette Models

Cigarette dose-response relationships have been studied quite extensively. A simple model [Lundy, 1977] describes the degree of risk as an exponential function of dosage, possibly in terms of packs per day, in the numerator. The denominator describes a repair function which rises asymptotically to some maximum level of activity to repair the damage.

$$M(x) = \frac{a e^{(bx)}}{1+c e^{(-d(x-x_0))}} \quad (2)$$

where $M(x)$ = health impact.

x_0 = age at onset of exposure.

a, b, c, and d are estimated coefficients.

The principal advantage of this model is that it is able to incorporate latent effects. The damage term is dependent on the previous exposure. Thus, when matched to the projected age distribution of the population, this model may account for accumulated damage, which occurs as the population grows older. Consequently, given constant pollution levels, a greater level of adverse health effects will be forecast for the far future, as compared to the close future, with respect to single coefficient models.

3.4.3 Infectious Epidemiological Models

Infectious epidemiological models provide a more sophisticated treatment in contrast to simple regression models. These models generally deal with diseases that are capable of transmission at some stage in the lifecycle of the infectious organism from infected host to an uninfected susceptible individual.

The simplest deterministic infectious health model assumes that new incidence is proportional to both the number of infectives I and the number of susceptibles S . This formulation assumes homogeneous mixing of all individuals in both groups. The number of individuals is n , and the epidemic begins at time $t=0$ with only one individual infected with the rest susceptible, but as yet uninfected. The process may be described by a differential equation:

$$\frac{dI}{dt} = -BSI = -BI(n-I+1) \quad (3)$$

where B = estimated proportionality constant.

The number of infectives at time t may be derived from this formulation, resulting in the epidemic curve shown earlier in Figure 3.7 where

$$S = \frac{n + 1}{1 + n e^{-(n+1) Bt}} \quad (4)$$

Infectious epidemic epidemiological models are among the best developed of the mathematical health models. Numerous variations are described elsewhere [Bailey, 1975].

3.4.4 Cancer Models

Cancer models differ markedly from other health models due to two properties of carcinogenesis. First, cancers are believed to be single cell in origin. A minimum number of damage events or "hits" to a cell are required to develop the cancer. Secondly, cancer development may not spontaneously follow initiation: other changes at later times (stages) may be necessary to promote the malignancy. The basic multistage multihit model assumes that cells respond similarly and independently, and that the cancer incidence rate increases as a power of time, after a minimum number of hits k . If the induction time is independent of the dose and stochastically independent of the time required for cancer initiation, then $P(d,t)$, the probability of observing cancer at time t in a single individual subject to a constant dose rate d , is approximately

$$P(d,t) = 1 - e^{[-Q(d)L(t)]} \quad (5)$$

$$\text{where } Q(d) = \sum_{t=1}^k (a + bd)$$

$$L(t) = c \int_0^t f(t-u) u^{k-1} du$$

a,b,c are estimated coefficients.

A number of modifications of this model have been suggested by a number of authors. For example, "cell-killing" terms may be added, as well as intermediate stages or immunological phenomenon. However, most cancer models that depict exposures as affecting an ongoing process will lead to linearity at low doses. Crump (1976) has shown that for low doses, and over a wide range of estimated coefficients, the dose responses are approximately linear.

Despite several decades of general agreement in the "zero threshold" linear dose responses for radiation damage, there exists substantial disagreement as to the form of the dose-response relationship for other carcinogens in air, water and food. Recent evidence has increased the controversy of the threshold concept for non-carcinogens as well, including ambient air pollutants.

3.5 Dispersion Models

It is beyond the scope of this study to review the details of air quality dispersion models (AQDM). Only the briefest discussion of the various types of models available and their implications for health effects assessment is permitted. At present, the major use of these models is in the implementation of regulatory standards and policies.

These uses include managing air quality among many contributing sources (in prevention of significant deterioration [PSD] areas, and non-attainment areas) and for new source applications. Of particular interest is the recent use of these models in risk-cost-benefit research (not policy determinations) of major emission sources, especially power plants.

There are three generic classes of AQDMs: 1) simulation/numerical models, including Gaussian, grid finite element, pseudo-spectral particle in cell, and box models, 2) statistical/empirical models, and 3) physical models, involving fluid modelling. These models are also tailored to the number and types of emission sources (i.e., point, roadway or urban), terrain features (flat, valley, complex), type of pollutant (including "visibility") and short-range or local conditions (less than 50 to 80km) or long-range (mesoscale) transport.

Models are highly specialized tools. A model applied improperly or with inappropriately chosen data can lead to serious misjudgements regarding the source impact or the effectiveness of an abatement strategy. Gaussian models are the most widely used technique for estimating concentration of non-reactive pollutants, and are the best choice for modelling sulfur dioxide and TSP from point sources (E.P.A., 1980). Other types of models may provide more accuracy, but at considerably more cost, due to large computational requirements.

As inputs, simulation models require emission data, including pollutant emission, stack height, and emission temperature; meteorological data, including wind speed and direction, stability class, mixing height

and temperature; and source data, including local emissions and long-range transport. Truly valid data are probably the exception. Meteorological data are scarce: one weather station may be used to represent conditions over many thousands of square miles. Similarly, source data suffer from uncertainties due to low monitoring densities. Models assume particular rates of dispersion, removal, chemical and physical transformation (as discussed in 3.5.4) in a predictor equation. Such equations usually model the physical processes under ideal conditions. For example, they are not necessarily mass-conserving and thus may forecast more or less pollution than is actually emitted.

Dispersion models tend to be very inaccurate. A benchmark for accuracy is predicting 50% of all values within a factor of two. One stumbling block in the development of better models is the lack of agreement as to which performance measures and standards should be used to evaluate models. Current policies require that the model simulate the second highest concentration for averaging times of 1, 3, 8 and 24 hours, which are equivalent to the 0.02 to 1 percentile (out of 8760 hours in a year). Such extreme values are not well behaved from a statistical viewpoint (American Meteorological Society, 1980). Consequently, the E.P.A requires 5 years of continuous meteorological data or at least one year of on-site data -- if it can be shown to be climatologically representative -- to predict these rare worst-case conditions with any degree of confidence. AQDMs are better able to predict median values, i.e., average annual concentrations. Median values still may be confounded by background concentrations, which are either unknown or not properly accounted for in the source inventory. Long-range transport may also

contribute a significant background level. Lastly, model bias may systematically over or under-predict median values, depending on the model application.

Risk-cost-benefit analyses of ambient air pollution require population-weighted estimates of exposure. For these uses, models must thus provide accurate spatial projections onto whatever geographic units (Section 4.4) are used in the analysis, in addition to a meaningful prediction of peaks, frequency of peaks, and other measures 1/. Of course, for risk-cost-benefit assessments, it is desirable to predict dosage measures which best predict potential health impacts. With the large likelihood of potentially significant errors in dispersion models results, field validation is probably necessary, and certainly advisable, for risk-cost-benefit determinations of pollution policies which use AQDM.

3.6 Exposure Estimates

Pollutant exposure or dosage is a measure of cumulative exposure to pollution which is often expressed as the product of pollutant concentration and time, e.g., parts per million per hour, or micrograms per cubic meter per year. For risk-cost-benefit and regional assessments of health impact, exposure estimates are projected onto a demographic

1/ Present regulatory practice is mostly concerned with only concentrations -- not their locations. Thus, if a "hot spot" is predicted over a little or unpopulated area, e.g., a bay or river, it may be attributed to statistical error and not affect the determination of policy.

distribution, perhaps by a simple multiplication of exposure and affected population to yield, for example, person-years of exposure.

As other information in epidemiological research, exposure estimates should be precise and valid. Ideally, an exposure estimate for each individual that captured exposure during every phase of his or her activity would be provided. At present, the desirability and feasibility of using personal monitors for ambient gases and particulates (which, in the future might be worn like "radiation badges") is limited. Without exception, ambient air pollution epidemiological studies rely upon only very approximate measures of exposure.

In the past, epidemiological studies have used many surrogates or substitutes for measures of ambient air quality. In the 1940's and 1950's, indirect measures of air pollution concentration were estimated using the density and proximity of combustion sources, particularly coal furnaces. For example, the distance to power plants and the density of coal furnaces for home heating in a given area provided a relative measure of pollutant exposure. In the mid 1950's, when air quality monitoring stations were established in major metropolitan areas, (described later in Section 3.1), providing the first "hard" quantitative pollution data, the standard epidemiological technique was to use the concentration measures of the nearest monitoring station to characterize exposure for the study population. As a result of today's more stringent pollution abatement regulations which mandate increased monitoring, monitor density is higher in certain areas. However, the spatial distribution of monitors is very irregular, and not ideally suited for epidemiological purposes.

Air pollution dispersion models provide an alternative to on site monitoring. Such models provide the best prospective estimate of ambient air pollution from new emission sources such as power plants. In addition to such "point source" models, "line source" and "urban" models are available to estimate concentrations near roads and urbanized areas, respectively. These models are used regularly to maintain compliance with ambient air quality standards. Generally, it is preferable to use actual monitor data, when it exists, rather than model forecasts. This is due to uncertainties in the dispersion models and the data inputs which drive them, mostly emission quantities and meteorologic data. The uncertainty in these models has been discussed in the previous section. Simple dispersion models also may be used to effectively supplement sparse or inadequate monitor information, as described in Section 4.4.

There are five major and unresolved issues in air pollution exposure estimates for epidemiological investigations. These include 1) the selection of the pollution species to be monitored, 2) the location of monitoring site and choice of geographic unit, 3) the representation of exposure or dose, 4) the interactions of pollutants, and 5) the handling of data quality and uncertainty. These are discussed in turn in the following sections. Technical issues of monitoring, such as measurement error and cost, can not be discussed.

3.6.1 Selection of Pollutant Species

It is impractical to simultaneously monitor and assess the health effects of even a few dozen of the existing 60,000 or so species of air

pollutants. Most research is concentrated on the seven "criterion" pollutants for several reasons 1/. First, previous toxicological and epidemiological investigations have screened the more prevalent of the pollutant species. The ones that appeared to pose significant risks of harm were, of course, embodied into law. Thus, much of the emphasis on these pollutants seeks to reconfirm the validity of the Ambient Air Quality Standards. Secondly, because monitoring of these pollutants is legally mandated, a vast amount of data is available on them. Useful measures of other pollutants rarely exist, because monitoring techniques are not perfected or in widespread use .

Such a limited selection of pollutants may be made tenable by two assumptions. First, the reasonableness of the prior scientific investigations may be assumed. This implies that these earlier studies were methodologically correct and adequate in scope. Thus, all said, the criterion pollutants represent the major risk factors for damages caused by ambient pollution. Secondly, measures of the criterion pollutants may be assumed to represent levels of pollutants which are not measured. This may well be the case for some local pollutants e.g., carbon monoxide and sulfur dioxide, in some circumstances and for very short periods of time. In these instances, atmospheric ventilation tends to similarly affect pollutant concentrations, and interpollutant concentrations may be well

1/ Ubiquitous but indirect measures of pollution exposure, e.g., property or vegetation damage, may not be irrelevant measures of dose for epidemiologic purposes, though it is unknown whether research to transform this data to meaningful human measures has been attempted.

correlated on a time dependent basis. In general, however, dispersion, removal and chemical and physical transformation rates may vary significantly.

Consider emissions of a pollutant a emitted from a single source. The ambient concentration may be expressed as follows:

$$C_a(t) = E_a(t) \Pi_i(D_{ai},t) \Pi_j(R_{aj},t) \Pi_k(T_{ak},t) \quad (6)$$

where: $C_a(t)$ = ambient concentration of pollutant a at time t.

$E_a(t)$ = emissions of pollutant a at time t.

$\Pi_i(D_{ai},t)$ = product of all dispersion factors i on pollutant a at time t, including vertical (buoyancy/settling) and horizontal diffusion.

$\Pi_j(R_{aj},t)$ = product of all removal forces j on pollutant a at time t, including wet precipitation (washout, rainout), scavenging (adsorption), dry deposition (gravity settling, impaction).

$\Pi_k(T_{ak},t)$ = product of all transformation factors k on pollutant a at time t, including chemical and physical process, including oxidation, reduction, and adsorption onto particles and water droplets.

For pollutant b, which is assumed to be a derivative of pollutant a, e.g., sulfates (from sulfur dioxide), an equation similar to equation (6) may represent ambient levels:

$$C_b(t) = P_b E_a(t) \Pi_i(D_{bi},t) \Pi_j(R_{bj},t) \Pi_k(T_{bk},t) \quad (7)$$

where: P_b = proportion of emissions of pollutant b to pollutant a.

Assuming that pollutants a and b are produced in a constant proportion, at the emission source, a measure of pollutant a can serve as a perfect

surrogate measure for pollutant b only if all dispersion, removal and transformation coefficients for a and b are the same. Because polluted areas tend to be highly urbanized with multiple pollutant sources of various types, a pollutant measure may serve as a very approximate generalized indicator of total pollution. However, there are many circumstances, where the concentration of local, i.e., proximate, and "conservative," i.e., non-interactive pollutants such as TSP (not ozone), may vary widely enough not to be useful for epidemiological or risk-cost-benefit purposes. In these cases, pollutant concentrations change significantly within distances of a dozen kilometers or so.

For example, most sulfur emitted from fossil fuel combustion is in the form of sulfur dioxide. Some 3 to 4% is emitted as sulfur trioxide. In the atmosphere both of these species are further oxidized to sulfuric acid, a more toxic irritant. The speed with which this conversion takes place depends on all transformation factors for these chemical species, especially microclimate influences such as temperature, insolation, rain and humidity. However, the E.P.A. (1978) and many dispersion models condense the variability of transformation and removal rates into one measure known as pollutant half-life. The E.P.A. (1978) provides the following estimates of pollutant half-lives for sulfur dioxide, when average downwind distances are greater than 80km:

	<u>East of 100° longitude</u>	<u>West of 100° longitude</u>
rural	8 hours	16 hours
urban	4 hours	4 hours

Thus, only on the average, concentration due to an initial concentration

$c(t)$ may be expressed as a constant function of time and half life h by an exponential growth equation:

$$c(t) = c(t_0)e^{-0.693t/h} \quad (8)$$

For shorter downwind distances, where travel times are sufficiently short, no decay terms are suggested.

Thus, if only sulfur dioxide is measured and the rate of transformation is high, presenting a relatively low level of sulfur dioxide but a high concentration of the more dangerous acid radical, an overbiased estimate of the health effects related to sulfur dioxide is made. The converse, an underbiased estimate of health effects, is made when the conversion rate is low.

At present, most monitoring in the U.S. for particulates uses the Hi Vol Sampler which assesses particulates on a total weight basis, yielding a measure of total suspended particulates (TSP). Mass measures are dominated by the comparatively few large particles in the sample, because small particles have relatively little mass. Fine particulates, those smaller than 1 or 2 micrometers in diameter (and especially the sulfuric acid "mist"), are generally regarded as more hazardous than larger (coarse) particulates. Fine particulates are capable of penetrating the lower the lung while coarse particulates usually are effectively removed in the nasal and pharyngeal cavities.

Particulates share similar chemical composition in either of the two generalized size classes: fine (0.002 to 1.5 micrometers), and coarse (1.5 to 100 micrometers) [Perera, 1979]. Fine particulates generally result from human activity, in contrast to coarse particulates, which result from both natural mechanical action, including windblown dust, sea

spray, volcanoes, and plant particles, as well as human generated emissions. Large particulates tend to quickly settle out by sedimentation and long range transport is generally minimal.

Fine particulates come from both primary and secondary sources. Primary sources include combustion emissions, especially electric utilities which produce almost 75% of all fine particulates, as well as industrial processing, and metal manufacturing. These particles contain soot (carbon), flyash (silicon, aluminum, iron, calcium, carbon, sulfur, trace elements 1/), metal oxides and salts, sulfates and nitrates. Secondary sources of fine particulates result from chemical reactions in the atmosphere. These compounds are generally classified as acid sulfates, nitrates and organic matter, and are formed in the presence of sunlight from precursor gases, including sulfur dioxide, nitrogen dioxide, hydrocarbons, carbon monoxide and ammonia. Very fine particulates (less than 0.1 micrometers) condense and form larger particles (0.1 to 1 micrometer) which have residence times of up to several months. Thus, long range transport, up to several thousands of miles, is a frequent occurrence that contributes an often significant "background" level to the local pollution situation.

Most air quality dispersion models simulate the differences of pollutant dispersion between particulates and gases by changing those

1/ The concentrations of beryllium, cadmium, chromium, manganese, nickel, lead, vanadium and zinc increase with decreasing particle size. Note also the collinearity between TSP measures, and sulfur gas measures, due to the sulfur constituent in ash, and precursor role sulfur gases play in the formation of secondary sources of fine particulates.

coefficients which represent vertical and horizontal diffusion. In general, models do not distinguish between coarse and fine particulates (since the NAAQS do not either).

Across wide geographic regions, TSP measures can not be used as a good surrogate for fine particulates because the fine particulate proportion of TSP varies widely. In Point Arguella, for example, it is only 2%, while in St. Louis it is 76% (by weight) (Perera, 1979). Additionally, farming, construction and highway dust in the vicinity of TSP monitors may result in high TSP concentrations that are not characteristic of the region, nor indicative of the level of fine particulates. Another source of error in TSP and sulfur measures is the fact that TSP contains a variable amount of sulfur salts. Lastly, the contribution of secondary sources of particulates, which contributes a very important 60 to 80% of the mass of fine particulates in urban areas (Perera, 1979), may fluctuate greatly, in large part due to weather influences.

In summary, the correlation between pollutants is complex and variable. In short term correlations between pollutants, the variability is mostly random, due to stochastic microclimate influences. In the long term, and across geographically distinct regions, such variations are less random, due to averaged weather conditions and the differences in pollutant source inventories between the regions. The use of several years of pollution data in order to represent most nearly average meteorologic conditions, may mitigate some of the uncertainty in short term pollution assessments due to weather variability. The existence of different transformation, removal and dispersion factors for different

pollutants suggests that surrogate indicators of concentrations for pollutants are, at best, only very approximate. The importance of these factors for epidemiological and risk-cost-benefit purposes is proportional to the extent to which they occur, and to the degree to which health effects are related to pollutants that are not measured.

3.6.2 Monitor Location and Selection of Geographic Unit

Population-weighted pollution concentrations are the most desirable air pollution measures for epidemiological and risk-cost-benefit research, given the impracticality of determining individual exposures for large numbers of people. However, given the high mobility of the American public, exposure estimates may be in serious error if the residence history of the sample population is not considered.

The selection of monitor(s) location and geographic unit (e.g., census tract, city, county, etc.) are the most important considerations for accurate determinations of pollutant concentrations among sample populations. Typically, pollutant levels are measured at only a few locations. Urban monitors are located in "hot spots" in order to provide early warning of possible air pollution violations. On the other hand, rural monitors are often situated in relatively remote and unpopulated sites in order to provide pollutant measures of large sources such as power plants which are not affected by small local sources. It is the rare investigator who is able to locate monitors optimally for the sample population (which, in fact, may frequently change as different cohorts are studied).

The selection of geographic unit frequently requires a compromise between statistical precision, which requires large units like states, and accuracy of air pollution exposure, which is easier with small units (Lipfert, 1980). Monitor density determines the smallest practical unit due to the (excessive) interpolation which is required to obtain pollution scores for units between monitors. Substantial variations in pollutant level may occur within a relatively compact geographic unit like a SMSA. Thus, interpolations may not be accurate or adequate. A second, minor constraint on the selection of geographic units concerns the difficulties involved in obtaining information on non-standard geographic boundaries. Given all limitations, most cross-sectional studies settle on cities, counties or SMSAs although there exists considerable debate about the merits of each (Lipfert, 1980).

Lipfert (1980) has shown, for example, that population weighted TSP scores for the standard metropolitan statistical area (SMSA) study units used in the Lave and Seskin (1978) cross-sectional epidemiological study, varied in a non-random fashion. Population weighted scores ranged from 54 to 80% of the actual measures (for central cities) used in this study. Moreover, sulfate measures were biased yet more seriously, also in a regional fashion. Eastern suburbs had 27% lower mean sulfate concentrations than eastern cities, while western suburbs had only slightly lower levels (3.5%) than western cities, for the several regions investigated by Lipfert ^{1/}. These regional differences between east and

1/ In contrast to these "conservative" pollutants, oxidant levels frequently peak well downwind, often in the suburbs.

west are related to the types of air pollution sources and the influence of local sources versus long range transport. Without correction or compensation, such biases are likely to affect the stability of the pollutant regression coefficients (due to multicollinearity with other regional variables), as well as their magnitude, statistical significance and, in some cases, their sign.

3.6.3 Representation of Exposure

Previous discussion described pollutant exposure or dose as a measure of the product of concentration and time. In fact, however, there are many types of measures used, with varying effect, to express pollutant dosage. Ambient concentrations frequently fluctuate dramatically from hour to hour, as a function of meteorologic conditions and source emissions. Moreover, they often may appear to defy characterization on a monthly or even annual basis. In general, however, pollutant concentrations tend to be log-normally distributed with a greater frequency of lower values than high values. High values often tend to be extremes. Thus, arithmetic means tend to be higher than medians.

The essential information which is desirable in exposure representation relates to the clinical significance or health effects due to the pollutant concentration and the pattern of exposure. As a result of the long induction period and cumulative nature of many air pollution related diseases, such measures should tend to be stable over time. Additionally, representations should be analytically simple to derive (either directly from monitors, or indirectly, from dispersion models),

be adeptly transformable (permitting combinations with other pollutant measures in order to assess a broad spectrum of pollutants, and for various health effects and dose-response hypothesis), and ideally possess physical significance (primarily for an intuitive appeal, but also for regulatory purposes).

There are four major areas of uncertainty in pollution exposure or dosage estimates. These include 1) the existence and/or determination of the threshold level, 2) the slope or functional form of the dose-response relationship, 3) changes in the dose-response relationship due to aging, weather, etc., and 4) the handling of previous exposure. Most investigators have either ignored or severely reduced the dimensionality of pollutant representation. Many epidemiological and risk-cost-benefit investigations use the yearly mean pollutant concentration as a measure of pollution exposure, with the assumption that each individual has experienced a similar concentration of pollution for an equal length of time. Such measures give little information about the distribution of pollutant concentrations. High concentrations that occur for short periods of time, which may result in acute health effects, and cleansing periods, times of low pollution which may give the respiratory system a chance to "recover," are of particular interest. Moreover, annual means do not give information about previous exposure or residence history which may be important if the regional patterns of pollution have changed or if the person has moved.

Pollutant concentrations may be expressed in an infinite combination of averaging times and percentile groups. For example, sulfur dioxide concentrations typically are measured each hour. The peak of the average

of 24 consecutive measures in a year would represent the highest (1st percentile) 24 hour running average in that year. Regulatory standards are based on such peak concentrations. For example, the 24 hour peak concentration is 365 micrograms/cubic meter for sulfur dioxide. The NAAQS have been shown earlier, in Table 2.1. At present, these levels are not to be exceeded more than once annually 1/. Annual averages are not necessarily indicative of shorter averaging times and vice-versa. However, because these measures have been calcified in the National Ambient Air Quality Standards (NAAQS), they are frequently used in epidemiological investigations.

Lower percentile groups may be selected at any averaging time. For example, the 14th percentile 8 hour concentration represents the running 8 hour average for which 14% of all 8 hour averages were higher in the sample year, and 84% of them lower. Such measures do not necessarily correlate well with peak measures most investigators are familiar with. It is sometimes hypothesized that a minimum averaging time, for all but extremely high exposures, is important in regards to the initiation and development of respiratory diseases. These concentrations can be endured for short periods of time with little adverse effect. It may require 24 or 72 hours of exposure for a breakdown in body defenses before permanent

1/ The justification for the NAAQS rests on an assumed distribution of pollutant levels which supposedly can be determined by knowledge of peak levels: the Standards were never intended, however, to be used as dosage measures. An indication of the assumed distribution is found in the EPA's screening process for new source applications. Multiplication factors to convert hourly peak measures into 3, 8 and 24 hour concentrations are $0.9+0.1$, $0.7+0.2$ and $0.4+0.2$, respectively [E.P.A., 1977].

damage can occur.

Long averaging times and percentile groups approaching 50%, i.e., median rather than peak exposures, tend to smooth out extreme values. Peak values often result from the effluent plume center line directly "hitting" the monitor site, a phenomenon whose frequency and severity is determined by the stochastic influence of the weather. As a result, these nonpeak concentration measures contain a smaller random element, and tend to be more stable over time.

Instead of using concentration measures, exposures may describe the amount of time that a particular threshold concentration was exceeded. The usefulness of these measures in a clinical and statistical sense is contingent upon the threshold selected and the level of pollution experienced in the given area. However, no information on just how far the threshold was exceeded is contained in such estimates. Nonetheless, the Council of Environmental Quality (CEQ, 1978, 1979) finds such measures useful in monitoring trends in the violations of the NAAQS by noting the number of days the 24 hour NAAQS was exceeded in selected cities. Additionally, CEQ selects a threshold about 25% higher than the 24 hour NAAQS to obtain information regarding more severe violations.

Exposures may be true measures of dose, that is, a product of concentration level and time. For example, a measure of sulfur dioxide in terms of micrograms-years/cubic meter may be derived by an integration of hourly concentrations $C(t)$ over any particular time. Yearly dosage measures are simply annual average concentrations. However, if the dose response function is non-linear, for example increasing, high concentrations are significantly more important than low concentrations.

Thus, the dosage measure may weight or transform the concentration measure with some function F , an approximation of the dose-response function, to produce a linear measure of dosage.

$$\frac{\sum_{t=1}^{8760} F(t)}{8760} \quad (9)$$

Another difficulty in dosage estimates relates to changes in dose-response relationships, such as seasonal variations, aging, difference between cohorts (susceptibility groups), pathological influences in the development of a disease, etc. It remains to be seen whether dosage estimates are more sensitive to such phenomenon than more approximate measures of exposure. Given all caveats, however, dosage estimates probably contain more information than "peak" measures, and should probably be used more, at least in research. Even so, their ability to associate or predict chronic progressive diseases like emphysema to air pollution is probably less than their ability to forecast more acute symptoms and diseases, for which more approximate indicators of pollution suffice, though without particular effectiveness.

The handling of exposure history is the last issue in representing pollutant exposure. Lifetime exposure estimates are generally not feasible for other than occupational settings. In general, air pollution epidemiology looks through a narrow time window to characterize ambient air pollution levels. This situation is especially unfortunate given the long induction periods and progressive nature of many of the respiratory diseases.

The exposure during the year in which health effects are investigated might be the most relevant for acute health effects. For chronic effects, it might be the previous several years of exposure. The pollution health effects association will be diluted to the extent to which inappropriate time frames for health and air quality data are examined, and to the variability of the air quality over the crucial disease initiation/induction/development period.

Several years of data may tend to reduce random fluctuations. Typically, available data is averaged. Nonaveraging weighting schemes, which might give the maximum weight to the most recent year, and increasingly "discount" previous years, may make sense in circumstances and still tend to diminish the statistical variability.

Other transformations to exposure information may be useful in epidemiological assessments, though not necessarily in risk-cost-benefit analyses. In cross-sectional studies, exposure scores based on averaged annual rankings or normalized values may indicate regions which consistently experienced the worst pollution in the investigation. Alternatively, annual pollution scores could be corrected or standardized for yearly fluctuations among all regions (especially in a study where all geographic units are relatively close and under the same influence) and then averaged or weighted as desired. Such treatments may be justified by linear dose-response functions with low thresholds and more or less additive effects.

Following sections will demonstrate just how critical the representation of exposure of ambient air pollution is in health effects research. It is likely that risk-cost benefit determinations, while not

as sensitive to such concerns as are research studies, could be seriously biased without the adequate representation of air quality.

3.6.4 Pollutant Interaction

From the health viewpoint, pollutant interaction may produce synergisms or antagonisms which potentially alter the health effect. For example, simultaneous exposure to respirable particulates and sulfur dioxide may aggravate respiratory effects beyond what either pollutant would cause alone. Or, if the two pollutants are both necessary causes, health effects will only be observed in the mutual presence of the other. The sulfuric acid "mist" or radical, which results from the transformation of sulfur oxides with water vapor and/or fine particulates seems to be an example of this phenomenon. The acid radical is considerably more irritant to the respiratory tract than sulfur dioxide or ISP alone. Of course, from an epidemiological or policy perspective, it is important to characterize exposure as a predictor of pollutant induced health effects.

Pollutant interaction is also relevant to atmospheric chemistry, where it concerns changes among pollutant species. It is also important in estimating secondary sources of pollutants in the atmosphere (especially with respirable particulates and oxidants, as well as the sulfur acid radical).

Measures of the sulfur radical are relatively scarce. Thus, it may be advantageous to estimate its concentration knowing concentrations of its constituent parts. Inaccuracies due to surrogate measures have been

previously discussed.

Synergisms and antagonisms are not the only health aspects of pollutant interaction. A more subtle issue concerns the nature and history of health effects induced by different pollutants, which need not have resulted from simultaneous exposures. As an extreme example, consider a person exposed to high levels of TSP in childhood, and to relatively low levels of sulfates in later years. Are the effects of the TSP exposure aggravated, reduced (eventually repaired in the years of low level exposure), or unaffected by the low level exposure? Most models assume additive effects for chronic damage.

In general, pollutant interaction is not well handled in toxicological, epidemiological and risk-cost benefit assessments. The large uncertainty of dose-response extrapolations to low doses is compounded for low level exposures of several pollutants. A second contributory reason is that little data is available concerning simultaneous measures of various pollutants.

Two general methods are used to represent interactive effects. The first method is entirely empirical: simply cross-products of pollutant concentrations between pollutant pairs. In several cases Lave and Seskin [1977] found that several cross-products increased the significance of the pollution coefficient. The following cross-products were found to be significant:

Annual averages were used in all cases, rather than an daily or hourly

products of pollutant level. Thus, it can not be claimed that these measures are the best predictor of synergisms -- unless perfect temporal and spatial correlation between pollutants is assumed. In other words, for any pair of pollutants, concentrations measured at any site must experience similar peaks at the same time. Ideally, cross-products should be constructed on a daily basis. However, TSP is often sampled only monthly. Many epidemiological investigations have used annual data, in which the chronological sequence of concentrations levels has been lost. Because monitor concentrations may not be representative of exposure to the population, and may be severely influenced by local sources, sequential hourly or daily correlations of peak measures among several pollutants may not be high. In any event, cross-products possibly provide some insight, although the statistical (as well as physical) meaning is unclear.

A second method of defining pollutant interaction uses an additive scaler index. The "Pollution Standards Index" (PSI) is in fairly wide use for daily reportings of pollution conditions. It transforms concentrations of the criterion pollutants to the PSI which is a nonlinear function of concentration level at 24 hour averaging times for the several pollutants. PSI definitions are shown in Table 3.1 1/. These PSI values may be added together to combine pollutant health effects. Thus, a PSI of 100 represents either the NAAQS for one pollutant, or more

1/ PSI values closely correlate to an power function of pollutant concentrations. An approximation of the PSI function is used in Section 4.4 and Appendix A

PSI	Air	24 Hour Pollutant Concentration	
Index	Quality	micrograms/cubic meter	
Value	Level	TSP	Sulfur
500	Significant Harm	1000	2620
400	Emergency	875	2100
300	Warning	625	1600
200	Alert	375	800
100	24 hour NAAQS	260	365
50	Annual NAAQS	75	80

Table 3.1. Definition of Pollution Standards Index [CEQ, 1979].

Author	Pollution related deaths	
	sulfur	TSP
Lave and Seskin (1977)	3.3	3.8
Mendolsohn and Orcutt (1978)	10.2	-0.28
Schwing and McDonald (1976)	-7	-
Lipfert (1978)	1.9	6.5
Smith (1976)	-	11
Gregor (1976)	-	26
Bozzo (1978)	3.56	-
Wilson (1980)	3.5	-

Table 3.2. Regression coefficients for several published air pollution regression models [Wilson, 1980]. Coefficients are for annual mean sulfur and TSP concentrations. These estimates are not strictly comparable due to differences in each study.

likely, relatively low concentrations of several pollutants. Problems with this sort of index include the the determination of the proper health functions, and the assumption that pollutant interaction is independent -- strictly additive -- rather than synergistic.

3.7 Uncertainty

Estimates of morbidity and mortality due to the effects of air pollution vary widely. As mentioned earlier, excess mortality estimates are easier to come by and are potentially more valid. Wilson, [1980] has collected and adjusted the air pollution coefficients from several published cross-sectional mortality studies. These are displayed in Table 3.2. The coefficients are not strictly comparable, due to the different combination of confounding and explanatory variables which were used. Even so, the variability and inconsistency of the coefficients is striking. Such observations have motivated several researchers to challenge such estimates. Viren [1980], for example, suggests that such estimates are little more than "spurious associations."

The uncertainty in these regression estimates arises from numerous influences. Mentioned earlier is the likely multiple causality of respiratory diseases, the variability of dose-response functions, effects of confounding and risk factors, difficulties of isolating and projecting relevant pollutant measures on to a population, and the various types of statistical errors possible in regression techniques.

Only the effects of variability and uncertainty of exposure estimates on the health effects analysis are described here. For epidemiological

purposes, this variability mostly results from errors in assigning relevant and valid exposures to a large and diverse population from the concentrations of several pollutants measured in relatively few places. Because pollutant related health effects are usually subtle (and thus, easily confounded), annual or statistical variability in exposure scores for different geographic units may produce large differences in the statistical strength and risk magnitude of the pollutant health effects association. This research problem is doubly compounded by the scientific uncertainty as to exactly what constitutes an appropriate dosage measure for ambient air pollution.

For risk-cost-benefit purposes, the problem of uncertainty is still more complex. It involves demographic projections of risk factors and exposure, as developed by air quality dispersion, source emission, and population models, as described in Section 2.3.

The consequence of uncertainty of air pollution exposures largely relates to the functional form of the dose-response relationship, as described in Section 3.2. Most analyses have examined only the most convenient case, that is, a linear dose-response curve with a "zero" threshold. Thus, for a given decrease or increase in pollutant exposure, or probability thereof, a corresponding change is expected in health effects. This lead Lave and Seskin [1977] to calculate that a 50% decrease in sulfur oxides and TSP emissions would lead to a 4.8 to 6.3% decrease in mortality (age-sex-race adjusted).

In a preliminary estimate of subjective probability distributions, Henrion [1981] has looked into the expected value of perfect information. As developed in Morgan [1978], Henrion used a subjective probability

distribution of health effect coefficients which is very similar to that contained in Table 3.2. Subjective probability distributions were developed for population weighted exposures surrounding a hypothetical 1000 MW(e) coal power plant (derived from a dispersion model) and for costs of pollution control. Due to the similarly broad distributions for both health effect coefficients and exposure estimates (in terms of their relative magnitude) and importance to the (health) outcome, the expected value of perfect information regarding exposure estimates is nearly half of the health effects coefficients (about 66%). Both dwarf that of the abatement costs, which are known with more certainty.

If dose-response functions are non-linear and/or with thresholds, the effects of uncertainty in exposure estimates is likely to be quite different from that demonstrated by Lave and Seskin or Henrion. The distribution of the estimates, from worst case to most likely case might have important policy implications. If, in fact, the language of the Clean Air Act, that is, "of protecting [the population] from adverse health effects with an adequate margin of safety" is to be implemented with any degree of effectiveness, the uncertainties in dose-response relationships and exposure estimates should be laid out in a reasonable fashion in the Ambient Air Quality Standards.

Chapter 4. The Chestnut Ridge Investigation

The Chestnut Ridge area epidemiological investigation is used as a case study to permit the development and testing of various air pollution characterizations. This 6400 square kilometer predominantly rural area in western Pennsylvania has been extensively studied, in part due to the four coal-fired power plants and one gassifier which comprise the Chestnut Ridge Power Complex, and the extensive network of 17 air pollution monitoring sites in the area. The mine-fed electric power plants form the largest concentration of coal-fired electricity in the U.S. [Maher, 1981] with a total peak capacity of over 4,700 MW(e). A proposed coal power plant near the Seward site would increase the generating capacity by an additional 625 MW(e). Figure 4.1 shows the study region and the location of the power plants. These plants are the dominant local pollution source although very tall stacks (up to 1000 feet) are used to aid pollutant dispersion. Table 4.2 shows the power plant capacity and plant stack height. Full load sulfur emissions exceed 19 tons/hour with the 2.4% sulfur coal typical of this area [Ruane, 1976].

1/ For a description of the power plants and details of the monitoring systems see Ruane, 1976. For amplification of pollutant representations, see Maher, 1980. "Progress Reports on the Health Effects of Air Pollution in the Chestnut Ridge Area" are available in Gruhl, 1980 and Gruhl, 1981. For preliminary results of the children's epidemiological study in this area, see Schenker, 1981.

Figure 4.1. Chestnut Ridge study region, located around Indiana County in Western Pennsylvania. Sites of the four coal power plants are indicated.

4.1 Health Study

Five separate health surveys have been carried out in the Chestnut Ridge area and are briefly described below.

1. The adult women study (September 1978 to January 1979) consists of a telephone survey of 5686 women, age 17 to 74. These women were elected at random. A standardized respiratory questionnaire was used in this cross-sectional investigation.
2. The child school survey (February to May 1979) is a second cross-sectional study using a respiratory disease questionnaire completed by parents of 4071 children. These children attended grades 1 to 6 at one of the 14 regional schools in the area. Children also had weight and spirometry (lung flow) tested in school. The 1906 children in grades 4 through 6 also completed confidential smoking questionnaires.
3. The acute effects prospective study in selected women (September to December 1979) is a case-control study involving 119 women selected from the adult women study. Included were 24 women with chronic cough or asthma, and 43 women with chronic phlegm or bronchitis. Controls for these two groups numbered 13 and 29 respectively, and were matched for residence location, age, and smoking habits. Lung function was tested every two weeks and questionnaires were completed at the beginning and end of the study.
4. The childrens acute air pollution study (November to December 1979) is a longitudinal investigation of 120 children in several high pollution areas. Children were tested for lung function before and after an air pollution alert. Questionnaires were also given to children.

5. The prospective surveillance study of selected children (September 1980 to May 1981) is a case-control study of children selected from the child school survey. Three groups of children were selected: 1) those with chronic cough or phlegm; 2) those with persistent wheeze without cough or phlegm, and 3) children without chronic respiratory symptoms, to serve as controls. Parents maintained a daily calendar of respiratory symptoms, and children had daily lung flow measurements.

Details of these studies may be found in the previously cited documents.

For the most part, the following sections discuss the child school survey. This data was used in the evaluation of various pollutant representations. In addition, the emphasis of this research concerns air pollution characterizations which are useful for cross-sectional epidemiological or risk-cost-benefit investigations, rather than longitudinal studies.

4.1.1 Confounding Factors

Questionnaires for the child school study concerned various respiratory health effects. These were then classified into five categories of respiratory symptoms, including 1) persistent cough, 2) persistent phlegm, 3) persistent wheeze, 4) chronic respiratory illness within the last year, and 5) chronic respiratory illness before age 2. The prevalence of these symptoms at each of the schools is shown in Table 4.2.

Questionnaires also attempted to isolate potential confounding and risk factors. These factors are listed below.

1. Socio-economic status (SES) derived from the parents occupation and education.
2. History of parent's allergic and respiratory disease history, based on M.D. diagnosis of bronchitis, emphysema, asthma, or hay fever.
3. Type of cooking stove, either coal, wood, gas, electric, etc.
4. Passive smoke exposure from parents and other household members.
5. Smoking history of the child.
6. Sex of the parent completing the survey.

Schenker [1981] discusses the relationships of these factors to the respiratory symptoms and illnesses or outcomes found in the children. In brief, SES is inversely related to respiratory outcomes and gas cooking stove use. Controlling SES reduced the significance of the association between gas stove use and persistent wheeze. In addition, the parents' respiratory disease history was a significant independent risk factor for all respiratory diseases. Passive smoke exposure was demonstrated to be a significant risk factor for chest illness both before age 2 and in the past year. Additionally, parental smoking is associated with SES. The last significant confounding element isolated was an apparent "reporting bias" by the parents who completed the questionnaires. Parents who report respiratory symptoms in themselves seem more likely to find them in their children. However, this association may be due to similar genetic factors or a common exposure to air pollution by both parents and children.

4.1.2 Statistical Analysis

Initial analyses of data were done by simple frequency distributions

and cross-tabulations. A multiple logistic regression model was then used for the analysis of combined health and air pollution data in a step-wise fashion. This method permits the simultaneous adjustment of qualitatively different risk factors, i.e., stove type and SES class, etc. Thus, it provides an effective method to control confounding. For each child, the logit model estimates the dichotomous dependent variable p , the health e.g., whether or not s/he has a chronic cough, illness, etc., as a function of several independent variables:

$$\log\left(\frac{p}{1-p}\right) = a + B_1X_1 + B_2X_2 + \dots + B_nX_n \quad (10)$$

where X_i = independent variable, either dichotomous, categorical or continuous, e.g., SES, pollution, etc.

$a_i, B_i \dots B_n$ = maximum likelihood estimates

Logistic models have disadvantages of large CPU core and time requirements. Thus, a simpler method was used for the preliminary screening of pollution indices, as described in the next section.

4.1.3 Screening Model

A simple correlation between prevalences at each of the schools (as displayed in Table 4.1) and pollution scores for the schools was used to efficiently test the many pollution indices against the health data. In addition to the prevalence, the logit transformation of prevalence, L , was also used in the correlation model, to roughly portray the behavior of the regression model.

$$L = \log\left(\frac{p}{1-p}\right) \quad (11)$$

This very simple screening model does not incorporate any of the known confounding or risk factors. For example, the model is likely to provide high estimates of the association between health effects and pollution levels. This is because SES, passive smoke exposure and gas stove use, all which are known to be associated with a high prevalence of respiratory symptoms, are also positively associated with the pollution variable. When controlled in the multiple regression, these independent variables should tend to weaken the association between the health and pollutant variables. However, this screening model may not be greatly affected by age and sex distribution found in the schools, since most schools have a similar distribution of these factors. This is important since these variables are significant risk factors for most respiratory symptoms ($p < 0.05$). Two schools, however, had only 1st and 2nd graders. One of these, number 13, had only 80 students and is located in a high pollution area. Correlations may be unduly influenced by this extreme observation. Because each school is a regional school, drawing students from several townships, additional confounding factors might tend to be randomized. However, it is obvious that SES, passive smoke exposure, and even age and sex are not ideally controlled in this screening model. Screening model results must be interpreted carefully: in all cases, they are inferior to the regression model.

A second screening model, a correlation which was weighted according to the number of students at each school, was also used. This may tend to increase the power of the screening model, since the influence of

<u>Unit</u>	<u>Capacity</u>	<u>Stack Height</u>
Keystone 1 & 2	1640 MW(e)	797 feet
Homer City 1 & 2	1200	796
Conemaugh 1 & 2	1700	1000
Seward 3, 4, & 5	218	225*

* Stack height raised to 1000 feet in September 1976. Unit shut down from March to September 1976.

Table 4.1 Power plant capacity and stack height in the Chestnut Ridge power complex.

	SCHOOL	SYMPTOM PREVALENCE - PERCENT					
	POP	COUGH	PHLEGM	WHEEZE	CI- 3dy	CI-1 yr	
O1	BEN FRA	143.	9.1	4.3	4.7	9.3	4.4
O2	UNITED.	818.	7.6	5.1	8.2	9.1	8.9
O3	BURREL.	277.	8.6	4.8	8.8	9.	9.4
O4	EASTPIK	354.	5.	2.3	6.3	9.4	7.8
O5	FAIRFLD	102.	2.1	3.	9.8	7.8	10.
O6	BLAIRSV	101.	9.3	5.1	8.1	12.9	5.
O7	CLYMER.	175.	10.	4.7	8.1	15.	12.3
O8	PENNRUN	159.	4.5	1.9	7.	11.3	7.7
O9	EISENHR	240.	3.5	1.7	7.2	12.2	6.5
10	HORMANN	344.	7.7	4.2	8.	11.	5.8
11	HOMER C	655.	6.3	4.3	8.3	12.9	6.5
12	LAURELV	333.	6.4	4.9	4.2	10.2	9.8
13	NEWFLOR	80.	10.3	6.4	5.1	25.	15.
14	3RDWARD	281.	6.4	4.5	7.5	10.4	5.4
	MAXIMUM...	818.	10.3	6.4	9.8	25.	15.
	MINIMUM...	80.	2.1	1.7	4.2	7.8	4.4
	MEAN.....	290.1	6.9	4.1	7.2	11.8	8.2
	STAN. DEV..	213.2	2.5	1.4	1.6	4.3	3.

Table 4.2. Number of children and prevalence rates for five respiratory symptoms at the fourteen schools in the Chestnut Ridge child school survey.

confounding factors may decrease among a larger sample size. The number of students at each school is shown in Table 4.2. In addition, this model eliminated the New Florence School, which contained only 1st and 2nd graders, as this was not representative of the other schools in the sample.

4.2 Pollutant Representation

Over four years of hourly sulfur dioxide concentrations and over five years of daily to weekly total suspended particulate (TSP) concentrations at each of the 17 monitors have been condensed into a data format known as the "arrowhead profile." The arrowhead profile or curve is a concise way of expressing pollutant concentrations and frequency of occurrence at various averaging times 1/. Suggested by Larsen [1981] and implemented by Maher [1981] for the Chestnut Ridge area, this format allows the convenient construction of various types of pollutant dosage measures 2/.

The arrowhead curve is a time collapsed representation of annual concentration statistics: the chronological sequence of the pollutant

1/ The data storage requirements for sulfur measures have been reduced by a factor of about 110 for sulfur and by about 8 for TSP. These factors are a function of the resolution desired in the pollutant representation, as described later.

2/ Larsen [1971] proposed a model of representing air pollution which could be constructed by using just two observations assuming 1) average concentrations are approximately log-normally distributed and 2) the median of the ranked concentration is proportional to the averaging time raised to a constant. However, Maher's construction of arrowhead curves refutes that these two assumptions are simultaneously valid.

concentrations has been lost in favor of a probabilistic strategy where isoquantiles describe the frequency of occurrences. Figure 4.2 displays a typical arrowhead curve for TSP concentrations at monitor 1 for 1978. The log-log plot has sulfur dioxide concentration in micrograms/cubic meter on the vertical axis and averaging times in hours on the horizontal axis. At each averaging time there are 6 points representing various isoquantiles or percentiles. The 100th percentile represents the highest peak value which occurred at that averaging time. The 1st percentile represents the lowest concentration which occurred at that averaging time. Both of these values occurred once. Figure 4.2 also displays the 90th, 84th, 50th and 16th percentile curves. The 84th percentile curve represents the concentration at various averaging times for which 84% (i.e., 306 out of 365 daily averages) were lower than the particular concentration and for which 16% (the remaining 60 averages) were higher. The other percentile curves may be interpreted in a similar fashion. The 50th percentile line represents the median concentration which remains fairly constant at various averaging times.

For sulfur, 99th and 1st percentiles were used, which represent the second highest and the second lowest concentrations, respectively. The 100th and 0th percentiles were used for TSP because these measures are taken less frequently. From this point, only the 99th and 1st percentiles are used when discussing either pollutant.

If the pollution concentrations are log-normally distributed, the 84th and 16th percentiles are equivalent to one standard deviation above and below the mean, respectively. At the appropriate averaging time, the 99th percentile may be interpreted directly as the NAAQS. For example,

Figure 4.2. Arrowhead profile for TSP concentrations at monitor 1 in 1978.

the 24 hour averaging time 99th percentile concentration in Figure 4.1 is 2.1, or about 125 micrograms/cubic meter (u.gm/cu.m), which is well below the primary standard of 260 micrograms/cubic meter ($\log(260) = 2.42$). Table 4.2 summarizes percentile information used in the the Chestnut Ridge study. Table 4.3 shows the various averaging times used in the construction of the arrowhead curves. Sulfur measures are available for shorter averaging times, because measures are taken every hour.

All points on the arrowhead curves are derived from running hourly averages in the case of sulfur, and running daily averages in the case of TSP. Thus, measures for the 1978 profile in Figure 4.2 are not necessarily confined to the calendar year 1978. The sampling period for any calendar year includes portions of both the preceding and following years. Intrusion into adjacent years is by an amount exactly equal to half the averaging time. This permits 365 averages to be used at each averaging time for the construction of the TSP arrowhead curve and 8760 averages for sulfur. Thus, in the case of the annual averaging time for Figure 4.2, the 99th percentile is the second highest running average of any 365 consecutive measures of which at least one occurred in 1977 1/. Percentile groups may still be constructed for the annual averages. Running averages are used instead of "block" averaging because the physiological response is a continuous phenomenon. Health effects should not be contingent upon an artificial structure of fixed or arbitrary

1/ Interpolation and other corrective measures were necessary to supplement missing data. If more than 1/3 the observations at any averaging interval were missing, no estimate was created. In general, monitor availability was good, generally exceeding 80%.

<u>TSP</u>		<u>Sulfur</u>
<u>Hours</u>	<u>Equivalent Period</u>	<u>Hours</u>
		1
		2
		3
		8
24	1 day	24
	2 days	48
	3 days	72
	1 week	168
336	2 weeks	336
730	1 month	730
1460	2 months	
2190	3 months	2190
4380	6 months	4380
8760	1 year	8760

Table 4.2. Averaging times used for the arrowhead profiles for TSP and sulfur dioxide.

<u>Percentile</u>	<u>Equivalent</u>
100th	highest (TSP only)
99th	second highest (sulfur only)
90th	+1.28 standard deviation
84th	+1 standard deviation
50th	median value
16th	-1 standard deviation
1st	2nd lowest value (sulfur only)
0	lowest value (TSP only)

Figure 4.3. Quantiles and equivalents displayed in arrowhead curves.

sampling periods, as represented by block averaging.

Advantages to operating in the arrowhead format include computational speed, ease of data handling and transformation, greater accuracy of spatial interpolation [Gruhl, 1980], as well as a visual interpolation of pollutant patterns. The principal disadvantage is the loss of chronological information. For example, it is impossible to determine exactly when the peak 24 hour average occurred [Maher, 1980]. Arrowhead curves may be constructed on a daily or weekly basis to provide more continuous measures needed for longitudinal (over time) epidemiological investigations 1/. The use of the arrowhead curve is described at length in following sections.

Arrowhead profiles have been constructed on an annual basis for all 17 monitors in the Chestnut Ridge area. Profiles for sulfur dioxide are available for four years (1975 to 1978) and for five years for TSP (1974 to 1978). These 153 profiles (4x17 + 5x17) provide the primary data base for the air pollution component of the research. More detailed arrowhead modeling using eleven percentiles (1, 10, 20, 30, . . . 90, 99) were used for validation purposes in certain cases. A number of pollutant measures were constructed from these arrowhead curves, as described in the next section.

1/ Arrowhead curves with 6 percentile groups obviously require at least 6 observations at each averaging time. For better resolution and robustness (in the case of missing data) several dozen observations are desirable. Thus, to construct daily arrowhead curves on a non-overlapping or block basis (which may be more useful for a longitudinal study where data smoothing is undesirable) using hourly sulfur measures, the longest possible averaging time is 8 hours, with the arrowhead converging to a single point for the 24 hour average.

4.3 Pollutant Indices

Some 70 different indices to represent pollutant exposure or dosage were developed for TSP and sulfur concentrations at each of the 17 monitors. In addition, to test various hypotheses concerning health effects, pollutant characterization and interaction, selected measures were manipulated and combined into yet more indices. These actions included ranking, normalization, standardization, logarithmic transformations, cross-products and cross-sums. Eleven indices for both TSP and sulfur which had been previously developed are shown in Table 4.4, and are discussed in the following sections.

4.3.1 NAAQS Measures

The arrowhead profiles display second highest peak values at each averaging time. At the appropriate averaging time, these 99th percentile values may be directly interpreted as the NAAQS. Thus, indices 4 and 5 for sulfur represent the fraction of the 3 and 24 hour NAAQS. For TSP, index 4 may be interpreted as the 24 hour standard. Index 5 for TSP represents the fraction of a hypothetical standard of 147 micrograms per cubic meter with a 15 day averaging time. This "standard" was derived by a logarithmic interpolation of the 24 hour and annual threshold standards.

Index 6 for each pollutant represents the fraction of the annual NAAQS standard for the two pollutants. This measure is based on a running annual average of the second highest peak level. Index 11 also represents an annual average, but on a calendar year basis. This latter index

<u>Index</u>	<u>Name</u>	<u>Description of Sulfur Index</u>	<u>Description of TSP Index</u>
1	General	Sum of 99, 84, 50, 16 and 1 percentiles for 1, 3, 8, 24, 72, 168, 730, 2190, and 8760 hours averaging times	Sum of 100, 84, 50, 16, and 0 percentiles for 1, 72, 168, 730, 1460, 2190, and 8760 hours averaging times
2	Short-term Acute	Sum of 99 and 84 for 1, 3, 8 and 24 hours	Sum of 100 and 84 for 24 hours
3	Short-term High	Sum of 99, 84 and 50 for 1, 3, 8 and 24 hours	Sum of 100, 84 and 50 for 24 hours
4	Short-term Standard	Ratio of 99 percentile to to 3 hour threshold	Ratio of 100 percentile to 24 hour threshold
5	Mid-term Standard	Ratio of 99 percentile to 24 hour threshold.	Ratio of 99 percentile to hypothetical 15 day standard of 147 ug/cu.m.
6	Annual Running Standard	Ratio of 99 percentile to annual threshold	Ratio of 100 percentile to annual threshold
7	Short-term Cleansing	Sum of 16 and 0 for 1, 3, 8 and 24 hours	Sum of 16 and 0 for 24 hours
8	Long-term Cleansing	Sum of 16 and 0 for 72, 168, 730 and 2190 hours	Sum of 16 and 0 for 168, 336, 730, 1460, 2190, 4380, and 8760 hours
9	Short-term Moderated Acute	Sum of 90 and 84 for 1, 3, 8, and 24 hours	Sum of 90 and 84 for 24 hours
10	Short-term Moderated High	Sum of 90, 84, and 50 for 1, 3, 8, and 24 hours	Sum of 90, 84 and 50 for 24 hours
11	Annual Calendar Standard	Ratio of calendar year average to 1 yr threshold	Ratio of calendar year average to 1 yr threshold

Table 4.4 Pollution indices for sulfur and TSP [Gruhl, 1981].

represents the actual annual NAAQS, which is confined to the calendar year. Index 6 is generally very similar to index 11, and is used throughout the text to represent annual averages and the NAAQS.

Index 6 should better characterize pollutant exposure for purposes of health assessments in any particular year than index 11. This is due to the arbitrariness of calendar years with respect to health effects. However, index 11 may be a more useful measure in examining an area's pollution "history." Figure 4.3 shows how calendar year averages may mislead exposure estimates. Calendar year estimates for each of the three years portrayed are exactly the same. However, it is clear that from mid-1976 to mid-1977 pollution levels were far above the norm. A running annual year average would capture this high exposure: running averages for both 1976 and 1977 would reflect this high exposure period.

4.3.2 90th Percentile Indices

In addition to the previously described peak measures, the 90th percentile concentrations for various averaging times were examined. These measures should characterize the higher range of pollutant concentrations and be less severely affected by the stochastic influence of the weather than peak measures, i.e., the 99th and 100th percentile measures which comprise the NAAQS.

4.3.3 Arrowhead Indices

The arrowhead profile permits a general characterization of pollution

patterns, as well as particular concentration, peak measures, etc., just described. For example, regions above the 50th percentile depict "acute" or higher than average pollution periods. Regions on the profile below the 50th percentile reflect lower than average concentrations, or cleansing periods. Figure 4.4 shows several arrowhead profiles. The differences in the shapes of the profiles, in terms of sharpness, asymmetry, spread, etc., is quite dramatic.

To characterize these differences, six indices for each of the pollutants have been developed for the study. These indices are labeled 2, 3, 7, 8, 9, and 10 and are explained in Table 4.4. Rather than "point" measures, like the peak and NAAQS measures described previously, these indices combine concentrations at selected percentiles and averaging times, depending on the effect one is trying to characterize. The summing of concentrations may tend to reduce the influence of extreme values. These combined measures have a complex mathematical interpretation which does not easily lend itself to a simple expression of concentration or time. However, they provide relative indications of cleansing and acute periods that might be useful in epidemiological investigations.

Index 1 is entitled the "general index" because it sums concentrations over five percentile groups and many averaging times. This index may help isolate extreme differences between monitors when peak or average values are similar.

4.3.4 Exceedence Measures

"Exceedence measures" were developed to estimate the amount of time

Figure 4.3. Hypothetical pollution concentrations over three years demonstrating difference between running and calendar year averages. Average annual concentrations in the three calendar years are the same.

1974 Arrowhead Curve

1976 Arrowhead Curve

1975 Arrowhead Curve

1977 Arrowhead Curve

Figure 4.4. Arrowhead curves at monitor 1 for TSP for four years.

that a particular threshold was exceeded. Three thresholds were selected for sulfur: 30, 80 and 150 parts per billion (ppb), and two thresholds for TSP: 75 and 150 micrograms/cu.meter. For both pollutants, the lowest threshold represents the annual NAAQS. Thus, these exceedence measures might display the percentage of time in a year that the annual NAAQS was exceeded. An exceedence time of 100% means that the NAAQS was violated.

These types of measures may be useful with a threshold-type dose-response function. They may also serve to isolate regions which have frequent exposures to relatively high concentrations of air pollution, but fairly typical mean and peak values. Exceedence measures have the advantage of reducing the dimensionality of percentiles since the threshold is fixed. However, exceedence measures may be constructed at any averaging time. Thus, to test the sensitivity of these measures to averaging time, they were constructed at all averaging times found on the arrowhead profile. Sulfur exceedences above 30 and 80 ppb were found for 13 averaging times at all monitors, and at 5 averaging times (up to 24 hours) for exceedences above 150 ppm. Only 5 averaging times were used for this threshold because very few exceedences were found with averaging times longer than 24 hours. (Thus, exceedences equaled 0 for all monitors.) TSP exceedences were found for both thresholds at 7 averaging times. Later sections describe the consistency found in the 45 indices developed.

Exceedence measures have the disadvantage of providing little information on the distribution of pollutant concentrations above or below the threshold. For example, two regions may exceed a particular threshold a similar amount of time, but one region might experience much

higher short-term concentrations. For this reason, several thresholds were used. Ideally, clinical hypotheses concerning respiratory effects should determine which thresholds are used. However the range of useful thresholds is restricted by ambient concentrations found in the area. A threshold which is too low will consistently give exceedences of 100%; one too high will give 0%. Neither is helpful for cross-sectional analyses or geographic comparisons.

The arrowhead curve was used to estimate exceedence times. Profiles for the case study displayed concentrations at 6 percentiles. Consequently it was necessary to extrapolate between percentile groups at particular averaging times to estimate the quantile that the threshold value represents. A linear interpolation between log concentrations was used in all cases. Thus, this procedure assumes a lognormal distribution of concentrations between percentiles. Because only six percentile groups were available, some error is possible if the lognormal assumption is seriously violated. However, these estimates were verified with extrapolations using arrowheads containing eleven percentiles.

Extrapolations between two points are quick and efficient. However, testing is required to find the appropriate points. One must lie above the threshold; the other below. If both are above the threshold, e.g., the 99th and 90th percentiles, then the next pair, e.g., the 90th and the 84th percentiles, must be examined. If all six points (i.e., the 99, 90, 84, 50, 14 and 0 percentiles) are below the threshold, a value of 0 is assigned. If all are above, a value of 100% is assigned. This procedure is inefficient. A faster way to accomplish the same function is to construct an arrowhead curve where isopleths represent threshold

concentrations (rather than percentiles of time) on a time vs. averaging time basis (rather than concentration vs. averaging time). This would permit exceedences to be directly read off from the profile at the appropriate isopleth. Appendix A discusses an alternative arrowhead model which incorporates this approach.

4.3.5 Non-Linear Dosage Measures

As mentioned earlier, a simple annual concentration represents a true dosage measure in the sense that it is a product of concentration and time. In order to develop other dosage measures which incorporate a non-linear or threshold dose-response function, more sophisticated treatments are necessary. Note here that we are not concerned with the structure of the health model, but rather with the development of a input variable. For example, a linear population health model may be used with a dosage measure which weights exposures to high pollution more than linearly than exposures to lower concentrations. This section describes the construction of such a non-linear dosage measure derived from the Pollution Standards Index.

As seen earlier in Section 3.5.3, a non-linear dose d may be calculated as

$$d = \frac{\sum_{t=1}^{365} H(c)}{365} \quad (10)$$

where $H(c)$ is the health risk function using daily measures of pollution concentration c .

This analysis uses a health function H derived from the Pollution Standards (PSI) scale outlined in Section 3.6.4. PSI scores are an approximately linear scale of the health risk posed by 24 hour averages of pollutant concentrations. (Of course, any averaging time may be used in such a measure.) This dosage measure is analogous to a yearly summation of daily PSI scores.

Having assumed that the PSI is a linear indicator of potential health damage, our next task is to build a model giving a continuum of PSI scores at any concentration c. Using a power function, the optimal least squares lines which fit the PSI definitions are shown below.

$$PSI_{tsp} = 0.90066 c^{0.90019} \quad (.984) \quad (11)$$

$$PSI_{sul} = 2.44587 c^{0.66121} \quad (.991) \quad (12)$$

The high correlation coefficients (in parenthesis) indicate that the equation structure is satisfactory. Since pollutant concentrations in the Chestnut Ridge region rarely exceeded the PSI equivalent of 100, the regression lines were re-optimized using points at or below this level:

$$PSI_{tsp} = 4.5032 c^{0.55755} \quad (13)$$

$$PSI_{sul} = 6.7594 c^{0.45666} \quad (14)$$

These functions closely model the PSI relationships for the pollutant levels found in the study region.

The equivalent of a piece-wise linear approximation, using the six

percentile points on the arrowhead curve, was used to estimate cumulative PSI scores. The basic procedure is to estimate the average concentration found in each of the six percentile periods in a year, i.e., the average concentration in the "dirtiest" 7 days, the next dirtiest 30 days, etc. These concentrations are then used in the appropriate PSI equation. Finally, the resulting six PSI scores are weighted by the amount of time that each represents, i.e., 7, 30, etc., days, and then summed. The result is an annual PSI index which fairly accurately reflects daily PSI scores. Details on the construction and verification of this scheme are shown in Appendix B. In effect, this dosage measure gives higher weight to lower percentiles. Figure 4.5 show the dose-response relationship implied by the PSI dosage measures.

4.3.6 Threshold Dosage Measures

Threshold dosage measures assume that exposures to pollutant concentrations below a certain threshold do not cause health effects. Thus, lower concentrations for whatever period of time are irrelevant to health effects (with this hypothesis) and must be "subtracted out." The measures developed in this section assume two dose-response functions, with a threshold level pinned at the annual NAAQS, as shown in Figure 4.5.

Threshold dosage measures were estimated using the exceedence measures described in Section 4.2.3. In brief, both normal and log-normal distributions of pollutant concentrations were assumed. Exceedence measure thresholds were multiplied by either a normal or lognormal adjustment factor to estimate the average concentration in each interval,

Figure 4.5 Dose-response functions implied for the dosage measures based on the PSI, and two threshold measures. Annual average concentration, which implies linear dose responses, is also shown.

i.e., the average concentration between thresholds of 30 and 80 ppm. Then, the concentration values for each interval were multiplied by the exceedence measure (the amount of time spent in that interval). Lastly, all such products were summed. The result is an estimate of concentration-year above the annual NAAQS.

For threshold concentration d , exposures may be constructed from hourly measures:

$$\text{For all } C(t) > d \quad \sum_{t=1}^{8760} (C(t) - d)/8760 \quad (15)$$

This would provide a "true" measure of dosage above a threshold. Probabilistic extrapolations may be more convenient, where the time variable is disaggregated into n fixed components or percentiles, where $T(n)$ is the percentile fraction of the year, $F(c)$ is the health function (equal to unity with linear dose-response relationships) and $C(n)$ is the average concentration in each percentile for each percentile:

$$\sum_{n=1}^n C(n) F(c) T(n)/n \quad (16)$$

This method was used to construct the non-linear dosage measure (based on the PSI). Pollutant levels may also be treated as the fixed variable, which is the method used for the threshold dosage measures. For concentration thresholds p_1 and p_2 , where p_1 is greater than p_2 and (as typical) time duration exceeding p_1 , i.e., $T(c > p_1)$, overlaps the time duration exceeding p_2 , with weights $w(p_1)$ tailored to distribution of

pollutant values and the dose-response function, dosage may be calculated as follows:

$$D = p_1 W(p_1) T(c > p_1) + p_2 W(p_2) (T(c > p_2) - T(c > p_1)) \quad (17)$$

Equation 17 is particularly easy to implement in this manner when the lowest exceedence threshold (p_2 in this case) is set at the health effects threshold and higher thresholds (e.g., p_1) are set at reasonably higher levels. The procedure is described in detail in Appendix C.

The threshold dosage with a log-normal distribution would tend to reflect an increasing (more than linear) dose-response relationship. The normally distributed threshold dose is closer to a linear dose-response function -- above the threshold. Neither threshold dosage measures includes low percentiles. Thus, both model dose-response functions that are only sensitive to high level "acute" exposures. In this sense they are similar to the exceedence measures. However, the dosage measures effectively combine exceedence measures at different thresholds. As mentioned earlier, the average annual concentration provides a linear dosage measure. Lastly, the dosage measure based in the PSI gives more emphasis to low percentiles.

The goal of these dosage measures is a linearized measure of exposure with respect to health effects. Logarithmic, other transformations, and/or selective use of various averaging times may help compensate for the log-normal distribution of pollutant concentrations. However, the physical significance (though perhaps not the clinical relevance) of such measures becomes murky with such actions.

4.4 Regional Characterization

The Chestnut Ridge area experiences substantial variation in pollutant concentration with small changes in distance. This is due to the complex terrain and the importance of the local sources. Thus, to accurately estimate the population exposure, 36 small geographic units, either townships or portions thereof, were defined in the study region. These "districts" were used to estimate exposure for the adult women study. Planer estimates, corrected for terrain and distance, were made for the population weighted center of each district from the three closest monitors. The procedure is shown in Figure 4.6, and further explained in Gruhl, 1981. The same triangulation procedure was used to derive pollution scores at the exact location of each of the schools. Monitor, district and school locations are shown in Figure 4.7.

Both the childrens' school and township of residence were known. Pollution scores for the 21 townships in the study region were found by averaging the district scores which comprised the township. Thus, to estimate the child's exposure E , any combination of school or township pollution score, respectively $C(s)$ and $C(p)$, may be selected with weight w :

$$E = w C(s)/2 + (1-w) C(p)/2 \quad (18)$$

School and residence scores were used separately in the regression model. The screening model used the school scores alone, and a combined township and school score. The latter averaged the township score in which the

Figure 4.6. Block diagram of derivation of indices from monitor scores.

Figure 4.7. Location of 17 monitors, 37 districts and 14 schools in the Chestnut Ridge area.

school was located with the schools score, that is, $w = 0.5$. This was the only way that residence scores could be incorporated. Either township or school score could be used in the regression model because each child's score could be computed individually from his or her residence location and school location. This was not possible in the screening model. The use of both scores was desirable since school scores did not have a range of pollution as great as the townships. For statistical purposes and because of the subtlety of pollution related health effects, it is advantageous to have a range as large as possible.

4.5 Evaluation of Indices

The ideal pollutant representation should meet several criteria. First, it should be clinically relevant to the health effect. Second, it should be relatively stable over time; the stochastic influence of weather and the particular monitoring site selected should not greatly affect the measure. Thirdly, it should be able to clearly distinguish "clean" from "dirty" areas. Secondary criteria relate to the data and computational requirements necessary to derive the measure, and the ease of physical interpretation in using the measure.

Health models used in this investigation have been described previously. The great danger in using these models to prove the clinical relevance of a pollution measure was noted at the onset of this paper. Far more research will be necessary to prove that the results of this study which relate to the clinical relevance of pollution measures are, in fact, transferable to other situations.

The coefficient of variation (COV) was examined to assess the stability of pollutant measures. The COV is the standard deviation of the (available) 4 or 5 annual pollution scores at a monitor, divided by the mean of the scores for the same period of time. COVs were found for many of the indices at each of the 17 monitors. The average of the 17 monitor COVs was also computed. Background concentrations of pollution, attributable to long range transport, should remain fairly constant and not seriously affect the COVs. In the absence of changes in area source emissions, low monitor COVs should indicate a measure which is stable over time. A low mean COV should indicate an index which tends to be stable. Thus, these COVs provide a measure of stability across monitors, as well as indices.

A pollutant measure must have a workable range of values to be useful in either health effects or policy assessments. This range should be significantly larger than the random variation found in the measure, at least for the majority of the points. Typically, a range is determined by the extreme outlayers, the maximum and the minimum values. For epidemiological purposes, however, the range is better established by the annual standard deviation of the monitor sites as compared to the mean of the monitor COVs. Otherwise, outlayers may represent a statistically insignificant population. For example, it is possible that only one relatively unpopulated geographic unit may have pollutant levels significantly different from the rest of the population. This would weaken the association between pollution and health effects.

The following chapter describes the results of the various pollutant characterizations using health and air quality data from the Chestnut Ridge area.

Chapter 5. Case Analysis

This chapter presents the analysis of various pollutant characterizations using data from the Chestnut Ridge health effects study. First, the patterns of pollution found in the region are examined with respect to known sources. Then, the stability of the various pollutant indices is explored. More stable or interesting indices undergo further analysis to investigate the effect of averaging times, percentiles, chronological effects, etc. A simple screening model is used to estimate the relevance of pollutant measures to health effects. Selected measures are then used in a multiple logistic regression model, which controls confounding and risk factors. Lastly, some of the potential health effects of a proposed power plant in the Chestnut Ridge area are investigated, using results from the regression model. Figure 5.1 shows a block diagram of the steps involved in this analysis.

5.1 Regional Evaluation

The rural Chestnut Ridge area is quite heavily affected by local and long range pollution sources. The four coal power plants, gassifier, and large hospital near Blairsville (in district 22) are the major stationary sources of energy-related pollutants. The primary source of the pollutants imported to the area is the Pittsburgh urban area, which is located some 50 to 70 km to the west. Prevailing winds carry Pittsburgh's (as well as the Ohio River Valley's) emissions directly over the region. Winds originate from the SW to the WNW direction some 53% of the time,

Figure 5.1. Block diagram of analysis of pollutant measures and health model in this chapter.

with an average speed from this sector of about 10 mph (as measured at Pennview meteorological tower). Winds blow from other directions with approximately equal frequencies, but with a lower mean speed of about 4 to 8 mph [U.S. Army Corps of Engineers, 1981]. The Johnstown urban area lies just 20 to 30 km to the southeast, but has a negligible effect on the Chestnut Ridge region. This is due to both the prevailing westerly winds, and the Laurel Hills, a chain of 2000 to 2500 foot high mountains which separate the two areas.

In general, the region experiences a pollution gradient which increases to the south and east. However, the patterns are rather more complex, due to several localized "hotspots." Peak concentrations, especially, may vary tremendously with only small changes in distance. This is due largely to the complex terrain. West facing slopes and locations on the top of the ridge tend to be severely impacted. East facing slopes are not as greatly affected by such short duration, high level exposures.

Tables 5.2 and 5.3 display annual concentrations (index 6) for the 17 monitors in the region. As shown by the statistics in these two tables, the region often violates the annual NAAQS for both pollutants. (Even so, the region presently is classified as an "attainment area" by the Commonwealth of Pennsylvania.) Over the period from 1974-5 to 1979, both TSP and sulfur concentrations have tended to decrease. Mean sulfur values tended to increase by 53% from 1975 to 1977, as seen in Table 5.1. For the same period of time, TSP levels increased only 6%. (Calendar year changes are less dramatic, 21% for sulfur and 2% for TSP. TSP concentrations also show a smaller range than the sulfur values. Such

long term macro level changes suggest that TSP readings at the monitors are not related to sulfur emission sources. One TSP source is dust stirred up by traffic on dirt and gravel roads, which are adjacent to monitoring sites. Six sites were relocated in 1975 due to this problem. However, discrepancies between sulfur and TSP concentrations are widespread and still suggest that different emission sources are involved for the two pollutants. This is corroborated by other evidence, as shown in Section 5.4.

Figures 5.2 to 5.5 map annual calendar average (index 11) and 3 hour average peak (index 4) concentrations on to the approximate locations of the 17 monitors for the years 1975 to 1978. Isopleths, or lines of equal pollutant concentration, have been drawn for the annual concentrations. As seen in any year, the area surrounding monitor 17 consistently experiences high concentrations, due to the presence of the Connemaugh and Seward power plants. Short-term concentrations at this monitor far exceeded -- by 250 to 300% -- levels at other monitors in 1975 and 1976. This is probably due to downwash from the short 230 foot stack at the Seward Plant. After 1976, when the stack was raised to 1000 feet, concentrations at this monitor were only slightly above the mean. At this time, monitors 14 and 16, located nearby on a ridge, experienced dramatic increases in annual concentrations. This seems due to the fumigation of these monitors by direct impaction of the now higher Seward plume. This southeastern area is consistently "dirty," in terms of annual concentrations. Dispersion of power plant emissions is apparently impeded by the location of the plants in a valley, as seen earlier in Figure 4.5.

Figures 5.6 to 5.9 show calendar year averages (index 11) and 24 hour

	POLLUTANT LEVEL IN YEAR				4 YEAR
	1975	1976	1977	1978	MEAN
MONITOR 01	.39473	.63667	1.01867	1.04413	.77355
MONITOR 02	.7258	.89133	.85313	.71307	.79583
MONITOR 03	.5348	.47113	.8404	.8786	.68123
MONITOR 04	.59847	.86587	1.146	.82767	.8595
MONITOR 05	.9168	.78947	1.00593	1.00593	.92953
MONITOR 06	.8786	1.04413	1.146	.955	1.00593
MONITOR 07	.67487	.77673	1.146	.94227	.88497
MONITOR 08	.58573	.89133	1.17147	1.09507	.9359
MONITOR 09	.77673	.8404	.98047	.98047	.89452
MONITOR 10	.4202	1.0314	1.1078	.86587	.85632
MONITOR 11	.5348	.44567	.8786	.96773	.7067
MONITOR 12	.39473	.63667	1.01867	1.04413	.77355
MONITOR 13	.81493	1.00593	.955	1.1842	.99002
MONITOR 14	.7258	1.08233	1.36247	1.27333	1.11098
MONITOR 15	.78947	.77673	.8404	.8404	.81175
MONITOR 16	.9168	1.04413	1.19693	1.01867	1.04413
MONITOR 17	1.28607	1.59167	1.6426	1.337	1.46433
MAXIMUM...	1.28607	1.59167	1.6426	1.337	1.46433
MINIMUM...	.39473	.44567	.8404	.71307	.59847
MEAN.....	.70408	.87186	1.07709	.99844	.91287
STAND.DEV.	.22923	.26748	.20577	.16056	.21576

Table 5.1. Rolling annual averages (index 6 -- fractions of annual NAAQS) concentrations at monitor locations for sulfur.

	POLLUTANT LEVEL IN YEAR					5 YEAR
	1974	1975	1976	1977	1978	MEAN
MONITOR 01	1.	1.08	1.08	1.05333	.90667	1.024
MONITOR 02	.73333	.76	.86667	.89333	.77333	.80533
MONITOR 03	.81333	.93333	.86667	.86667	.73333	.84267
MONITOR 04	.89333	.86667	.82667	.93333	.90667	.88533
MONITOR 05	1.01333	.90667	.84	1.12	1.17333	1.01067
MONITOR 06	.92	.93333	1.	1.01333	.81333	.936
MONITOR 07	.93333	1.12	.98667	.88	.84	.952
MONITOR 08	1.08	1.17333	1.14667	1.05333	.96	1.08267
MONITOR 09	1.05333	1.04	1.54667	1.48	.90667	1.20533
MONITOR 10	.96	.86667	.96	1.12	.98667	.97867
MONITOR 11	.78667	.77333	.74667	.76	.64	.74133
MONITOR 12	.77333	.81333	.8	.81333	.68	.776
MONITOR 13	.84	.90667	.88	.82667	.68	.82667
MONITOR 14	.92	.93333	.77333	1.02667	.85333	.90133
MONITOR 15	1.14667	1.13333	1.06667	1.09333	.97333	1.08267
MONITOR 16	1.04	1.10667	1.02667	1.05333	.90667	1.02667
MONITOR 17	1.04	1.14667	1.21333	1.38667	1.21333	1.2
MAXIMUM...	1.14667	1.17333	1.54667	1.48	1.21333	1.312
MINIMUM...	.73333	.76	.74667	.76	.64	.728
MEAN.....	.93804	.9702	.97804	1.02196	.87922	.95749
STAND.DEV.	.11925	.13651	.19877	.19189	.15885	.16105

Table 5.2. Rolling annual averages (index 6 -- fractions of annual NAAQS) concentrations at monitor locations for TSP.

Figure 5.2. 1975 Isopleths and pollution scores for sulfur at monitors.*

Figure 5.3. 1976 Isopleths and pollution scores for sulfur at monitors.*
 * 3 hour peaks above annual(year) averages. Isopleths for the latter.

Figure 5.4. 1977 Isopleths and pollution scores for sulfur at monitors.*

Figure 5.5. 1978 Isopleths and pollution scores for sulfur at monitors.*

* 3 hour peaks above annual (year) averages. Isopleths for the latter.

Figure 5.6. 1975 Isopleths and pollution scores for TSP at monitors.*

Figure 5.7. 1976 Isopleths and pollution scores for TSP at monitors.*
 * 24 hour peaks above annual(year) averages. Isopleths for the latter.

Figure 5.8. 1977 Isopleths and pollution scores for TSP at the monitors.*

Figure 5.9. 1978 Isopleths and pollution scores for TSP at the monitors.*
 * 24 hour peaks above annual (year) averages. Isopleths for the latter.

peaks (index 4) for TSP at monitor locations. Isopleths for TSP are somewhat similar to those drawn for sulfur, although less clearly defined. The lack of gradient is noticeable throughout most of the region. Monitors 3 and 9, lying in the flat, northwest area of the region away from known major combustion sources, show quite high TSP levels. Particulate levels are also high surrounding the Homer City area, near monitors 1, 2, 3, and 6. This evidence suggests that TSP values are not good indicators of fossil fuel pollutants, since we would expect the dispersion of the heavy particulates (which primarily show up on the Hi-Vol sampler) to be more limited than the dispersion of sulfur dioxide. These particulates probably result from agricultural, roadway and urban dust, in addition to emissions from the power plants.

Table 5.3 shows correlation coefficients for the data in Figures 5.1 to 5.9. There is only 34% correlation between all annual TSP-sulfur concentrations. There is practically no correlation between 24 hour peaks of these pollutants: correlation coefficients range from -0.140 to 0.015. Sulfur concentrations at averaging times of 24 hours and a year tend to show higher correlations than TSP concentrations for the same averaging times. The next sections delve into these problems of stability and consistency in the pollution patterns.

5.1.1 Characterization of Geographic Units

Pollution scores for each geographic unit, that is, either a school or the population-weighted centroid of each district, are derived from three monitors. In part, as a consequence of the planar estimate or

	<u>TSP -- Sulfur</u>		<u>Annual -- 24 hour</u>	
<u>Year</u>	<u>Annual</u>	<u>24 hour</u>	<u>TSP</u>	<u>sulfur</u>
1974	-	-	-0.067	-
1975	0.551	0.015	0.334	0.814
1976	0.358	-0.140	0.762	0.893
1977	0.342	-0.059	0.512	0.675
1978	0.108	-0.111	0.366	0.607
all years	0.341	-0.028	0.481*	0.764

*1975 to 1978 only.

Table 5.3 Correlation coefficients for the 17 monitors using annual calendar average (index 11) and 24-hour 99th percentile concentrations.

	<u>Districts</u>		<u>Schools</u>		<u>Townships</u>	
	<u>Sulfur</u>	<u>TSP</u>	<u>Sulfur</u>	<u>TSP</u>	<u>Sulfur</u>	<u>TSP</u>
Index 4	0.949	0.920	0.898	0.900	0.621	0.608
Index 5	0.892	0.980	0.729	0.689	0.682	0.548
Index 6	0.895	0.991	0.701	0.760	0.699	0.708

Table 5.4. Comparison of district and school range to monitor range using ratios of standard deviations of three different indices. (For sulfur: Index 4: 99th percentile, 8 hour mean; index 5: 99 percentile, 24 hour mean; Index 6: annual mean. See Table 4.4 for TSP index definitions.

averaging in the case of the townships, schools, and districts, pollutant scores had slightly smaller ranges than those of the monitors. (Derivation of these scores is shown in Chapter 4.) The magnitude of this "compression" may be estimated by comparing the standard deviation of the districts and schools to that of the monitors. Table 5.4 shows these ratios, based on the 4 or 5 year average of the annual standard deviations for the respective indices. On the average, the triangulation procedure reduces the range in the districts (which should be as diverse as the monitors) by less than 10%. Schools have been compressed more, up to 30%, due in part to the triangulation, but mainly due to their more restricted geographic location; few schools are in very dirty areas. Townships have been compressed up to 40%, due to averaging. These estimates, however, belie the actual range. A population weighted estimate would be a better indicator of range than standard deviations. For example, although township scores appear to be compressed more than school scores, they in fact exhibit a larger overall range. It is advantageous to maintain as great a range as possible in epidemiological investigations. Thus, it may be advantageous to use district or township scores rather than school scores in the cross-sectional analysis. The use of three or more monitors also tended to smooth out fluctuations, thus, to a limited extent, to dampen year to year variations.

In some cases, the triangulation procedure is very sensitive to the monitors used. The isopleths drawn earlier take into account concentrations at all monitors. There are several situations in the Chestnut Ridge area where district centers or school locations were located about the same (effective) distance from several monitors. In one

of these cases the triangulation procedure predicted a different result than indicated by isopleths. Monitor selection was changed accordingly. Room for error increases further when the extrapolation point is outside the triangle formed by the three monitors used. It is possible that this might actually increase the range, although this was not observed. This occurs mostly in the periphery of the study region, including districts 1, 2, 3, 17, and 24. Pollution scores in these areas may be weighted to reflect their uncertainty, as discussed in Section 5.5.

Figure 5.10 shows a normal probability plot of four years of annual TSP concentrations at the 36 districts. These concentrations are approximately normally distributed, as demonstrated by their proximity to the regression line, indicated by asterisks. Four high outlayers at the extreme high end of the distribution, are the major deviation from normality. Figure 5.11 shows a similar plot for TSP concentrations at the 99th percentile at an averaging time of 8 hours. These concentrations are approximately log-normally distributed, as will be seen in Section 5.3.2. The varying distributions of pollutant measures play a major role in determining the stability and usefulness of indices.

5.2. Analysis of Stability

The development of a stable, precise and valid pollution measure is a major problem in air pollution epidemiology and risk-cost-benefit assessments. This section discusses stability, that is, the year to year variation of pollutant levels.

Figure 5.10. Normal probability plots of running annual mean sulfur concentrations. 4 years of data at the 36 districts. (Asterisks denote regression line.)

Figure 5.11. Normal probability plots of 8 hour 99th percentile sulfur concentrations. 4 years of data at the 36 districts. (Asterisks denote regression line.)

The basic problem of chronological stability is illustrated in Figures 5.12. Five years of pollution data at the 36 districts have been ranked by the means at each district and plotted. Figure 5.12a shows annual TSP concentrations. Figure 5.12b shows 24 hour 99th percentile concentrations. Concentrations do show a continuum, though it is difficult to distinguish any particular district -- except perhaps for the extremes -- as being either "clean" or "dirty." The majority of the districts, for example, those between ranks 10 and 30, display only a faintly increasing trend. Discerning any trend at all is difficult with the 24 hour peak measures. Categorizations as "clean" or "dirty" are quite arbitrary.

Annual variations can be significant and cause dramatic changes in pollution scores or rankings of the geographic units. Tables 5.5 and 5.6 show correlations among the mean, first year (1974 or 1975), and last year (1978) pollution scores as well as two chronological weighting functions. These linear weighting functions are shown below. The integers signify the relative contribution of the various years.

<u>Year</u>	<u>Sulfur Pollution Scores</u>		<u>TSP Pollution Scores</u>	
	Increasing	Decreasing	Increasing	Decreasing
	<u>Weight</u>	<u>Weight</u>	<u>Weight</u>	<u>Weight</u>
1974	-	-	1	5
1975	1	4	2	4
1976	2	3	3	3
1977	3	2	4	2
1978	4	1	5	1

These weightings are designed to smooth out extreme variations due to

Figure 5.12a. Five years of annual TSP concentrations (index 6) at the 36 districts, ranked by 5 year means.

Figure 5.12b. Five years of 24 hour peak TSP concentrations (index 4) at the 36 districts, ranked by 5 year means.

24 hour 99th Percentile Sulfur Concentrations (index 5)

	<u>1975</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1975	1				
1978	-.153	1			
Mean	0.832	0.173	1		
Decreasing	0.930	0.107	0.953	1	
Increasing	0.520	0.574	0.882	0.697	1

Rolling Annual Mean Sulfur Concentrations (index 6)

	<u>1975</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1975	1				
1978	0.205	1			
Mean	0.765	0.558	1		
Decreasing	0.859	0.429	0.980	1	
Increasing	0.607	0.688	0.969	0.901	1

Sulfur Exceedences above 30 ppm

	<u>1975</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1975	1				
1978	0.446	1			
Mean	0.883	0.655	1		
Decreasing	0.949	0.511	0.976	1	
Increasing	0.622	0.845	0.902	0.785	1

Sulfur Exceedences above 80 ppm

	<u>1975</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1975	1				
1978	0.381	1			
Mean	0.978	0.353			
Decreasing	0.903	0.301	0.997	1	
Increasing	0.959	0.434	0.993	0.981	1

Table 5.5. Correlation matrices for sulfur measures at the 14 schools over different years and weighting functions.

24 hour 99th Percentile TSP Concentrations (index 4)

	<u>1974</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1974	1				
1978	0.130	1			
Mean	-.323	0.103	1		
Decreasing	0.089	-.232	0.828	1	
Increasing	-.465	0.234	0.971	0.671	1

Rolling Annual Mean TSP Concentrations (index 6)

	<u>1974</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1975	1				
1978	0.791	1			
Mean	0.880	0.802	1		
Decreasing	0.867	0.691	0.983	1	
Increasing	0.866	0.880	0.986	0.938	1

TSP Exceedences above 75 micrograms/cu.m.

	<u>1974</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1974	1				
1978	0.584	1			
Mean	0.756	0.858	1		
Decreasing	0.810	0.805	0.994		
Increasing	0.752	0.899	0.994	0.976	1

TSP Exceedences above 150 micrograms/cu.m.

	<u>1974</u>	<u>1978</u>	<u>Mean</u>	<u>Weighted Concentrations</u>	
				<u>Decreasing</u>	<u>Increasing</u>
1974	1				
1978	0.538	1			
Mean	-.052	0.317	1		
Decreasing	0.222	0.033	0.798	1	
Increasing	-.198	0.424	-.940	0.543	1

Table 5.6. Correlation matrices for TSP measures at the 14 schools over different years and weighting functions.

error or "freak" conditions in any particular year, but also to maintain the chronological significance which is lost by averaging.

Tables 5.5 and 5.6 indicate that negative correlations occurred for 24 hour peak measures of both sulfur and TSP. Comparisons of weighted scores show a correlation of 0.7. Still, this indicates that different sites experienced high peaks in 1974-5 than in 1978. Changes in annual concentrations changed less dramatically as shown by a correlation coefficient of 0.9 between weighted averages. However, sulfur concentrations in 1975 had a correlated coefficient of only 0.205 to concentrations in 1978. These tables also show similar changes for the exceedence measures (based on the shortest averaging time). All exceedence measures show less variation than their analogous concentration measures.

In epidemiology, there are several approaches which may be used to minimize the effects of year to year variations in pollutant concentration. First, only the extremely dirty and clean geographic units may be employed. This may tend to increase the power of the health effects association. Table 5.7 ranks the cleanest 9 and the dirtiest 9 districts for all 11 indices. (Rankings are based on mean annual rankings of sulfur concentrations, as detailed later in section 5.2.3.) This method is unsatisfactory because of the imbalance in the resulting populations sample size. For example, for annual sulfur values (index 6), the total population in the cleanest group of 9 districts is 2147, but only 908 women reside in the 9 dirtiest districts. Table 5.8 uses a similar approach, but clean or dirty districts are selected until a population constraint of a minimum of 1000 women is met. The clean group

DISTRICT			INDEX NUMBER										
NO.	POP		1	2	3	4	5	6	7	8	9	10	11
DISTRICT:	1	5			6		3						
DISTRICT:	2	31											
DISTRICT:	3	77		-9			-7			6	8		
DISTRICT:	4	24	9	3	3	2							
DISTRICT:	5	9		4	8	4							
DISTRICT:	6	37	-9		-9						-8	-8	-9
DISTRICT:	7	79	6			8		3	4	5		9	5
DISTRICT:	8	66	3	7	7			5	9	7	7	8	7
DISTRICT:	9	68				9	9						9
DISTRICT:	10	46	5	9	9		8	6			5	5	6
DISTRICT:	11	41	8	8	6	6	4				9	7	
DISTRICT:	12	114	7	5	4	7	3	7			6	6	8
DISTRICT:	13	110	-7	-7	-7	-7		-8	8		-9	-9	
DISTRICT:	14	103					5			9			
DISTRICT:	15	220							3	4			4
DISTRICT:	16	524					-9		2				
DISTRICT:	17	284	1	1	1	1	2	1	5	1	3	3	1
DISTRICT:	18	53	2	2	2	5	1	2	7	2	2	2	2
DISTRICT:	19	385						-7	-9	-9	-7	-7	-8
DISTRICT:	20	77					7	9			8		
DISTRICT:	21	129											
DISTRICT:	22	370							-6				
DISTRICT:	23	161							1	3			
DISTRICT:	24	52								-8			
DISTRICT:	25	66								-7			
DISTRICT:	26	60	-8	-8	-8	-8	-8	-6	-5	-6	-6	-6	-6
DISTRICT:	27	63	4		5		6	4			1	1	3
DISTRICT:	28	191	-5	-6	-5	-6	-5	-9	-4	-5			-7
DISTRICT:	29	322	-3	-4	-4	-4	-3	-3	-1	-1	-2	-2	-3
DISTRICT:	30	126						8			4	4	
DISTRICT:	31	233	-2	-3	-3	-2	-1	-2	-3	-3	-3	-3	-2
DISTRICT:	32	192	-4	-2	-2	-3	-4	-4	-8	-4	-4	-4	-4
DISTRICT:	33	233	-1	-1	-1	-1	-2	-1	-2	-2	-1	-1	-1
DISTRICT:	34	421	-6	-5	-6	-5	-6	-5			-5	-5	-5
DISTRICT:	35	279				-9				6			
DISTRICT:	36	197							-7				
NO. CLEAN			9	9	9	9	9	9	9	9	9	9	9
POP CLEAN			770	642	700	677	849	908	1574	1322	870	872	993
NO. DIRTY			9	9	9	9	9	9	9	9	9	9	9
POP DIRTY			1799	1839	1799	2041	2253	2147	2183	1734	1993	1993	2074

Table 5.7. Ranked cleanest 9 districts (1 is cleanest) and dirtiest 9 districts (negative numbers, -1 is dirtiest) for the mean ranked sulfur concentrations of the 11 arrowhead indices. (Rankings described in Section 5.2.3). Population in districts from the adult women study.

DISTRICT			INDEX NUMBER										
	NO.	POP	1	2	3	4	5	6	7	8	9	10	11
DISTRICT:	1	5	12	6	12	3							
DISTRICT:	2	31				14							
DISTRICT:	3	77											
DISTRICT:	4	24	9	3	3	2							
DISTRICT:	5	9	11	4	8	4	10						
DISTRICT:	6	37											
DISTRICT:	7	79	6	12	11	8		3	4	5	11	9	5
DISTRICT:	8	66	3	7	7	11		5			7	8	7
DISTRICT:	9	68	10	10	10	9	9				10	10	9
DISTRICT:	10	46	5	9	9	10	8	6			5	5	6
DISTRICT:	11	41	8	8	6	6	4				9	7	10
DISTRICT:	12	114	7	5	4	7	3	7			6	6	8
DISTRICT:	13	110											
DISTRICT:	14	103					5	10					
DISTRICT:	15	220	14	14	14	15			3	4			4
DISTRICT:	16	524							2				
DISTRICT:	17	284	1	1	1	1	2	1	5	1	3	3	1
DISTRICT:	18	53	2	2	2	5	1	2		2	2	2	2
DISTRICT:	19	385											
DISTRICT:	20	77	13	13	13	12	7	9			8	11	
DISTRICT:	21	129											
DISTRICT:	22	370											
DISTRICT:	23	161							1	3			
DISTRICT:	24	52											
DISTRICT:	25	66											
DISTRICT:	26	60							-5				
DISTRICT:	27	63	4	11	5	13	6	4			1	1	3
DISTRICT:	28	191	-5		-5		-5		-4	-5			
DISTRICT:	29	322	-3	-4	-4	-4	-3	-3	-1	-1	-2	-2	-3
DISTRICT:	30	126						8			4	4	
DISTRICT:	31	233	-2	-3	-3	-2	-1	-2	-3	-3	-3	-3	-2
DISTRICT:	32	192	-4	-2	-2	-3	-4	-4		-4	-4	-4	-4
DISTRICT:	33	233	-1	-1	-1	-1	-2	-1	-2	-2	-1	-1	-1
DISTRICT:	34	421		-5		-5		-5			-5	-5	-5
DISTRICT:	35	279					11			6			
DISTRICT:	36	197											
NO. CLEAN			14	14	14	15	11	10	5	6	11	11	10
POP CLEAN			1149	1149	1149	1180	1137	1011	1268	1076	1017	1017	1034
NO. DIRTY			5	5	5	5	5	5	5	5	5	5	5
POP DIRTY			1171	1401	1171	1401	1171	1401	1039	1171	1401	1401	1401

Table 5.8. Ranked cleanest districts (1 is cleanest) and dirtiest districts (negative numbers, -1 is dirtiest) which contain a minimum of 1000 women, using mean ranked sulfur concentrations of the 11 arrowhead indices. (Rankings described in Section 5.2.3). Population in districts from the adult women study.

of 1000 requires about twice as many districts -- up to 15 -- as the dirty group of 1000. We saw earlier how difficult it was to characterize these districts as dirty. Additionally, such categorizations draw fairly arbitrary divisions between clean and dirty regions. The pollution gradient in a small geographic area like Chestnut Ridge seems continuous, and not well represented by such a categorical variable. Other approaches are described in the next section.

5.2.1 Relative Stability

A second method to minimize the effects of year to year variation might be a relative approach, that is, finding areas which are clean or dirty with respect to the other areas in that year, for as many years of data as is relevant or available. Thus, pollution scores in each year exert the same influence, regardless of the actual concentration. Such relative methods lose information about the absolute level of the concentrations: consequently, no particular type of dose-response function is involved. However, they may be useful if they lead to more stable characterizations, particularly for cross-sectional epidemiologic investigations. Both parametric and non-parametric methods are available.

Parametric methods maintain information concerning the distribution of pollutant levels. Two methods were attempted. First, the pollution scores were normalized on an annual basis. Then, the several years of normalized scores are averaged by geographic units. Table 5.9 displays such normalized scores for the districts using annual sulfur concentrations. Clean districts, that is, those lying below the mean for

DISTRICT		NORMALIZED ANNUAL VALUE				MEAN
NUMBER		1975	1976	1977	1978	VALUE
DISTRICT RANK						
1.....	17.	-.974	-1.738	-1.115	-.446	-1.068
2.....	18.	-.711	-1.247	-1.104	-.756	-.955
3.....	27.	.05	-.835	-1.623	-1.079	-.872
4.....	8.	-.893	-1.703	-.902	.072	-.857
5.....	7.	-1.243	-1.156	-.685	-.247	-.833
6.....	10.	-.057	-.053	-1.033	-1.393	-.634
7.....	12.	.045	.137	-1.04	-1.542	-.6
8.....	20.	-.054	-.037	-.916	-1.344	-.588
9.....	35.	-1.583	-.98	-.13	.57	-.531
10.....	11.	.167	.265	-.95	-1.414	-.483
11.....	30.	.344	-.415	-1.131	-.694	-.474
12.....	14.	-.839	-.769	-.293	.105	-.449
13.....	15.	-.883	-.235	.327	-.861	-.413
14.....	9.	.132	.068	-.825	-.977	-.401
15.....	3.	-1.525	.672	.246	-.528	-.284
16.....	23.	-1.02	.216	.799	-1.093	-.275
17.....	22.	-.238	-.367	.135	-.383	-.213
18.....	36.	-.574	-.62	.454	-.072	-.203
19.....	16.	-.686	-.346	.474	-.155	-.178
20.....	2.	-.377	.155	-.167	-.062	-.113
21.....	21.	-.256	-.409	.629	-.37	-.101
22.....	5.	.247	-.097	-.471	-.009	-.083
23.....	4.	.252	-.191	-.407	.225	-.03
24.....	1.	.273	-.141	-.389	.222	-.009
25.....	24.	1.018	-.502	-.35	.43	.149
26.....	25.	1.051	-.381	-.288	.43	.203
27.....	28.	.11	.087	.02	.659	.219
28.....	6.	-.69	.091	.625	1.011	.259
29.....	19.	.472	.393	.247	-.056	.264
30.....	13.	-.592	.574	1.048	.367	.349
31.....	26.	.784	.441	.112	.406	.436
32.....	34.	.092	.589	1.074	1.379	.783
33.....	32.	1.185	1.005	1.038	.712	.985
34.....	29.	1.544	1.806	1.108	2.259	1.679
35.....	31.	2.536	2.616	2.451	2.09	2.423
36.....	33.	2.892	3.109	3.035	2.543	2.895

Table 5.9. Normalized pollution scores for rolling annual average (index 6) for sulfur at the districts. Clean districts have negative scores.

NORMALIZED POLLUTION SCORE FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
01 BEN FRA	-0.92	-0.81	-0.84	-0.78	-1.01	-1.07	-0.56	-1.01	-0.74	-0.83	-0.96
02 UNITED.	-0.36	-0.31	-0.30	-0.10	-0.32	-0.34	0.54	0.44	-0.70	-0.63	-0.32
03 BURREL.	0.39	0.23	0.26	0.12	0.36	0.11	0.68	0.58	0.08	0.17	0.32
04 EASTPIK	-0.60	-0.54	-0.56	-0.53	-0.61	-0.58	-0.66	-0.79	-0.34	-0.46	-0.65
05 FAIRFLD	-0.10	0.14	0.11	0.33	-0.29	-0.45	-0.16	-0.76	-0.44	-0.43	-0.42
06 BLAIRSV	0.24	0.10	0.13	0.02	0.22	-0.05	0.74	0.56	-0.11	-0.03	0.16
07 CLYMER.	-0.33	-0.27	-0.31	-0.31	-0.00	-0.13	-1.44	-1.01	-0.12	-0.35	-0.72
08 PENNRUN	-0.18	-0.25	-0.23	-0.27	-0.23	0.38	0.23	0.65	0.36	0.39	0.37
09 EISENHR	-0.44	-0.43	-0.43	-0.44	-0.04	-0.44	-0.69	-0.19	-0.38	-0.34	-0.36
10 HORMANN	-0.12	-0.13	-0.15	-0.18	0.38	0.15	-1.03	-0.17	0.06	-0.01	-0.18
11 HOMER C	-0.00	-0.19	-0.15	-0.28	-0.18	0.44	0.50	0.53	0.57	0.67	0.63
12 LAURELV	0.00	0.22	0.20	0.39	-0.19	-0.31	-0.07	-0.63	-0.31	-0.30	-0.28
13 NEWFLOR	2.09	2.05	2.05	1.94	1.58	2.29	1.12	1.17	2.08	2.08	2.16
14 3RDWARD	0.33	0.18	0.21	0.08	0.33	-0.00	0.77	0.61	-0.01	0.07	0.27

Table 5.10. Mean normalized pollution scores for 11 sulfur arrowhead indices at the schools. Clean schools have negative scores.

NORMALIZED POLLUTION SCORE FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
01 BEN FRA	-0.26	-0.09	-0.11	0.05	0.09	-0.33	-0.54	-0.35	-0.09	-0.12	-0.32
02 UNITED.	1.31	0.58	0.79	0.22	0.33	1.01	1.79	1.65	0.96	1.10	1.11
03 BURREL.	0.65	0.63	0.66	0.45	0.78	0.68	0.23	0.06	0.39	0.47	0.45
04 EASTPIK	-0.03	0.02	0.01	0.05	0.26	-0.01	-0.32	0.07	0.13	0.09	0.02
05 FAIRFLD	-0.78	-0.49	-0.62	-0.18	-0.46	-0.86	-0.27	-0.69	-0.87	-0.90	-0.98
06 BLAIRSV	0.15	0.12	0.14	0.05	-0.11	0.23	-0.08	-0.20	-0.02	0.08	0.10
07 CLYMER.	-0.51	-0.55	-0.57	-0.66	-0.63	-0.26	-0.27	0.40	-0.13	-0.17	-0.13
08 PENNRUN	0.95	0.43	0.50	0.46	1.09	0.94	-0.12	0.21	0.89	0.84	1.09
09 EISENHR	-1.32	-0.34	-0.50	-0.09	-0.79	-1.52	-0.80	-1.47	-1.13	-1.34	-1.37
10 HORMANN	-0.98	-0.45	-0.55	-0.40	-0.80	-0.95	-0.49	-0.82	-0.71	-0.58	-0.76
11 HOMER C	0.20	-0.18	-0.10	-0.22	-0.28	0.24	0.31	0.41	0.12	0.21	0.33
12 LAURELV	-0.71	-0.47	-0.58	-0.18	-0.42	-0.77	-0.22	-0.62	-0.79	-0.82	-0.88
13 NEWFLOR	1.11	0.63	0.76	0.37	0.99	1.31	0.86	1.32	1.22	1.30	1.22
14 3RDWARD	0.21	0.15	0.18	0.07	-0.05	0.29	-0.07	-0.18	0.01	0.12	0.13

Table 5.11. Mean normalized pollution scores for 11 TSP arrowhead indices at the schools. Clean schools have negative scores.

NORMALIZED POLLUTION SCORE FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
01 BEN FRA	-1.17	-0.90	-0.94	-0.72	-0.92	-1.41	-1.10	-1.36	-0.83	-0.94	-1.28
02 UNITED.	0.95	0.27	0.49	0.12	0.01	0.68	2.34	2.08	0.26	0.47	0.79
03 BURREL.	1.04	0.87	0.92	0.56	1.14	0.79	0.92	0.64	0.47	0.64	0.76
04 EASTPIK	-0.63	-0.52	-0.55	-0.48	-0.36	-0.59	-0.98	-0.71	-0.22	-0.36	-0.64
05 FAIRFLD	-0.88	-0.35	-0.50	0.15	-0.76	-1.31	-0.43	-1.45	-1.30	-1.34	-1.40
06 BLAIRSV	0.37	0.22	0.27	0.07	0.12	0.17	0.66	0.37	-0.13	0.06	0.26
07 CLYMER.	-0.84	-0.81	-0.88	-0.97	-0.63	-0.39	-1.71	-0.60	-0.25	-0.53	-0.85
08 PENNRUN	0.78	0.18	0.27	0.19	0.85	1.32	0.11	0.87	1.25	1.23	1.46
09 EISENHR	-1.76	-0.77	-0.93	-0.53	-0.83	-1.96	-1.49	-1.66	-1.50	-1.68	-1.74
10 HORMANN	-1.10	-0.58	-0.70	-0.58	-0.41	-0.80	-1.52	-0.78	-0.65	-0.89	-0.94
11 HOMER C	0.20	-0.37	-0.25	-0.49	-0.45	0.68	0.81	0.94	0.69	0.88	0.96
12 LAURELV	-0.71	-0.24	-0.38	0.21	-0.61	-1.08	-0.29	-1.24	-1.09	-1.12	-1.16
13 NEWFLOR	3.20	2.68	2.81	2.32	2.58	3.60	1.99	2.49	3.30	3.38	3.38
14 3RDWARD	0.53	0.34	0.39	0.15	0.28	0.29	0.70	0.43	0.00	0.20	0.39

Table 5.12. Mean normalized pollution scores for combined sulfur and TSP arrowhead indices at the schools. Clean schools have negative scores.

the respective year, have negative scores. Tables 5.10 and 5.11 show mean normalized scores for all indices for both sulfur and TSP at school locations. This format makes it particularly convenient to locate outlayers, that, is, points which are more than one standard deviation above or below the mean. Normalization has the additional advantage of allowing pollutants to be combined by the simple addition of normalized values (with indices that have similar types of distributions). Table 5.12 is the sum of the data in Table 5.10 and 5.11, effectively combining sulfur and TSP scores. School number 9, located west of Indiana City, is the only school which consistently stays clean. School number 13, near the Connemaugh and Seward Power Plants (just hundreds of feet away) is consistently dirty.

A clustering analysis, which is used parametrically, provides a second approach. Cluster analysis is an iterative technique which provides categorical outcomes that meet specified measures of similarity or dissimilarity, called proximity measures. The proximity measures used here are designed to efficiently minimize the sum of squares between categories. Then, an aggregate hierarchical clustering of the proximity matrix is performed. Figure 5.13 shows the cluster tree: the 36 districts are ranked from cleanest to dirtiest, and are connected with branches whose horizontal length represents the proximity to the next group or cluster. In this figure, the lowest four districts comprise a groups far dirtier than the remaining 32. Graphically, the number of categories increases as one moves to the left of the tree. Clustering into 6 categories was done for several indices. Results are shown later, in Table 5.26. These partitions are then optimized in an iterative process.

Six partitions explained about 90% of the variance between the groups. Any number of partitions may be selected, but unfortunately, the size of categories can not be selected directly. To some extent, this limits the use of this technique. In addition, because this "experimental" technique is probably not available to or well understood by most researchers, its use is de-emphasized in this report.

The simplest non-parametric method which adjusts for year to year variation is an annual ranking of pollution scores. This is demonstrated for annual sulfur concentrations in Table 5.13. An examination of the average rank with the standard deviation of the district in mind may assist the investigator in making decisions as to what is "clean" or "dirty." Note that the average rank is different from the average normalized rank, shown earlier in Table 5.9, for all but extreme districts.

Relative methods are not in common use in epidemiology. They are not likely to be much use in risk-cost-benefit assessments. However, they may offer some insight into cross-sectional epidemiologic studies. In this investigation, they are used primarily to assess and compare the stability of indices. However, normalized scores facilitate studies of changes over time, as well as studies of pollutant interaction and may make more sense than cross-products. Normalized scores require a second level of transformation, which may be a slight disadvantage 1/.

1/ In the Chestnut Ridge investigation, they required a third level of transformation, since monitor indices from the arrowhead curve could be directly extrapolated into scores for the geographic units. Normalized scores required that monitor indices be extrapolated to school or district scores, which could then be normalized. (It is impossible to normalize monitor scores and then directly extrapolate these to normalized district scores.)

Figure 5.13. Clustering Tree for rolling annual sulfur concentrations in the 36 districts. Districts numbers are listed at the right and are ranked from cleanest to dirtiest. Horizontal length of the branches represent the proximity to the next group or cluster.

DISTRICT		RANK IN YEAR				MEAN	STANDARD
NUMBER		1975	1976	1977	1978	RANK	DEVIATE
DISTRICT	RANK						
1.....	17.	5.	1.	3.	12.	5.25	4.79
2.....	18.	9.	3.	4.	9.	6.25	3.2
3.....	7.	3.	4.	11.	15.	8.25	5.74
4.....	27.	20.	6.	1.	6.	8.25	8.18
5.....	8.	6.	2.	9.	21.	9.5	8.19
6.....	10.	17.	19.	6.	3.	11.25	7.93
7.....	12.	19.	24.	5.	1.	12.25	11.
8.....	30.	28.	10.	2.	10.	12.5	11.
9.....	20.	18.	20.	8.	4.	12.5	7.72
10.....	14.	8.	7.	16.	22.	13.25	7.09
11.....	35.	1.	5.	19.	29.	13.5	12.9
12.....	15.	7.	15.	25.	8.	13.75	8.3
13.....	11.	24.	27.	7.	2.	15.	12.36
14.....	9.	23.	21.	10.	7.	15.25	7.93
15.....	22.	16.	13.	22.	13.	16.	4.24
16.....	36.	13.	8.	26.	17.	16.	7.62
17.....	23.	4.	26.	30.	5.	16.25	13.67
18.....	3.	2.	32.	23.	11.	17.	13.19
19.....	16.	11.	14.	27.	16.	17.	6.98
20.....	21.	15.	11.	29.	14.	17.25	8.02
21.....	5.	25.	18.	12.	20.	18.75	5.38
22.....	2.	14.	25.	18.	18.	18.75	4.57
23.....	4.	26.	16.	13.	24.	19.75	6.24
24.....	1.	27.	17.	14.	23.	20.25	5.85
25.....	24.	31.	9.	19.	27.	20.5	10.25
26.....	25.	32.	12.	17.	28.	22.25	9.32
27.....	6.	10.	23.	28.	32.	23.25	9.57
28.....	28.	22.	23.	20.	30.	23.5	4.43
29.....	13.	12.	30.	32.	25.	24.75	9.
30.....	19.	29.	28.	24.	19.	25.	4.55
31.....	26.	30.	29.	21.	26.	26.5	4.04
32.....	34.	21.	31.	33.	33.	29.5	5.74
33.....	32.	33.	33.	31.	31.	32.	1.15
34.....	29.	34.	34.	34.	35.	34.25	.5
35.....	31.	35.	35.	35.	34.	34.75	.5
36.....	33.	36.	36.	36.	36.	36.	.
MEAN....	18.5	18.5	18.5	18.5	18.5	18.5	6.98

Table 5.13. Ranks of districts for rolling annual sulfur concentrations. Districts are ranked by mean rank. Standard deviation of change in rank over 4 years of data. Mean standard deviation for 36 districts is also shown.

5.3 Index Consistency

The first part of this section compares coefficients of variation (COV) among the various indices. The second part examines the stability of measures which have been constructed at several averaging times. These include the exceedence measures and concentrations at the 90th and 99th percentiles. This information is then used to select the most stable indices.

5.3.1 Stability of Arrowhead Indices and NAAQS

Annual coefficients of variance (COV) were calculated for all indices. Tables 5.14 and 5.15 show COVs at each monitor as well as the mean COV. Several trends are apparent. First, as expected, similar indices, i.e., 2 and 3, 4 and 5, 6 and 11, and to some extent, 7 and 8, track each other closely. Second, indices derived from high percentile groups, i.e., 4 and 5 show considerable heteroscedasticity. For example, the dirtiest monitor, number 17, shows a large COV with respect to the mean COV. This suggests that a (logarithmic) transformation will greatly reduce the annual variation. This is shown in Section 5.3.2. The annual means, indices 6 and 11, consistently show the lowest COVs, suggesting that these are the most stable measure. Note however, that the range is quite high -- up to 0.485 with sulfur and 0.239 with TSP (index 6). Lastly, the calendar year average, index 11, is more stable than the rolling annual average, index 6, as one would expect, since index 6 is designed to pick up peaks that may occur but which are not confined to

calendar year boundaries.

Table 5.16 displays mean standard deviations for changes in the ranks of the 36 districts for the 11 sulfur and TSP indices. Highest standard deviations are also shown. This provides a nonparametric comparison to the COV measures of stability. In general, results parallel the COV measures, though sulfur indices with unstable COVs appear more stable with this measure. This may be due to a reduction of heteroscedasticity in the rankings. The high standard deviations are evidence that the districts shifted around considerably over the 4 or 5 year period. Some districts shifted from the almost the cleanest to the almost the dirtiest. The high standard deviations are either due to highly sensitive or unstable pollution measures, large changes in ambient pollution, or to the small difference in pollution levels in many of the districts.

5.3.2 Exceedence Indices

Tables 5.17 and 5.18 show COVs for the two lowest thresholds, at 30 and 80 ppb, of sulfur exceedence measures. Both tables show a definite trend of greater stability with shorter averaging times. This may be due to the greater frequency of occurrences that short averaging times permit, which pushes the percentiles of this measure into more stable median measures. The most surprising and significant aspect of these measures is their relative insensitivity to averaging time, even from periods of one hour to one year, in the case of the lower threshold. Long averaging times in the 80 ppb exceedence index resulted in high COVs due to the many "zeros" which occur in this measure. Remarkably, one hour

	INDEX NUMBER										
	1	2	3	4	5	6	7	8	9	10	11
MONITOR 01	.337	.336	.332	.382	.417	.406	.484	.418	.316	.322	.12
MONITOR 02	.132	.162	.165	.226	.167	.113	.301	.162	.143	.15	.141
MONITOR 03	.174	.129	.129	.107	.256	.305	.291	.241	.214	.204	.195
MONITOR 04	.17	.175	.159	.18	.304	.261	.299	.224	.173	.145	.109
MONITOR 05	.104	.139	.132	.164	.145	.11	.049	.101	.112	.109	.107
MONITOR 06	.052	.065	.059	.041	.194	.115	.135	.193	.062	.057	.054
MONITOR 07	.362	.637	.591	.867	.318	.233	.056	.099	.083	.076	.064
MONITOR 08	.291	.433	.402	.639	.256	.28	.121	.107	.184	.178	.166
MONITOR 09	.16	.171	.175	.137	.24	.115	.168	.166	.153	.163	.168
MONITOR 10	.212	.153	.165	.175	.283	.359	.486	.542	.194	.214	.247
MONITOR 11	.201	.095	.114	.048	.167	.361	.38	.304	.298	.303	.299
MONITOR 12	.337	.336	.332	.382	.417	.406	.484	.418	.316	.322	.334
MONITOR 13	.056	.138	.127	.209	.121	.154	.154	.163	.113	.111	.101
MONITOR 14	.173	.21	.204	.244	.31	.254	.253	.215	.231	.218	.191
MONITOR 15	.073	.112	.098	.219	.296	.041	.197	.126	.083	.085	.09
MONITOR 16	.136	.245	.225	.32	.378	.111	.139	.109	.084	.076	.065
MONITOR 17	.319	.429	.408	.464	.485	.122	.166	.203	.267	.237	.214
MEAN.....	.193	.233	.225	.283	.28	.22	.245	.223	.178	.175	.157

Table 5.14. Coefficients of variation for annual variations in sulfur scores at the 17 monitors for the 11 arrowhead indices.

	INDEX NUMBER										
	1	2	3	4	5	6	7	8	9	10	11
MONITOR 01	.113	.201	.172	.274	.38	.072	.153	.168	.135	.114	.108
MONITOR 02	.073	.196	.169	.299	.14	.087	.283	.076	.085	.079	.077
MONITOR 03	.063	.143	.116	.221	.189	.088	.181	.147	.113	.084	.068
MONITOR 04	.071	.139	.117	.227	.099	.046	.267	.108	.12	.084	.065
MONITOR 05	.228	.614	.529	.858	.628	.138	.271	.103	.245	.204	.149
MONITOR 06	.11	.175	.148	.265	.163	.085	.225	.189	.073	.074	.117
MONITOR 07	.099	.089	.074	.192	.102	.114	.137	.202	.162	.131	.119
MONITOR 08	.097	.197	.166	.273	.217	.078	.212	.186	.173	.149	.105
MONITOR 09	.234	.307	.282	.369	.62	.239	.263	.196	.199	.185	.213
MONITOR 10	.07	.114	.073	.148	.17	.093	.244	.128	.15	.113	.091
MONITOR 11	.094	.292	.243	.392	.215	.079	.166	.187	.183	.144	.108
MONITOR 12	.076	.301	.251	.417	.197	.072	.191	.095	.033	.028	.072
MONITOR 13	.111	.139	.122	.187	.172	.106	.231	.187	.102	.089	.103
MONITOR 14	.169	.512	.438	.693	.211	.105	.258	.187	.146	.126	.114
MONITOR 15	.059	.178	.148	.239	.08	.064	.11	.125	.1	.086	.089
MONITOR 16	.061	.092	.058	.169	.172	.072	.248	.152	.064	.064	.07
MONITOR 17	.133	.27	.227	.313	.209	.105	.414	.184	.209	.173	.113
MEAN.....	.109	.233	.196	.326	.233	.097	.227	.154	.135	.113	.105

Table 5.15. Coefficients of variation for annual variations in TSP scores at the 17 monitors for the 11 arrowhead indices.

	1	2	3	4	5	6	7	8	9	10	11
Sulfur											
Std.Dev.	6.9	6.8	6.8	6.8	8.7	7.0	7.3	7.3	6.7	6.7	6.6
Highest	12	13	12	14	15	13	14	16	14	13	13
TSP											
Std.Dev.	5.6	8.6	8.3	9.1	6.8	5.0	8.3	5.6	7.0	6.7	5.6
Highest	13	15	15	14	14	13	15	11	16	15	14

Table 5.16 Mean standard deviation and highest standard deviation for ranked districts, as seen earlier in Table 5.13.

AVERAGING TIME -- HOURS													
	1	2	3	8	24	48	72	168	336	730	2190	4380	8760
MONITOR 01	.12	.12	.12	.11	.15	.14	.15	.21	.2	.31	.73	1.31	2.
MONITOR 02	.17	.19	.18	.22	.35	.28	.3	.4	.43	.61	.68	.92	.
MONITOR 03	.2	.2	.2	.29	.5	.32	.39	.5	.81	1.16	1.2	1.29	.
MONITOR 04	.11	.11	.1	.12	.13	.14	.17	.17	.22	.26	.76	1.79	2.
MONITOR 05	.08	.08	.08	.09	.11	.13	.14	.18	.21	.27	.67	.68	1.15
MONITOR 06	.06	.06	.06	.07	.08	.09	.09	.13	.14	.2	.21	.33	1.58
MONITOR 07	.06	.06	.06	.07	.08	.09	.07	.11	.08	.13	.39	1.58	2.
MONITOR 08	.12	.13	.13	.15	.17	.2	.21	.28	.31	.38	.73	1.07	1.34
MONITOR 09	.16	.16	.18	.18	.23	.25	.29	.32	.46	.59	.68	.91	.
MONITOR 10	.19	.2	.2	.22	.27	.32	.35	.4	.43	.45	.73	.71	1.33
MONITOR 11	.45	.46	.47	.48	.49	.34	.38	.45	.46	.62	1.28	2.	.
MONITOR 12	.42	.44	.44	.5	.57	.48	.54	.75	.99	1.19	1.25	1.18	1.26
MONITOR 13	.1	.1	.1	.11	.13	.15	.17	.21	.21	.29	.15	.29	1.87
MONITOR 14	.11	.12	.13	.13	.15	.16	.18	.21	.24	.28	.34	.69	.98
MONITOR 15	.1	.1	.1	.11	.15	.16	.18	.21	.35	1.15	1.1	2.	.
MONITOR 16	.05	.05	.06	.05	.07	.08	.09	.12	.1	.12	.14	.32	1.41
MONITOR 17	.06	.07	.07	.08	.13	.17	.19	.26	.27	.29	.31	.4	.37
MEAN.....	.15	.16	.16	.18	.22	.21	.23	.29	.35	.49	.67	1.03	1.02

Table 5.17. Coefficients of variation for annual variations in sulfur exceedence scores, % of time over 30 ppb for various averaging times.

AVERAGING TIME -- HOURS													
	1	2	3	8	24	48	72	168	336	730	2190	4380	8760
MONITOR 01	.1	.1	.11	.28	1.25	2.	2.	2.	.	.	.	.	.
MONITOR 02	.17	.14	.22	.34	.73	2.	.	.	.	.	.	.	.
MONITOR 03	.11	.14	.23	.61	1.3	2.	2.	.	.	.	.	.	.
MONITOR 04	.15	.16	.2	.32	.97	1.33	2.	2.	2.	.	.	.	.
MONITOR 05	.1	.09	.12	.22	.43	.91	2.	.	.	.	.	.	.
MONITOR 06	.04	.04	.06	.14	.33	.88	1.56	.	.	.	.	.	.
MONITOR 07	.09	.1	.15	.26	.58	.82	.84	2.	.	.	.	.	.
MONITOR 08	.15	.14	.16	.21	.84	1.49	2.	2.	.	.	.	.	.
MONITOR 09	.18	.16	.18	.34	.83	2.	2.	.	.	.	.	.	.
MONITOR 10	.12	.16	.18	.29	.68	.77	1.04	2.	.	.	.	.	.
MONITOR 11	.18	.22	.23	.36	.74	2.	2.	.	.	.	.	.	.
MONITOR 12	.18	.21	.23	.39	1.23	2.	2.	2.	.	.	.	.	.
MONITOR 13	.06	.06	.06	.06	.16	.75	1.16	.	.	.	.	.	.
MONITOR 14	.38	.37	.4	.41	.47	.72	.85	1.49	2.	2.	.	.	.
MONITOR 15	.07	.06	.06	.11	.76	1.47	2.	2.	.	.	.	.	.
MONITOR 16	.04	.04	.05	.09	.36	.71	.72	2.	.	.	.	.	.
MONITOR 17	.38	.4	.42	.49	.73	.68	.69	.81	.88	1.09	2.	.	.
MEAN.....	.15	.15	.18	.29	.73	1.33	1.46	1.08	.29	.18	.12	.	.

Table 5.18. Coefficients of variation for annual variations in sulfur exceedence scores, % of time over 80 ppb for various averaging times.

	AVERAGING TIME -- HOURS						
	24	72	168	336	730	2190	8760
MONITOR 01	.124	.226	.252	.496	.482	.611	1.092
MONITOR 02	.223	.889	1.319	1.412	2.236	.	.
MONITOR 03	.123	.43	.763	1.143	1.407	.	.
MONITOR 04	.105	.187	.421	1.57	1.377	2.236	.
MONITOR 05	.223	.22	.468	.801	.903	1.379	.959
MONITOR 06	.136	.136	.419	.609	.621	.963	2.126
MONITOR 07	.171	.481	.454	.59	.623	1.471	2.236
MONITOR 08	.217	.254	.298	.468	.723	.811	.841
MONITOR 09	.225	.172	.279	.377	.644	.713	.717
MONITOR 10	.168	.201	.318	.415	.62	1.023	2.236
MONITOR 11	.568	.691	1.192	2.236	2.236	.	.
MONITOR 12	.12	.462	.914	.	2.236	.	.
MONITOR 13	.193	.736	1.338	2.236	2.236	2.236	.
MONITOR 14	.229	.422	.514	.567	1.008	2.236	2.236
MONITOR 15	.078	.196	.246	.245	.254	.262	.563
MONITOR 16	.118	.083	.113	.272	.297	.5	.779
MONITOR 17	.171	.151	.252	.283	.307	.384	.189
MEAN.....	.188	.349	.562	.807	1.071	.872	.822

Table 5.19. Coefficients of variation for annual variations in TSP exceedence scores; % of time over 75 u.gms/cu.m. for various averaging times.

	AVERAGING TIME -- HOURS						
	24	72	168	336	730	2190	8760
MONITOR 01	.596	1.395	2.25	.	.	.	.
MONITOR 02	2.236	.	.	.	.	.	.
MONITOR 03	1.731	.	.	.	.	.	.
MONITOR 04	1.64	.	.	.	.	.	.
MONITOR 05	1.766	2.236	2.236	2.236	.	.	.
MONITOR 06	1.407	.	.	.	.	.	.
MONITOR 07	.596	.	.	.	.	.	.
MONITOR 08	.521	1.092	2.236	.	.	.	.
MONITOR 09	.76	2.236	2.236	2.236	2.236	.	.
MONITOR 10	.467	1.465	.	.	.	.	.
MONITOR 11	1.206	.	.	.	.	.	.
MONITOR 12	1.165	.	.	.	.	.	.
MONITOR 13	2.236	.	.	.	.	.	.
MONITOR 14	1.546	2.236	.	.	.	.	.
MONITOR 15	.607	.	.	.	.	.	.
MONITOR 16	1.783	2.236	.	.	.	.	.
MONITOR 17	.915	1.215	.	.	.	.	.
MEAN.....	1.246	.83	.527	.263	.132	.	.

Table 5.20. Coefficients of variation for annual variations in TSP exceedence scores; % of time over 150 u.gms/cu.m. for various averaging times.

averaging times at the lower threshold resulted in a measure more stable than any of the previous 11 sulfur indices. Moreover, the range of COVs is quite restricted -- only several monitors show COVs much greater than the mean. Table 5.18 also shows considerable heteroscedasticity, which is not surprising since this measure is derived from concentrations at high percentiles and at short averaging times. Based on this information, it was decided to use exceedence measures based on the lowest thresholds.

Tables 5.19 and 5.20 show COVs for the TSP exceedence measures. As with the sulfur measures, many zeros result at long averaging times. This problem is exacerbated with TSP, since the range of values is smaller than that of sulfur's. At the high threshold of 150 micrograms, only 4 monitors, numbers 8, 9, 10, and 15, had no zeros for the five years, even at the shortest averaging time of 24 hours. This implies that this threshold was too high for the construction of this measure.

5.3.3 90th and 99th Percentile Measures

Tables 5.21 shows the COVs for sulfur concentrations in the 90th and 99th percentile, at five averaging times. 90th percentile concentrations appear to have a COV some 50 to 60% lower than 99th percentile's COV. The logarithm of the concentration values was taken, and COVs developed from the transformed data, as shown in Table 5.22. Now, the COVs of the 90th and 99th percentiles are almost the same. In both cases, there is a slight trend towards greater stability with shorter averaging times.

	COV - 90TH PERCENTILE					COV - 99TH PERCENTILE				
	1HR	2HR	3HR	8HR	24HR	1HR	2HR	3HR	8HR	24HR
MONITOR 01	.14	.14	.15	.14	.14	.12	.1	.05	.19	.45
MONITOR 02	.14	.14	.14	.14	.15	.22	.09	.23	.15	.17
MONITOR 03	.21	.21	.22	.22	.24	.05	.02	.11	.34	.26
MONITOR 04	.19	.18	.19	.19	.21	.11	.13	.18	.23	.32
MONITOR 05	.12	.12	.12	.11	.13	.14	.13	.17	.23	.15
MONITOR 06	.06	.07	.07	.06	.06	.08	.03	.04	.16	.2
MONITOR 07	.09	.09	.09	.09	.09	1.18	1.01	1.12	.62	.33
MONITOR 08	.19	.19	.19	.2	.2	.76	.91	.74	.38	.26
MONITOR 09	.15	.15	.15	.15	.14	.18	.12	.14	.21	.25
MONITOR 10	.2	.2	.2	.19	.16	.1	.22	.18	.25	.3
MONITOR 11	.31	.32	.32	.32	.31	.03	.1	.05	.12	.17
MONITOR 12	.32	.32	.32	.33	.37	.33	.43	.41	.45	.45
MONITOR 13	.1	.11	.11	.11	.12	.14	.17	.21	.25	.12
MONITOR 14	.23	.22	.23	.25	.26	.21	.24	.25	.36	.32
MONITOR 15	.08	.09	.09	.09	.08	.3	.19	.22	.12	.31
MONITOR 16	.09	.1	.1	.09	.09	.26	.3	.33	.35	.4
MONITOR 17	.32	.32	.32	.31	.3	.49	.48	.52	.62	.55
MEAN.....	.17	.18	.18	.18	.18	.28	.28	.29	.3	.29

Table 5.21. Coefficients of variation for annual variations in sulfur dioxide 90th and 99th percentile concentrations at various times.

	COV - 90TH PERCENTILE					COV - 99TH PERCENTILE				
	1HR	2HR	3HR	8HR	24HR	1HR	2HR	3HR	8HR	24HR
MONITOR 01	.035	.036	.036	.034	.035	.021	.017	.01	.037	.085
MONITOR 02	.037	.038	.038	.039	.041	.04	.017	.042	.032	.04
MONITOR 03	.053	.054	.056	.057	.065	.009	.004	.023	.068	.059
MONITOR 04	.047	.044	.046	.048	.051	.02	.025	.034	.045	.065
MONITOR 05	.028	.029	.028	.027	.033	.025	.026	.034	.045	.032
MONITOR 06	.016	.017	.017	.014	.016	.015	.006	.008	.032	.04
MONITOR 07	.022	.021	.022	.023	.024	.126	.118	.13	.097	.069
MONITOR 08	.045	.045	.045	.047	.05	.097	.109	.096	.063	.056
MONITOR 09	.037	.037	.037	.038	.037	.034	.023	.026	.043	.053
MONITOR 10	.053	.053	.052	.051	.044	.019	.038	.034	.052	.064
MONITOR 11	.081	.085	.085	.087	.085	.005	.018	.01	.026	.035
MONITOR 12	.079	.08	.08	.083	.093	.056	.073	.071	.078	.083
MONITOR 13	.025	.028	.027	.029	.032	.024	.029	.037	.046	.026
MONITOR 14	.05	.049	.051	.055	.06	.037	.042	.045	.059	.062
MONITOR 15	.022	.024	.024	.023	.023	.051	.039	.044	.024	.065
MONITOR 16	.023	.024	.025	.022	.022	.039	.044	.05	.06	.078
MONITOR 17	.076	.076	.076	.074	.075	.074	.073	.079	.091	.106
MEAN.....	.043	.044	.044	.044	.046	.041	.041	.045	.053	.06

Table 5.22. Coefficients of variation for annual variations from logarithmic transformations of 90th and 99th concentrations.

5.4. Inter-Index Correlations

This section explores correlations between different measures of the same pollutant as well as between different pollutants. Even small differences in correlations create tremendous differences in the health effects association. In general, index correlations seem more dependent upon percentile groups than averaging times. Most correlations in this section are between four or five year means, rather than annual concentrations. This tends to improve correlations by damping out extreme values which can lead to negative correlations when only one year of data is examined (as seen in Section 5.1.1). This subject is explored further in Section 5.4.5 (Running versus Calendar Year Averages). This section describes correlations between the NAAQS, arrowhead indices, exceedence and dosage measures. These correlations are used to select representative and stable indices for use in the health effects analysis.

5.4.1. Arrowhead Indices and NAAQS

Table 5.23 shows correlations for 4 or 5 year means of all 11 sulfur and TSP indices. In general, short-term high percentile indices, e.g., 2, 3, 4, 9 and 10 correlated fairly closely (about 0.85 or more). Similarly, indices with long averaging times, e.g., 1, 6, and 11, correlate closely. Cleansing indices 7 and 8 do not correlate particularly well with any other index. Index 4, the short-term 99th percentile, is the most volatile measure. Index 4 for both sulfur and TSP has the lowest correlations (0.519 for sulfur and 0.363 for TSP) with cleansing index 8.

	ARG1	ARG2	ARG3	ARG4	ARG5	ARG6	ARG7	ARG8	ARG9	ARG10	ARG11
ARG1	1.000										
ARG2	0.968	1.000									
ARG3	0.975	0.999	1.000								
ARG4	0.918	0.980	0.976	1.000							
ARG5	0.894	0.855	0.859	0.772	1.000						
ARG6	0.717	0.565	0.582	0.476	0.651	1.000					
ARG7	0.450	0.305	0.330	0.261	0.316	0.605	1.000				
ARG8	0.656	0.542	0.564	0.500	0.498	0.611	0.846	1.000			
ARG9	0.858	0.722	0.740	0.607	0.800	0.807	0.475	0.643	1.000		
ARG10	0.853	0.710	0.729	0.596	0.785	0.815	0.536	0.690	0.996	1.000	
ARG11	0.850	0.723	0.742	0.626	0.754	0.770	0.574	0.728	0.942	0.951	1.000

Table 5.23a* Correlation matrix for all 11 sulfur arrowhead indices, including NAAQS.

	ARG1	ARG2	ARG3	ARG4	ARG5	ARG6	ARG7	ARG8	ARG9	ARG10	ARG11
ARG1	1.000										
ARG2	0.723	1.000									
ARG3	0.790	0.993	1.000								
ARG4	0.612	0.980	0.959	1.000							
ARG5	0.727	0.760	0.772	0.724	1.000						
ARG6	0.845	0.412	0.487	0.292	0.571	1.000					
ARG7	0.627	0.154	0.233	0.082	0.278	0.588	1.000				
ARG8	0.753	0.291	0.377	0.183	0.164	0.644	0.690	1.000			
ARG9	0.812	0.652	0.697	0.502	0.617	0.674	0.297	0.481	1.000		
ARG10	0.874	0.628	0.689	0.471	0.610	0.745	0.413	0.594	0.984	1.000	
ARG11	0.952	0.541	0.621	0.402	0.569	0.871	0.633	0.823	0.799	0.872	1.000

Table 5.23b.* Correlation Matrix for all 11 TSP arrowhead indices, including NAAQS.

* Arg = Index. Mean of four or five years of observations at the 17 monitors.

	Sulfur Indices				TSP Indices			
	4	5	6	8	4	5	6	8
Sulfur								
4	1							
5	0.934	1						
6	0.870	0.923						
8	0.711	0.697	1					
TSP								
4	0.060	-.020	0.047	0.034	1			
5	0.189	0.168	0.274	0.245	0.844	1		
6	0.429	0.403	0.525	0.490	0.592	0.850	1	
8	0.546	0.409	0.560	0.558	0.270	0.528	0.862	1

Table 5.24. Correlation coefficients between Sulfur and TSP for NAAQS and cleansing index 8. Four years of monitor data.

This is the lowest correlation found for TSP and the second lowest for sulfur.

Table 5.24 displays correlations of several indices, including the NAAQS for both pollutants. In general, correlations increase with averaging time, but still remain quite low, below about 0.5. Figure 5.14 and 5.15 graphically illustrate the lack of correlation between the two pollutants for 24 hour peaks and annual averages. The apparent randomness of the observations makes classifications into "dirty TSP and sulfur" and "clean TSP and sulfur" problematical, since there is no consistent pattern and many observations are excluded from this classification.

Table 5.25 shows the results of the cluster analysis, which was designed to maximize the variance between groups. Districts have been grouped into 6 partitions, ranked from 1 to 6 in terms of increasing concentrations. Districts show little consistency among different measures.

5.4.2 Exceedence Measures

As seen in the previous section, sulfur exceedences above 30 ppb are remarkably stable. This section shows that they are also insensitive to averaging time. Correlation coefficients for this measure between any averaging time from 1 hour and 336 hours are above 0.9. Correlations are above 0.78 for averaging times up to a year. Sulfur exceedences above 80 ppb showed similar behavior. Averaging time up to 168 hours were correlated over 0.9. Averaging times up to 3 months were correlated 0.735. Longer averaging time with this threshold resulted in lower

Figure 5.14. Scatterplot of TSP vs. Sulfur 24-hour peak concentrations. Four years of district data. (Sulfur index 5; TSP index 4).

Figure 5.15. Scatterplot of TSP vs. Sulfur annual (calendar year) concentrations. Four years of district data. (index 11).

	DIST NO.	SULFUR INDICES				TSP INDICES			
		4	5	6	8	4	5	6	8
DISTRICT: 1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
DISTRICT: 2.	1.	1.	2.	1.	2.	1.	1.	2.	1.
DISTRICT: 3.	2.	2.	3.	2.	5.	3.	5.	6.	4.
DISTRICT: 4.	1.	1.	1.	1.	1.	1.	1.	1.	1.
DISTRICT: 5.	1.	1.	1.	1.	1.	1.	1.	1.	1.
DISTRICT: 6.	3.	3.	2.	1.	5.	5.	5.	5.	1.
DISTRICT: 7.	1.	2.	4.	5.	5.	2.	3.	3.	3.
DISTRICT: 8.	1.	1.	4.	6.	5.	2.	3.	3.	3.
DISTRICT: 9.	1.	1.	3.	2.	2.	4.	6.	2.	2.
DISTRICT: 10.	1.	1.	3.	2.	2.	2.	3.	3.	3.
DISTRICT: 11.	1.	1.	3.	2.	2.	4.	6.	2.	2.
DISTRICT: 12.	1.	1.	3.	2.	2.	2.	3.	3.	3.
DISTRICT: 13.	4.	2.	2.	2.	5.	5.	5.	1.	1.
DISTRICT: 14.	4.	4.	3.	6.	3.	5.	5.	1.	1.
DISTRICT: 15.	1.	1.	3.	5.	4.	2.	3.	4.	4.
DISTRICT: 16.	2.	2.	1.	6.	5.	2.	3.	3.	3.
DISTRICT: 17.	1.	4.	4.	5.	3.	2.	3.	2.	2.
DISTRICT: 18.	1.	4.	4.	6.	3.	2.	3.	2.	2.
DISTRICT: 19.	2.	2.	1.	3.	4.	2.	6.	1.	1.
DISTRICT: 20.	1.	1.	3.	2.	2.	2.	3.	3.	3.
DISTRICT: 21.	3.	3.	1.	2.	3.	5.	5.	1.	1.
DISTRICT: 22.	3.	3.	1.	2.	3.	3.	6.	2.	2.
DISTRICT: 23.	1.	1.	2.	5.	4.	2.	3.	4.	4.
DISTRICT: 24.	2.	2.	1.	1.	6.	6.	4.	4.	4.
DISTRICT: 25.	2.	2.	1.	1.	6.	6.	4.	4.	4.
DISTRICT: 26.	2.	2.	1.	3.	3.	3.	6.	4.	4.
DISTRICT: 27.	2.	4.	4.	1.	5.	3.	5.	5.	5.
DISTRICT: 28.	4.	2.	1.	1.	3.	3.	6.	2.	2.
DISTRICT: 29.	6.	5.	6.	4.	3.	3.	6.	1.	1.
DISTRICT: 30.	2.	4.	3.	1.	5.	3.	5.	5.	5.
DISTRICT: 31.	5.	6.	5.	4.	3.	5.	2.	6.	6.
DISTRICT: 32.	6.	5.	6.	3.	3.	5.	5.	1.	1.
DISTRICT: 33.	5.	6.	5.	4.	3.	5.	2.	6.	6.
DISTRICT: 34.	6.	5.	2.	1.	3.	3.	6.	1.	1.
DISTRICT: 35.	2.	1.	4.	6.	4.	2.	3.	3.	3.
DISTRICT: 36.	3.	3.	3.	2.	3.	3.	6.	2.	2.
NO. ITERA.	-	1.	2.	5.	2.	1.	.	1.	2.
%BETWEEN GRP.									
VAR..	-	96.9	94.3	93.5	88.9	96.2	94.3	92.2	93.6

Table 5.25. Districts partitioned into 6 groups, ranked in order of increasingly high concentrations. using indices 4, 5 and 6 (the NAAQS). Results of cluster analysis which maximizes between group variance.

correlations due to the many zeros in the exceedence score. Sulfur exceedences above 150 ppm showed correlations above 0.9 for averaging times from 1 to 8 hours, and about 0.8 for 1 hour to 24 hours. Again, longer averaging times resulted in many zeros and thus lower correlations.

Table 5.26 shows correlations between the sulfur and TSP exceedence measures at several averaging times. There is a lack of correlation between the thresholds for either pollutant. Thus, exceedence measures at different thresholds vary enough to be considered separate indices.

Figure 5.16 plots the three (1 hour) sulfur exceedence measures against the exceedence measure at the lowest threshold (30 ppb). Figure 5.17 is a similar plot, showing the relation of the TSP exceedence measures. Exceedence measures at higher thresholds exhibit an increasing trend but with considerable fluctuations. For example, the first peak on the sulfur 150 ppb threshold line (points J, K) represents monitors 2 and 3 which are located near Homer City. These monitors experience fairly low annual concentrations, but frequent high peak concentrations. The two pronounced valleys on the 150 ppb threshold line represent monitors 4 and 1 (points G, J). Both monitors are somewhat isolated from the plume line of the Seward and Connemaugh power plants, because they are located on the downwind slope of the ridge. Monitor 5, which is located just several kilometers from monitor 4, but on the upwind side of the ridge slope, shows a similar measure for the low sulfur exceedence threshold as monitor 5 (two points to the right). However, it shows a much higher exceedence measure for a threshold of 150 ppb. (It is to the immediate right of point G on the 150 ppb line). Examinations of peak concentrations (indices 4 and 5) showed no such pattern, probably

	Sulfur					TSP			
	Exceedence			NAAQS		Exceedence		NAAQS	
	30	80	150	24hr	Annual	75	150	24hr	Annual
Sulfur									
30	1								
80	.719	1							
150	.782	.840	1						
24hr	.762	.914	.930	1					
Annual	.855	.933	.831	.923	1				
TSP									
75	.551	.580	.430	.532	.610	1			
150	.090	.164	-.044	-.123	.211	.524	1		
24hr	.125	.013	-.008	-.020	.047	.315	.717	1	
Annual	.506	.449	.286	.403	.525	.921	.709	.592	1

Table 5.26. Correlation matrix of exceedence measures and NAAQS.
24-hour 99th percentiles and running annual averages.

Figure 5.16. Relation of the sulfur exceedence measures. Mean of 4 year monitor exceedence measures for various thresholds (on a log scale) plotted against low (30 ppb) exceedence measures.

Figure 5.17. Relation of the TSP exceedence measures. Mean of 4 year monitor exceedence measures for two thresholds (on a log scale) plotted against low (75 u.gms/cu.m.) exceedence measures.

Figure 5.18. * Sulfur exceedences above 80 ppb versus calendar year mean concentrations.

Figure 5.19.* Sulfur exceedences above 150 ppb versus calendar year means.
* Mean monitor scores are plotted. Regression line is shown in asterisks.

Figure 5.20.* TSP exceedences above 75 u.gms/cu.m versus calendar year mean concentrations.

Figure 5.21. TSP exceedences above 150 u.gms/cu.m vs. calendar year mean.
* Mean monitor scores are plotted. Regression line is shown in asterisks.

because the stochastic influence of the weather plays a large role in determining peak concentrations. Exceedence measures, however, provide information concerning the frequency of peaks which may be more relevant than the concentration of the highest or second highest peak.

Correlations between exceedence measures (at the shortest, most stable averaging time), and NAAQS are shown in Table 5.26. In general, low threshold exceedences correlate better to annual concentrations, and high threshold measures to peak concentrations, as expected.

Figures 5.18 to 5.21 plot exceedence measures against the annual concentrations (index 11). In most instances, the exceedence measures track annual concentrations quite closely. (The scatter in the high TSP threshold probably results from the lack of observations.) Interestingly, exceedence measures at high thresholds, which were thought to be peak measures and thus approximately lognormally distributed, have been linearized. This may help explain their insensitivity to different averaging times.

Equations for the linear regression lines plotted in Figures 5.18 to 5.21 are shown below, where C = annual concentration:

$$C_{\text{sulfur}} = 15.11 (\% \text{ exceedence over } 80\text{ppb}) - 5.02 \quad (19)$$

$$C_{\text{sulfur}} = 8.89 (\% \text{ exceedence over } 150\text{ppb}) - 3.83 \quad (20)$$

$$C_{\text{TSP}} = 79.46 (\% \text{ exceedence over } 75 \text{ u.gms}) - 33.81 \quad (21)$$

$$C_{\text{TSP}} = 12.49 (\% \text{ exceedence over } 150 \text{ ugms}) - 7.91 \quad (22)$$

The equations will be used to compare health predictions using the exceedence measures with predictions using average annual concentrations.

5.4.3. 90th and 99th Percentiles

Modelers and regulators are very interested in using a more stable measure to predict and assess air pollution from new sources, and have suggested the use of concentrations at the 90th percentile. Thus, the stability of the 90th percentile concentrations were investigated in depth, and compared to the 99th percentile, which is the current NAAQS. Sulfur concentrations at averaging times of 1, 2, 3, 8, and 24 hours were examined for both percentiles. Table 5.27 shows correlation coefficients which use logarithms of concentration values. All 90th percentile values correlated very closely: the lowest correlation is 0.986, between 1 and 24 hour averaging times. Correlations among the 99th percentile were considerably lower, with the lowest coefficient of 0.691, again between 1 and 24 hours. Correlation coefficients between the percentile groups are only 0.593 for 1 hour, and 0.802 for 24 hours. This suggests that the two percentile groups are significantly different measures of pollution. This is shown in other tests, which follow.

Table 28 shows other differences between the percentile groups. The range of the 99th percentile is a magnitude larger than the 90th percentile's, while mean values are about 5 times greater. 99th percentile concentrations may tend to exaggerate more subtle differences that exist in 90th and lower percentile values, but not in a consistent manner. As expected, mean concentrations and ranges tend to decrease as averaging time is increased. The change is especially pronounced for 99th percentile values. However, as seen earlier in Table 5.22, COVs for the logarithms of the values remained constant.

One and two way analysis of variance (ANOVA) test were used to further investigate the comparative stability of the percentile groups. Results are shown in Table 5.29. Logarithms of concentration values were used in all cases. The two way ANOVA, which accounts for differences among monitors as well as years, is the more stringent test to check for year to year variation. Surprisingly, only the 99th percentile showed no significant differences ($p < 0.05$) across the 4 years of data, for averaging times up to 8 hours. This suggests that the variance is too large in 99th percentile concentrations for averaging times of less than 8 hours to observe yearly differences. Thus, the 99th percentile concentrations did not show significant changes in pollution levels, even after the Seward stack was raised. 90th percentile measures do not show this behavior. These concentrations did show a significantly change over time. One way ANOVA results showed mean concentrations for the 17 monitors at 90th percentile varied significantly on a yearly basis. However, 99th concentrations were not significantly different until an averaging time of 24 hours. This suggests that if 99th percentile concentrations are used, a minimum averaging time of 24 hours is necessary to portray changes in pollution patterns which occur over time. However, 90th percentile values seem to capture the same information, regardless of the averaging time.

In summary, logarithms of 90th and 99th percentile concentrations are far more stable measures of air pollution than their untransformed values. 99th percentile values at short averaging times show a great deal of random variation. However, 90th percentile concentrations at any averaging time seem to give a better indication of changes in air pollution.

	90th Percentile			99th Percentile		
	<u>1</u>	<u>8</u>	<u>24</u>	<u>1</u>	<u>8</u>	<u>24</u>
90th percentile						
1 hour	1					
8 hours	0.997	1				
24 hours	0.986	0.991	1			
99th Percentile						
1 hour	0.593	0.606	0.595	1		
8 hours	0.702	0.709	0.714	0.866	1	
24 hours	0.787	0.792	0.802	0.691	0.875	1

Table 5.27. Correlations with 90th and 99th percentile concentrations at 1, 8, and 24 hours averaging time, using 4 years of data from 17 monitors.

	Averaging Time in Hours					Mean
	<u>1</u>	<u>2</u>	<u>3</u>	<u>8</u>	<u>24</u>	
90th percentile						
Range (u.gms/cu.m)	90	92	92	88	73	89.1
Mean (u.gms/cu.m)	55	55	54	52	47	53.0
99th percentile						
Range (u.gms/cu.m)	858	663	635	527	250	587
Mean (u.gms/cu.m)	263	237	210	156	107	195

Table 5.28. Mean and range of 90th and 99th percentile sulfur concentrations. 4 year mean of monitor data.

	Averaging Time in Hours				
	1	2	3	8	24
TWO WAY ANOVA					
90th percentile					
F (years)	8.8	8.7	9.0	8.8	9.3
Probability	0	0	0	0	0
F (monitors)	5.0	5.0	5.2	5.1	4.6
probability	0	0	0	0	0
99th Percentile					
F (years)	.72	.34	.30	1.2	7.0
probability	.55	.80	.83	.31	.001
F (monitors)	5.8	5.3	4.5	4.2	2.7
probability	0	0	0	0	0
ONE WAY ANOVA					
90th percentile					
F (years)	4.4	4.3	4.4	4.3	4.9
Probability	.008	.007	.008	.007	.004
99th percentile					
F (years)	.32	.16	.16	.69	4.9
Probability	-	-	-	-	0.004

d.f. for years = 3.

d.f. for monitors = 16.

Table 5.29. One and two way analysis of variance for 90th and 99th percentile sulfur concentrations over 4 years at the 17 monitors.

	Sulfur Measures				TSP Measures				Combined
	Threshold		PSI	Index 6	Threshold		PSI	Index 6	PSI dosage
	normal	lognormal			normal	lognormal			
SULFUR									
normal	1								
log	.719	1							
PSI	.782	.840	1						
Idx 6	.855	.933	.831	1					
TSP									
norm	.810	.669	.659	.822	1				
log	.551	.580	.430	.610	.538	1			
PSI	.090	.164	-.04	.211	.181	.524	1		
Idx 6	.506	.449	.286	.525	.490	.921	.709	1	
COMBINED									
PSI	.499	.540	.435	.560	.558	.965	.475	.862	1

Table 5.30. Correlation matrix for the dosage measures, including threshold dosages based on normal and log-normal distributions, and the Pollution Standards Index (PSI) as a non-linear dosage measure. Average concentration, index 6 is shown as reference. Correlations based on 4 year means of monitor scores.

Year	Correlations between annual TSP and Sulfur Concentrations	
	Running averages (Index 6)	Calendar averages (Index 11)
1975	0.368	0.551
1976	0.317	0.358
1977	0.450	0.342
1978	0.239	0.108
All years	0.525	0.551

Table 5.31. Correlations between sulfur and TSP, contrasting running versus calendar year averaging, using monitor data.

5.4.4 Dosage Measures

Table 5.30 shows correlations of the several dosage measures and annual average concentrations. In general, TSP measures correlate more closely to one another than the sulfur dosage measures. Sulfur measures vary considerably from the annual average, with correlation coefficients between 0.75 and 0.81. For both TSP and sulfur the log-normal distribution of the threshold dose correlates slightly less than the normal distribution of the dosage measure. The combined non-linear (or PSI) dose, that is, the sum of the PSI and sulfur measures, correlates highly with both TSP and sulfur mean concentrations, as expected.

5.4.5 Running versus Calendar Year Means

Table 5.31 shows correlations between sulfur and TSP measures for running and calendar year averages for the years 1975 to 1978. Higher correlations with calendar year averages as compared to running averages would tend to indicate that high TSP and sulfur concentrations are experienced in the same yearly period. Higher correlations occurred in 1975 and 1976. However, in 1977 and 1978, the opposite occurred. This may reflect a lack of simultaneity. The phenomenon is quite inconclusive.

5.5 Health Effects

The simple screening model is used to test the association of pollution measures with the health data from the children's school study.

Then, stable pollution measures were used in the multiple logistic regression model, which controlled potential confounding and other risk factors. Screening models also were used to evaluate chronological effects, pollutant interaction, and some initial weightings for the certainty of pollutant characterization. The last part of this section is a simple risk-analysis of some of the potential health effects posed by a proposed coal power plant at the Seward site which uses the results of the regression health model.

The screening model, which is used extensively in this section, does not control confounding and risk factors. High correlations with this model do not imply that the pollution measure is clinically relevant unless one makes the assumption that the health effects are related only to pollutant levels. With this assumption, either the screening or the regression model may provide a means to compare the various pollution measures. However, the prevalence of respiratory symptoms is known to be influenced by other factors. Therefore, the regression model provides the superior test.

5.5.1 Selection of Screening Model

Four screening models were evaluated. All models were correlations of health scores at the 14 schools (shown earlier in Table 4.2) with pollution scores. The first two models used either prevalence rates or the logit transformation of these rates, with school pollution scores. The third model used averaged school and township pollution scores. It showed lower correlations between pollution and health scores than

the model which used only the school scores. This formulation assumes that children live in the township where the school is located. However, these are regional schools: children are bused from several townships. This, perhaps, was reflected in the lower correlations. The fourth screening model weighted the correlations by the population at each school, and omitted school number 13, which contained only 1st and 2nd graders, from the correlation. This approach was thought to decrease the influence of risk and confounding factors, which were likely to be randomly distributed with respect to school size. This model resulted in extremely high correlations. Returning the one school to the correlation increased correlations even further. This suggests that school size is related to symptom prevalence. Correlations between health effects, population size, and pollution scores were not consistent, and did not fully explain the results. This approach was abandoned.

Thus, the first two models were used to screen the pollution measures. These models may not be very good indicators for wheeze and chronic phlegm since there is little variation (COV about 0.2) for school averages of these symptoms.

5.5.2 Correlations with Mean Pollution Scores

Table 5.32 shows correlations between the 11 arrowhead indices and prevalence rates and logit transformations of these rates. One quickly notes that these correlations are not impressively high. A visual interpretation of one of the correlations is seen in Figure 5.22. Annual sulfur concentrations (index 6) are plotted against the prevalence of

serious chest illness before age 2. There is only a very general positive trend seen in the plot, which has a correlation coefficient of 0.573. However, as will be seen in Section 5.5.4, serious chest illness before age two is the only symptom for which a significant and consistent association with air pollution was found.

Sulfur correlations tend to be considerably higher than those for TSP. Correlations for wheeze are almost always zero or negative, which suggests that risk factors other than ambient air pollution are the determining causes for these symptoms. Correlations for recent chest illness (in the past year) tend to be quite high. These correlations are highly influenced by one extreme observation. School number 13, which is located in the dirtiest district and contains only 80 children, has a prevalence rate for recent chest illness of over twice the average. Thus, correlations for this symptom may be either invalid or truly a function of air pollution. The logit model showed results similar to these correlations, as seen in Section 5.5.4.

Of the 11 sulfur indices, numbers 6, 9 and 11 show the highest correlations with health effects. For TSP, cleansing indices 7 and 8 have the highest correlations (excepting recent chest illness). This indicates that it is the schools with lower TSP concentrations that experience the highest rate of respiratory symptoms.

Correlations using the prevalence rate were generally slightly higher than those which used the logit transformation. However, due to the similarity between the two correlations, results from the latter model are not shown for correlations using other pollution measures.

Table 5.33 shows correlations with logarithms of concentrations of

Figure 5.22. Prevalence rates for serious chest illness by age two vs. mean annual sulfur concentrations. Correlation coefficient of 0.573.

Figure 5.23. Prevalence rates for serious chest illness by age two vs. mean TSP exceedences above 75 ug.m. Correlation coefficient of 0.573.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.406	.398	.395	.366	.405	.394	.09	.163	.427	.408	.361
PHLEGM	.61	.628	.625	.633	.532	.476	.418	.344	.433	.447	.484
WHEEZE	-.055	-.051	-.044	-.038	-.021	-.059	.011	.124	-.13	-.092	-.071
CI-3DY	.707	.627	.628	.579	.8	.858	.21	.425	.901	.871	.814
CI-1YR	.493	.481	.474	.493	.517	.573	.13	.15	.535	.498	.476

Table 5.32a. Correlation coefficients between prevalence rates and 4 year means of sulfur scores, including arrowhead indices and NAAQS. CI-3dy is 3 days of chest illness in the past 2 years; CI-1yr is serious chest illness by age 2.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.356	.337	.337	.301	.356	.365	.099	.197	.394	.383	.34
PHLEGM	.519	.545	.542	.552	.422	.374	.368	.258	.333	.347	.381
WHEEZE	-.052	-.057	-.049	-.051	-	-.032	-.004	.151	-.096	-.059	-.047
CI-3DY	.638	.547	.551	.491	.755	.824	.139	.4	.875	.845	.774
CI-1YR	.412	.402	.396	.422	.436	.501	.111	.131	.445	.414	.401

Table 5.32b. Correlations between logit transformation of prevalence rates and 4 year means of sulfur scores.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.32	-.097	.031	-.321	.006	.416	.403	.521	.378	.42	.393
PHLEGM	.324	-.156	-.014	-.353	-.109	.397	.615	.559	.309	.376	.337
WHEEZE	.022	.062	.058	.043	-.06	.024	.04	.012	-.029	-.017	.014
CI-3DY	.254	-.049	.035	-.181	.068	.33	.313	.415	.261	.288	.338
CI-1YR	.264	.008	.074	-.109	.039	.297	.401	.443	.275	.286	.286

Table 5.32c. Correlations between prevalence rates and 4 year means of TSP scores, including arrowhead indices and NAAQS.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.358	-.031	.092	-.257	.06	.447	.4	.527	.414	.452	.431
PHLEGM	.269	-.211	-.072	-.401	-.161	.342	.564	.511	.26	.328	.279
WHEEZE	.053	.096	.093	.066	-.025	.058	.04	.036	.008	.017	.055
CI-3DY	.208	-.07	.007	-.196	.046	.288	.234	.364	.22	.245	.305
CI-1YR	.257	.055	.106	-.047	.053	.275	.38	.411	.261	.265	.268

Table 5.32d. Correlations between logit transformation of prevalence rates and 4 year means of TSP scores.

	90th Percentile		99th Percentile	
	1 hour	24 hour	1 hour	24 hour
Cough	0.475	0.440	0.300	0.322
Phlegm	0.407	0.387	0.630	0.406
Wheeze	-.032	-.001	0.020	0.138
CI-3dy	0.830	0.833	0.494	0.679
CI-1yr	0.412	0.461	0.464	0.356

Table 5.33. Correlations between health effects and logarithms of 90th and 99th percentile sulfur concentrations at two averaging times.

	Sulfur Exceedence		TSP Exceedence	
	30 ppb	80 ppb	75 ugms	150 ugms
Cough	0.416	0.405	0.556	0.014
Phlegm	0.437	0.520	0.551	0.121
Wheeze	0.176	-.253	0.001	0.085
CI-3dy	0.509	0.906	0.403	0.167
CI-1yr	0.166	0.642	0.373	0.191

Table 5.34. Correlations between health effects and exceedence measures at two thresholds for TSP and sulfur.

	<u>Sulfur</u>			<u>TSP</u>			Combined
	<u>Threshold dose</u>		<u>PSI dose</u>	<u>Threshold dose</u>		<u>PSI dose</u>	PSI dose
	<u>normal</u>	<u>lognormal</u>		<u>normal</u>	<u>lognormal</u>		
Cough	0.426	0.379	-.003	0.541	0.479	0.489	0.337
Phlegm	0.549	0.587	0.406	0.546	0.512	0.491	0.601
Wheeze	0.033	-.014	0.060	0.008	-.027	0.049	-.003
CI-3dy	0.709	0.733	0.508	0.404	0.397	0.278	0.515
CI-1yr	0.422	0.532	0.522	0.370	0.348	0.267	0.516

Table 5.35. Correlations between health effects and dosage measures. Threshold dosages above annual NAAQS using normal or log-normal distributions; and non-linear dosage measure based on PSI.

the 90th and 99th sulfur concentrations at two averaging times. 90th percentile correlations are some 20 to 30% higher than the 99th percentile's. Actual concentrations of the 99th percentile, seen in Table 5.30a as index 5, have higher correlations than do logarithms of the values. This may result because actual values, while unstable, tend to emphasize small differences which may be found between school locations.

Table 5.34 show correlations with the exceedence measures. Sulfur exceedences above 80 ppb resulted in the highest correlations observed with the screening model. Exceedences above 75 micrograms also resulted in the highest correlations for TSP. Figure 5.23 plots the prevalence of serious chest illness with this measure, showing a scattered but positive trend. These preliminary results suggest that exceedence measures may be the most clinically relevant measure among those tested. High correlations with the health effects may indicate, with the caveats expressed earlier, that pollution scores which differentiate concentrations above a threshold may be the most useful in health models. This may be due to the difficulty of observing differences at low pollution levels, i.e., low percentiles, or because a threshold dose-response function exists.

Table 5.35 shows correlations with the dosage measures. For sulfur, threshold dosage based on a logarithmic distribution of pollution concentrations showed higher correlations. The normal distribution gave higher correlations with TSP. This tends to support the evidence that areas with high TSP concentrations are in fact the cleaner areas because the normal distribution of this measure gives less weight to periods of high exposure. The threshold dosages give results similar to, but

slightly lower than, indices 5 and 6. The non-linear dosage measures, which gave more weight to lower percentiles (based on the Pollution Standards Index), had lower correlations than other measures, in most cases. The combined sulfur and TSP non-linear dosage measure gave correlations higher than the TSP measure alone, as expected, but not significantly lower than the sulfur measure. However, none of these correlations are particularly high. Thus, it seems doubtful that these particular dosage measures captured significantly more information than annual averages or exceedence measures. However, the evidence tends to support the hypothesis that high percentile groups are important, with respect to health effects. Thus, both a median and upper percentile measure, perhaps the 84th or 90th (but not necessarily a peak or 99th percentile measure), seem the minimum data necessary to characterize air pollution in epidemiologic investigations.

Table 5.36 displays results for the mean of the 11 normalized arrowhead indices. Correlations for sulfur are nearly the same as for actual means, as seen earlier in Table 5.32. Correlations for TSP have increased markedly. This may suggest that the prevalence rates of respiratory symptoms are related to schools which consistently experience high TSP levels in relation to the other schools. Thus, prevalence rates may not be related as well to the arithmetic mean of the pollution level, which may be thrown off by occasional or atypical high peaks or high means, possibly caused by agricultural or roadway dust. Therefore, average TSP measures may not reflect the risk posed by either harmful energy related pollutants or more constant but moderate TSP levels. Firmer evidence for these hypotheses could be provided by better TSP

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.367	.322	.323	.259	.397	.347	.155	.215	.327	.29	.27
PHLEGM	.587	.585	.586	.577	.522	.455	.449	.336	.388	.391	.441
WHEEZE	-.104	-.163	-.152	-.179	.026	-.096	.027	.146	-.174	-.136	-.09
CI-3DY	.782	.771	.768	.715	.773	.861	.26	.44	.871	.831	.76
CI-1YR	.595	.677	.661	.718	.526	.591	.106	.096	.54	.491	.441

Table 5.36a. Correlations between prevalence rates and 4 year means of normalized sulfur scores, including arrowhead indices and NAAQS.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.429	.326	.361	.158	.378	.436	.309	.428	.434	.432	.385
PHLEGM	.566	.6	.62	.518	.423	.486	.588	.511	.413	.453	.448
WHEEZE	-.011	-.035	-.014	-.102	.006	-.033	.042	.089	-.108	-.079	-.033
CI-3DY	.582	.631	.623	.548	.571	.664	.316	.484	.7	.664	.619
CI-1YR	.496	.639	.615	.65	.439	.499	.275	.321	.498	.46	.409

Table 5.36b. Correlations between prevalence rates and 4 year means of normalized TSP scores, including arrowhead indices and NAAQS.

CORRELATION COEFFICIENTS FOR INDEX											
	1	2	3	4	5	6	7	8	9	10	11
COUGH.	.418	.264	.311	.078	.369	.436	.315	.456	.441	.441	.399
PHLEGM	.478	.504	.525	.427	.326	.395	.528	.436	.322	.364	.357
WHEEZE	.007	-.046	-.02	-.136	.034	-.001	.039	.125	-.068	-.042	.001
CI-3DY	.511	.53	.526	.428	.513	.616	.24	.448	.658	.619	.573
CI-1YR	.448	.586	.565	.607	.394	.447	.248	.289	.441	.405	.362

Table 5.37. Correlations between prevalence rates and 4 year means of combined normalized sulfur and TSP scores, i.e., sum of normalized sulfur and TSP pollution indices.

measures, perhaps by measures of respirable particulates.

Table 5.37 shows correlations of combined normalized TSP and sulfur indices. Most scores are below 0.3; thus, associations between TSP or sulfur and health effects are cancelled out. Other combinations of the two pollutants, e.g., the product of the annual means, generally gave similarly poor results. This results either from a lack of synergism between the pollutants, erroneous or clinically irrelevant pollution measures, or different areas which experience TSP and sulfur levels. The latter possibility fits the Chestnut Ridge area; the other hypotheses require more sensitive tests.

5.5.3 Chronological Effects

Table 5.37 shows the effects of chronological weightings of the NAAQS with selected health indicators. The change in correlations for different individual years is particularly striking. For example, correlations for serious chest illness is a very high 0.726 in 1975 for 8-hour 99th percentile concentration (index 4), but only -0.230 in 1978. In most cases, mean scores give the highest correlations. This may be evidence of the long incubation periods for these symptoms and/or of the difficulty of characterizing pollution using just one year's data.

Interestingly, correlations for annual means of sulfur and TSP (index 6), with increasing and decreasing chronological weights, follow the same trend. For example, for serious chest illness, increasing the weighting or influence of more recent years gives correlations with health effects above the mean. Correlations of decreasing weighting to later years

	Sulfur Indices			TSP Indices		
	4	5	6	4	5	6
1974-5	0.240	0.260	0.306	-.269	0.077	0.271
1978	0.278	0.199	-.170	-.242	0.054	0.216
Mean	0.336	0.405	0.396	0.321	0.006	0.416
Decreasing	0.298	0.351	0.439	-.365	0.060	0.438
Increasing	0.339	0.410	0.320	-.270	-.025	0.383

Table 5.38a.* Chronological Effects of Correlations for Chronic Cough.

	Sulfur Indices			TSP Indices		
	4	5	6	4	5	6
1974-5	0.518	0.412	0.245	0.058	0.076	0.194
1978	0.299	0.271	0.120	-.355	-.244	0.026
Mean	0.632	0.531	0.482	-.353	-.109	0.397
Decreasing	0.582	0.453	0.474	-.159	0.005	0.455
Increasing	0.518	0.551	0.460	-.399	-.175	0.322

Table 5.38b.* Chronological Effects of Correlations for Chronic Phlegm.

	Sulfur Indices			TSP Indices		
	4	5	6	4	5	6
1974-5	0.726	0.575	0.178	0.092	-.196	0.153
1978	0.230	-.245	0.503	-.246	-.132	0.273
Mean	0.504	0.520	0.544	-.109	0.039	0.297
Decreasing	0.725	0.583	0.453	0.056	0.115	0.243
Increasing	0.173	0.323	0.626	-.168	-.006	0.337

Table 5.38c.* Chronological Effects of Correlations for Serious Chest Illness before age two.

* Chronological effects of NAAQS pollution scores on health screening model's correlations. Health data is untransformed. 1974 for TSP; 1975 for sulfur. Decreasing represents decreasing weighting of later years; increasing is the opposite.

	<u>Sulfur Exceedence Threshold</u>		<u>TSP Exceedence Threshold</u>	
	<u>30</u>	<u>80</u>	<u>75</u>	<u>150</u>
1974-5	0.504	0.450	0.450	-.041
1978	-.052	0.120	0.491	-.222
Mean	0.416	0.405	0.556	0.014
Decreasing	0.499	0.426	0.561	0.219
Increasing	0.193	0.367	0.544	-.104

Table 5.39a.* Chronological Effects of Correlations for Chronic Cough.

	<u>Sulfur Exceedence Threshold</u>		<u>TSP Exceedence Threshold</u>	
	<u>30</u>	<u>80</u>	<u>75</u>	<u>150</u>
1974-5	0.455	0.593	0.409	0.121
1978	0.196	0.317	0.296	-.305
Mean	0.437	0.520	0.551	0.121
Decreasing	0.450	0.532	0.595	0.511
Increasing	0.348	0.494	0.502	-.122

Table 5.39b.* Chronological Effects of Correlations for Chronic Phlegm.

	<u>Sulfur Exceedence Threshold</u>		<u>TSP Exceedence Threshold</u>	
	<u>30</u>	<u>80</u>	<u>75</u>	<u>150</u>
1974-5	0.092	0.655	0.518	0.068
1978	0.053	-.017	0.060	-.195
Mean	0.166	0.642	0.373	0.191
Decreasing	0.131	0.657	0.417	0.428
Increasing	0.211	0.608	0.325	0.023

Table 5.39c.* Chronological Effects of Correlations for Serious Chest Illness before age two.

* Chronological effects of sulfur and TSP exceedence measures on health screening model's correlations. Health data is untransformed. 1974 for TSP; 1975 for sulfur. Decreasing represents decreasing weighting of later years; increasing is the opposite.

	<u>1975</u>	<u>1976</u>	<u>1977</u>	<u>1978</u>	<u>Mean*</u>
Cough	0.305	0.510	0.424	-.170	0.268
Phlegm	0.237	0.595	0.446	0.120	0.251
Wheeze	-.036	0.059	0.050	-.369	0.167
CI-3dy	0.559	0.821	0.740	0.532	0.663
CI-1yr	0.178	0.610	0.548	0.489	0.445

Table 5.40a. Correlations with health effects using running annual average of sulfur concentrations (index 6).

* Mean of 4 years of correlations.

	<u>1975</u>	<u>1976</u>	<u>1977</u>	<u>1978</u>	<u>Mean*</u>
Cough	0.449	0.534	-.245	0.048	0.197
Phlegm	0.562	0.445	0.064	0.295	0.342
Wheeze	-.600	0.038	-.268	0.049	-.045
CI-3dy	0.722	0.809	0.475	0.346	0.588
CI-1yr	0.405	0.360	0.493	0.107	0.366

Table 5.40b. Correlations with health effects using calendar year average of sulfur concentrations (index 11).

* Mean of 4 years of correlations.

years give correlations below the mean. Strangely, the effect is opposite but smaller for the symptoms of cough and phlegm. This may demonstrate effects due to large scale weather related changes in pollutant dispersion, which have the same effect on TSP and sulfur. Also, 1974 and 1975 peak measures (indices 4 and 5) correlated highly with serious chest illness before age 2. These pollution scores date to the time these children were that age.

Table 5.39 shows chronological effects for the exceedence measures. Results are similar to those for the NAAQS. The table demonstrates the instability of the exceedence measure above 150 micrograms.

Table 5.40 shows correlations for individual years of running and calendar year averages of sulfur concentrations. With the exception of correlations for chronic phlegm, correlations using the running averages are significantly higher and more consistent than those which use calendar year averages. This may tend to confirm the intuitive notion that running averages are more appropriate measures of pollution for health effects research and analysis.

5.5.4. Logit Model Results

On the basis of pollutant index stability and consistency, the Harvard researchers selected two sulfur pollution measure for use in the logistic regression model: mean running annual averages and exceedences over 80 ppb. One TSP measure was selected: exceedences over 75 micrograms/cubic meter. Scores for school locations (those used in the screening model) and residence locations, i.e., townships, were used

separately. These six pollution scores were used for four classes of health effects. Chronic cough and phlegm were combined into one indicator. The model was first optimized without the pollution scores with confounding and risk factors, which include age and sex of the child, sex and social class of the parent completing the questionnaire, parents' allergic and respiratory history, and parental smoking. Pollution scores were then added to the model. Coefficients and probabilities are shown in Table 5.41. Regression coefficients for annual averages and exceedences are easily compared. Sulfur exceedence scores, and thus coefficients, may be multiplied by 15.11 to obtain approximate equivalents in terms of annual concentrations, as seen in Section 5.4.3. In all cases, coefficients are comparable and usually very close.

There are only six results which are significant at about the 0.05 confidence level. Of these, one is invalid. Persistent wheeze was significant at the 0.06 level, but has the wrong sign (also seen in the screening model results). Recent chest illness is related to sulfur levels at the school at the highly significant 0.002 level. The results indicate a strong relationship in terms of the confidence level and the high magnitude of the coefficient. However, the coefficients are so large, in fact, that these results are questionable. Moreover, results using pollution scores at the residence locations had a statistically insignificant association, and much smaller coefficients. Recent chest illness may be an inappropriate health measure for children since outbreaks of infectious disease are common, and quickly spread at schools. Thus, these results may indicate a merely spurious association.

	Township Pollution Scores			School Pollution Scores		
	annual	exceedences		annual	exceedences	
	<u>sulfur</u>	<u>sulfur</u>	<u>TSP</u>	<u>sulfur</u>	<u>sulfur</u>	<u>TSP</u>
Chronic Cough or Phlegm						
coefficients	-.060	-.0069	-.0016	0.29	0.030	0.0075
standard error	0.47	0.033	0.0079	0.72	0.061	0.0071
probability	-	-	-	-	-	-
Persistent Wheeze						
coefficients	-.72	-.069	0.0039	0.29	-.037	0.0047
standard error	0.51	0.038	0.0083	0.75	0.069	0.0074
probability	0.15	0.06	-	-	-	-
Recent Chronic Illness						
coefficients	0.32	0.020	0.0043	2.3	0.18	0.0014
standard error	0.39	0.024	0.0067	0.58	0.045	0.0060
probability	-	-	-	0.002	0.002	-
Serious Chest Illness by Age 2						
coefficients	0.98	0.064	0.014	0.37	0.075	0.0094
standard error	0.44	0.030	0.0078	0.72	0.056	0.0071
probability	0.03	0.04	0.08	-	0.20	0.19

Table 5.41. Result of logistic regression health model. Coefficients, standard error and probabilities for pollution variables in the 24 models, i.e., six pollution measures for each of 4 health indicators.

	No Passive Smoke		Passive Smoke	
	Relative		Relative	
	<u>Risk</u>	<u>Risk</u>	<u>Risk</u>	<u>Risk</u>
Low Pollution 22 ppb (58 ug)	0.146	1.00	0.223	1.51
High Pollution 37 ppb (100 ug)	0.223	1.52	0.320	2.19

Table 5.42. Risk and Relative risk for children, either with or without exposure to cigarette smoke. [Communication from M.S. Schenker, 1981].

In addition, they emphasize the importance of the pollutant characterization. The use of different geographic units produced startling different results.

Sulfur measures did show a significant association with serious chest illness before age 2. This was the only model where TSP results were of significance, albeit borderline, but with a strong trend. The magnitude of the effect may be seen in Table 5.42. In a low sulfur pollution area (annual average of 22 ppb), the risk of serious chest illness is 14.6%, while in a high pollution area (37 ppb sulfur), the risk is 22.5%. The relative risk for a child in a high pollution area in this study is just slightly greater than the risk posed by passive smoke exposure, i.e., if the child's parents smoke cigarettes. Lastly, these risks are multiplicative: a child that lives in a high pollution area has a 32% chance of developing serious chest illness if his or her parents smoke.

5.5.6 Uncertainty in Pollution Characterization

Results of the logit model varied greatly depending whether school or township scores were used. This suggests that either school exposures are not equivalent to residence exposure, planer estimates for schools or townships caused significant errors, or the association between pollution and health effects is spurious or not related to the parameters of the case study. The first possibility seems counterintuitive. A different estimation scheme for school locations, possibly taking into account the direction of the prevailing wind or combining school and residence pollution scores, might reduce the difference between school and

residence results.

Section 5.5 attempted to select measures of pollution which appeared most relevant to respiratory health effects. Undoubtedly much has been missed. Effects of other pollutants, for which sulfur dioxide and TSP are not good surrogates, including ozone and nitrogen dioxide, have been missed, probably entirely. Furthermore, elimination of pollution measures was accomplished using a very approximate screening model. Several pollutant measures were very crude approximations. Lastly, pollutant characterizations which have been missed by this research may prove to be clinically more relevant than any of those tested here.

Sections 5.2 and 5.4 demonstrated that pollutant levels changed greatly from year to year, with a corresponding change in the strength of the health effects association. In general, the average of four or five years of pollutant data is found to be the most representative measure, as well as the best health predictor. Using less data, even in the case of stable pollutant measures like the average annual concentration, caused considerable variation and error. This study could not directly check the cumulative effects of exposure to air pollution over many years. However, the uncertainty in the characterization of pollutant scores over several years can be assessed by an examination of the variance over the available data. Large coefficients of variance (COV) indicate unstable locations, which could be weighted less to reflect this uncertainty. However, if this instability occurs in a very dirty or very clean location, it should not be penalized. That is, if even with the extreme variations, the area is still cleaner (or dirtier) than the other areas in the (cross-sectional) study, it should not be discounted.

	<u>Sulfur Indexs</u>		<u>TSP Index</u>	
	<u>6</u>	<u>9</u>	<u>6</u>	<u>9</u>
Cough	0.954	0.974	0.639	0.551
Phlegm	0.962	0.978	0.672	0.804
Wheeze	0.862	0.894	0.368	0.785
CI.3dy	0.947	0.961	0.644	0.739
CI-1yr	0.967	0.941	0.795	0.789

Table 5.43 Correlations between indices 6 and 9 using weights based on the annual variance of the pollutant characterization.

	<u>Present Air Quality</u>	<u>Air Quality with Proposed Plant</u>
Sulfur		
annual average	73 ugms/cu.m.	78 u.gms/cu.m
24 hour highest peak	276	321
8 hour highest peak	1010	1179
TSP		
annual average	61	61
24 hour highest peak	117	118

Table 5.44. Air quality predictions from proposed coal plant in the Chestnut Ridge area [Army Corps of Engineers, 1981].

Appendix D describes a simple way to weight locations knowing the COV and mean of all geographic units in the study. These weights have been developed for the 14 schools, for running annual average concentrations (index 6) and arrowhead index 9, a high percentile, short averaging time measure, for both TSP and sulfur. (Weights are shown in Appendix D.) Correlations with health effects using these weights are shown in Table 5.43. Correlation coefficients with sulfur are very high. Correlations with TSP have also improved. Inaccuracies and limitations in the simple screening model may cause these effects. However, no systematic association between health effects, pollution scores, and weights was noted. This approach may have promise. Experimentation and validation is needed in several areas: 1) Proper measures of uncertainty are required, which may not be the COV of the pollutant score used in the health model. For example, correlations between the two sulfur weights was only 0.338; between the TSP weights, only 0.218. Uncertainty scores should be as stable as the pollutant measure itself. 2) An appropriate range for weights, perhaps over a 2 to 1 ratio, must be found. 3) An appropriate distribution of the weights is also necessary. A logarithmic transformation might be useful to heavily weight only extreme differences.

5.6. Potential Health Effects from Proposed Power Plant

The result from the logit analysis is used to examine potential health effects of a proposed coal power plant in the Chestnut Ridge area. The Penn Electric Company, which operates the four existing power plants, has proposed a new 625 MW(e) coal fired unit. The favored location is

near the Seward plant. This project is in the early planning stages as of July, 1981. The draft environmental impact statement (EIS) [Army Corps of Engineers, 1981] contains air pollution forecasts. A Gaussian dispersion model was used to estimate pollutant levels attributable to this plant. One year of meteorological data (1975) and the most recent monitoring data (1979) were used to generate the pollutant concentrations shown in Table 5.44. As detailed in the EIS, the plant will burn cleaned 2.4% sulfur coal, use electrostatic precipitators for particulate control, flue gas desulfurization (employing the best available technology) for sulfur oxides control, and a 645 foot high stack to aid pollutant dispersion. The proposed unit would increase annual sulfur concentrations by 6.85%.

Health effects are estimated using the coefficients for serious chest illness in the logistic regression model, shown earlier in Table 5.40. The coefficient is approximately 1.0 for annual sulfur concentrations. The logit formulation is shown below. P is a proportionality constant selected to balance the left side of the equation. The average prevalence rate for serious chest illness in the children's study is 0.082.

$$\ln\left(\frac{0.082}{1 - 0.082}\right) = P \quad (23)$$

We may easily find the increase in prevalence D attributable to an increase of 6.85% of P:

$$\ln\left(\frac{0.082 + D}{1 - 0.082 - D}\right) = P/1.0685 \quad (24)$$

In this instance, the model predicts that the prevalence rate will

increase by 2.92%, a 16% increase.

This model may be "fine tuned" to the high pollution and high prevalence found in this area. The four closest schools to the power plant, numbers 2, 5, 12, and 13, which are all within several miles of the plant, have a population weighted prevalence of serious chest illness of 9.68%. (Developed from Table 4.3.) Using this lower baseline prevalence in (4), the model calculates a 1.42% increase in the prevalence rate, which represents a total increase of 14.7%. In a deterministic fashion, we may calculate that 19 children, out of the 1333 children who go to these 4 schools, may experience serious chest illness which is related to the emissions of the power plant. 129 children already report this symptom.

This brief analysis has not examined other diseases or health effects felt outside the immediate area. Typically, a large power plant will incrementally affect air quality within a 50 mile radius. Thus, a health effects assessment would need to consider far more factors, e.g., demographic distributions, than this analysis. However, this analysis does have the considerable advantage of using a health model developed from the population at risk. Therefore, it may reflect the actual risk to young children in this section of Indiana County. However, all such predictions must be qualified by the limitations of the data and the methodology used in the study. These results may be in serious error, for the reasons of confounding, bias, etc., and probably are not transferable to other situations.

Chapter 6. Conclusion

Much of this research consisted of data exploration, searching to find a stable and clinically meaningful measure of air pollution which would be useful in a cross-sectional epidemiologic study of respiratory symptoms in young children. The study demonstrated that results of air pollution epidemiological investigations may be drastically influenced by the use of different pollution measures. Manipulations and perturbations of the air pollution data were shown to have a large effect on both the stability of the pollution measure and the statistical significance and magnitude of the association with the health data.

The arrowhead profile was used to separate the myriad of possible pollution measures by averaging time and percentile. This analysis indicates that the percentile, not the averaging time, of the pollution measure appears to make the more important difference with respect to stability and health effects assessment. Of course, many effects are interdependent. Short averaging times, i.e., less than 24 hours, and high percentile concentrations, i.e., peaks, second highest peaks, and even 90th percentile concentrations, tend to vary greatly from year to year. These measures do not reflect changes in pollution patterns which were known to occur. They also had a weaker association with the prevalence of respiratory symptoms in children than more stable measures, e.g., annual mean concentrations. In general, such measures are caused by relatively unfrequent events, for example, stack downwash and fumigation. High percentile concentrations are likely to happen several times each year, but are not representative of a more constant exposure. This suggests

that the peak measures in the National Ambient Air Quality Standards may not be particularly useful for epidemiologic studies. It can not be claimed, however, that these measures are irrelevant. Quite to the contrary, it seems important to incorporate this or similar information into policy and risk assessments, for it contributes information regarding high percentile concentrations which is not contained in annual averages.

Sulfur dioxide gave more consistent results than total suspended particulates (TSP) or combined measures in the case study. TSP does not seem to be a good indicator of either sulfur pollution or health effects. In the study area, peak concentrations of TSP primarily result from small local sources, e.g., agricultural and roadway dust, and, in fact, were found to show a slight negative association in the screening model for health effects. Only one TSP "exceedence" measure, the amount of time that the annual ambient standard of 75 micrograms per cubic meter was exceeded, showed a significant degree of association with the health effects. Several measures of sulfur dioxide gave strong associations with certain health effects, i.e., the prevalence of serious chest illness before age two.

With an appropriately selected threshold (to give about 15 to 45% exceedence), exceedence measures appeared to be somewhat more stable estimates than the annual average concentrations. This suggests that measures derived from two percentiles, perhaps the 50th and 84th, will provide more stable estimates than measures based on one concentration. In addition, exceedence measures are remarkably insensitive to averaging time. Short averaging times appear to be slightly more stable, possibly

due to the greater number of observations that occur.

This study could not directly evaluate the clinical relevance of the pollution measures without making the critical assumption that the variation in the prevalence of respiratory symptoms is caused by the pollution variable alone. Although this is not true, a relative comparison of measures is possible. No measure of a particular pollutant clearly surpassed the others in terms of clinical meaning in the study. Air pollution dosage measures, which were based on various physiological hypotheses related to the nature of the toxicity of pollutants, did not necessarily give higher correlations with respiratory symptoms. This is probably due to the importance of other risk factors in causing these symptoms, inappropriate dose-response functions, or the use of imprecise estimates and testing techniques. The exceedence measures, however, did tend to give associations which were as high or higher than more common measures, such as annual average concentration.

Difficulties encountered in an epidemiologic study do not invalidate the application of risk-cost-benefit analysis to air pollution policies. They do, however, point out several areas of major uncertainty, which are easily missed, and which may be extremely difficult to handle. Foremost among these is the selection of the proper measure of ambient air pollution. Major differences between pollutants and characterizations have already been discussed.

A second problem, highlighted by this research, is the difficulty of accurately characterizing geographic units and individual exposures, even with a high monitor density. The approach which used school exposures gave a statistically insignificant association with health effects.

Residence scores, which had a greater range, did give a meaningful association. The effect of uncertainty in estimating an individual's dosage may have been increased in the case study, due to the relatively small size of the Chestnut Ridge region and the complex terrain. On the other hand, this small size gives power to the health estimates, since potentially confounding factors, which exist along geographic patterns, tend to be minimized.

A third critical issue for both risk assessments and epidemiologic studies concerns the importance of using several years of pollution information -- from either monitors or weather driven dispersion models. Pollutant levels may vary considerably from year to year. In the study, averages of four or five years of data consistently gave more stable results than less data. Long term averages may help characterize the exposure history of the study population. Several years of data also seem necessary to obtain valid estimates of pollutant levels for risk assessments. For prospective assessments using dispersion models, a conservative recommendation is the use of several years of weather data, as generally suggested in regulatory practice (for new source applications).

Lastly, risk assessments, especially those using dispersion models, should incorporate the considerable error which is likely to exist in air pollution data and forecasts. This error may be compounded by the use of non-linear health models. For example, a 7% increase in sulfur levels in the Chestnut Ridge area is associated with a 15% increase in the prevalence of certain respiratory diseases. Moreover, the effect of the

uncertainty varies as a function of existing pollution and disease prevalence.

This study shows that better measures of particulates are desirable. Measures of fine or respirable particulates would probably give more useful results. A further extension of this work might be the analysis of coefficient of haze data as a surrogate for fine particulates. This data is available for the Chestnut Ridge area. Also, several new indices might be constructed which would combine concentrations at two or more percentiles using log-normal (or empirically determined) correction factors so that equal weight is given to each observation. This may produce a more stable and precise measure than the exceedence measures, which were based on interpolations. Lastly, new cleansing and acute indices might be constructed by using the ratio of the logarithms of the 50th to 16th percentile concentrations, and the 84th to 50th percentiles, respectively. These measures, used simultaneously with the mean concentration in the multiple regression model, might provide insight into the importance of cleansing and acute measures.

The major contribution of this study has been a description and evaluation of methods of characterizing air pollution for cross-sectional analysis. The use of more stable and representative pollution measures should improve studies of pollution related health effects. Health effects are now observed at pollutant levels, which, a decade ago, were thought to incorporate a margin of safety. More precise data collection and analytic techniques may continue to lower the threshold at which health effects are observed. This may cast doubt upon the effectiveness

of the threshold standards embodied in the National Ambient Air Quality Standards. Results of health models used in epidemiologic investigations will remain controversial. However, improvements in the methodologies of these investigations -- including the characterization of ambient air pollution -- may help reduce the scientific uncertainty of the models. This may help orient the policy debate more towards the value questions of what a tolerable level of air pollution is, and which mechanisms for policy implementation are most appropriate, rather than on the constraints of scientific and technological knowledge.

Appendix A. A Suggested Arrowhead Approach for Dosage Measures

A modification of the arrowhead profile used in this study might prove useful for dosage estimates. The duration of concentrations or "exceedences" which are of particular interest due to clinical hypotheses, could be taken directly from the suggested format. As illustrated in Figure A.1, the plot lacks the symmetry of the arrowhead curve. However, it would be derived quite similarly. Isopleths are constructed for pollutant concentrations of interest, in this case for 10, 20, 30, 50, 100 and 200 ppm. The horizontal axis remains the logarithm of averaging time. The vertical axis now represents fractions of a year, in terms of percentages, when isopleth concentrations are exceeded, in the bottom half of the plot, and, in the top half, the amount of time when concentrations are below the isopleth. For example, at an averaging time of 24 hours, the 200 ppm concentration is exceeded 10% of the time. It can be seen that ambient concentrations were below 30 ppm 20% of the time. Conversely, it was above 30 ppm some 80% of the time.

This construction permits time exceedences of clinically meaningful thresholds to be read directly from the plot. Dosages may be quickly estimated. For example, assume that health effects are proportional to the amount of time spent above 30 ppm (a threshold hypothesis), and that concentrations above 100 ppm cause more severe effects (non-linear effects). A dosage score D which combines the two effects with weight w given to the higher level exposure may be quickly constructed:

$$D = E(t > 30) + E(t > 100) \quad (A1)$$

Figure A.1. Suggested modified arrowhead profile for dosage measures.

Lastly, this format permits several years of plots to be combined directly. This can be done indirectly for several concentrations using averages of exceedence measures.

Appendix B. Derivation of Non-Linear Dosage Estimates

This appendix describes a method of estimating annual dosage estimates based on a non-linear, non-threshold dose-response function. (An exponential function derived from the Pollution Standard Index's definition was described in the text.) Basically, annual concentrations are disaggregated into 6 averages, using concentrations at the 6 arrowhead percentiles. Then, the health function is imposed on these 6 averages, weighted by the duration of time that each average represents, and summed. This represents a piece-wise integration of the health function over the concentrations found in the course of the year.

There are two assumptions made: six points are sufficient to describe 365 days of pollution concentrations; and a lognormal distribution of pollutant concentrations exists. A check on the estimates revealed that six points did in fact give very good results. Departures from normality are known to occur at both high and low tails of the distribution. However, for the purposes of this estimate, its effect is judged to be negligible.

Figure B.1 shows a normal distribution applied to the arrowhead curve's percentile groups. Also shown are the percentile "intervals" which become the interval boundaries in this analysis. They are selected to roughly center the available percentiles, and to give approximately

Figure B.1 Standard normal distribution applied to pollutant concentrations on percentile basis. Represented by dashed lines are the average concentrations found in each interval.

Interval	Interval Boundary	Z St. Dev.	$P(Z > z)$	$P(Z > z_1) - P(Z > z_2)$
1	0	-	1	0.0202
	2	2.05	0.9798	
2	2	-2.05	0.9798	0.3505
	37	-0.33	0.7023	
3	37	-0.33	0.7023	0.1293
	50	0	0.5	
3b	50	0	0.5	0.1293
	63	0.33	0.3707	
4	63	0.33	0.3707	0.2497
	88	1.17	0.1210	
5	88	1.17	0.1210	0.1008
	98	2.05	0.0202	
6	98	2.05	0.0202	0.0202
	100	-	0	

Table B.1 Interval definitions.

equal weight to the 16th, 50th and 84th percentile groups. The idea is to use the normal distribution to find the percentile value, in terms of a z value, or number of standard deviations from the median, which best represent the average concentration in each interval.

To do this, we first find the probability of each interval. Table B.1 shows the boundaries of the six intervals, their z values, the probability of each interval, and the mid-interval probability. Z values are found knowing that the 16th and 84th percentiles represent one standard deviation from the median, or in the general case,

$$P(Z > z) = P(Z > 0.01 \text{ percentile}) \quad (\text{A2.1})$$

Next, we need to find the "mid-point probability" Z_m of each interval, that is:

$$P(z_1 > z) - P(z_2 > z_m) = P(z_1 > Z_m) - P(z_2 > Z_m) \quad (\text{A2.2})$$

where z_1 and z_2 represent the probabilities at the boundaries of the interval. Table B.2 also shows the standard normal densities at each Z_m . Note that the interval surrounding the 50th percentile is disaggregated into two symmetrical parts. Only one mid-point need be derived.

Lastly, the standard normal densities of the interval midpoint $D(Z_m)$ and the normal densities for the known percentiles $D(z)$ are used to adjust the pollution concentrations to the representative interval average $C(i)$:

<u>Interval</u>	<u>D(Zm) Standard Mid-point</u>	<u>Normal Density</u>
1	-2.32	0.0271
2	-1.16	0.2036
3	0.17	0.3932
4	0.69	0.3144
5	1.47	0.1354
6	2.32	0.0271

Table B.2 Definition of interval mid-point and normal density.

<u>Existing Percentile</u>	<u>Z Percentile</u>	<u>Standard Normal</u>	<u>Fraction of Year</u>	<u>Adjustment Factor</u>
1	-2.33	0.0264	0.02	1.023
16	-1	0.2420	0.25	1.189
50	0	0.3989	0.26	0.986
84	1	0.2420	0.25	1.299
90	1.28	0.1758	0.10	0.771
99	2.33	0.02643	0.02	1.023

Table B.3 Results for 6 part approximations.

$$C(i) = C(z) D(z)/D(Z_m)$$

(A2.3)

These values are displayed in Table B.3. These concentration values are used in the health function. Then, the health index is multiplied by the the adjustment factor shown in Table B.3.

A simple check proved the accuracy of this scheme. The products of the adjustment factor and the time interval, i.e., the last two columns of Table B.3, were multiplied by the concentrations at the appropriate percentiles on the arrowhead curves. This product should give an annual linear dosage -- exactly equal to the average annual concentration. Four years of sulfur data and five years of TSP data were checked at each of the 17 monitors. Correlations were very high; about 0.95 for sulfur and over 0.98 for TSP. Thus, this approximation would appear to be quite accurate.

Appendix C. Derivation of Threshold Dosages

This appendix describes the method used to develop threshold dosage measures from exceedence measures. Three assumptions are made. First, pollutant concentrations are lognormally or normally distributed. Second, all monitors are assumed to have approximately the same variance. Third, only two or three points are necessary to estimate threshold dosages. The first two assumptions are not terribly confining in the Chestnut Ridge area. Annual concentrations among the 17 monitors were generally within 20 or 30% of one another. Also, most monitors had a similar coefficient of variance.

Below are shown the 4 or 5 year average and mean 84th percentile values for the 17 monitors.

	TSP <u>ugms/cu.m</u>	Sulfur <u>ppb</u>
Mean concentration	67	26.3
84th percentile concentration	100	45
Standard deviation (upper)	33	18.7

Exceedence scores provide the percentage of time a monitor's concentrations exceed a particular threshold. This is illustrated in Figure C.1. The shaded area represents the portion of time that the thresholds -- in this case 30, 80 and 150 ppb -- were exceeded. The dashed lines display a three part approximation to the normal curve, which is used to represent pollutant concentrations. However, we know neither the values actually experienced nor the mean of the values in any of the three intervals. However, if we assume that the pollutant concentrations are normally or lognomally distributed, we may calculate the average concentration in each interval.

Initially we will assume that concentrations are normally distributed using the means and standard deviations just given. The procedure is quite similar to that in Appendix B. For each threshold, let z = number of standard deviations above the mean. We then find the probability of each interval, then the mid-point probability, then the Z_m which represents this probability. Then, Z_m is transformed to pollutant concentration $C(Z_m)$. These values are shown in Table C.1.

These concentration estimates represent the best estimate of the

Figure C.1. Upper tail of distribution representing pollutant concentrations for derivation of average concentration in interval.

<u>Pollutant</u>	<u>Threshold</u>	<u>Z (threshold)</u>	<u>Zm</u>	<u>C(Z)</u>
Sulfur	30ppb	0.198	0.84	42.0ppb
	80	2.87	3.07	83.7
	150	6.61	6.72	152.0
TSP	75ug/cu.m	0.24	0.81	93.7ug/cu.m
	150	2.52	2.75	157.8

Table C.1. Average concentrations in intervals using a normal distribution.

<u>Pollutant</u>	<u>Threshold</u>	<u>Z (threshold)</u>	<u>Zm</u>	<u>C(Z)</u>
Sulfur	30ppb	0.447	0.82	37.3ppb
	80	1.360	1.49	104.8
	150	1.645	1.97	346.6
TSP	75ug/cu.m	0.595	0.89	89.5
	150	1.264	1.63	365.6

Table C.2 Average concentrations in intervals using a lognormal distribution.

average interval concentration assuming the distribution is correct. The threshold dosage may then be calculated as

$$D = \sum_{i=1}^n [T(c > z_i) - T(c > z_{i+1})] C(z_{mi}) \quad (A3.1)$$

where $T(c > z)$ is the exceedence measure, with n such measures, and thresholds are overlapping and increasing in concentration with increasing i .

We may also attempt the same exercise assuming a lognormally distribution of pollutant concentrations. In this case

$$z = \log(\text{threshold} - \text{mean}) / \log(\text{standard deviation}) \quad (A3.2)$$

$$C(z) = e^{z \log(\text{standard deviation}) + \text{mean concentration}} \quad (A3.3)$$

Otherwise the analysis is parallel to the previous case. Calculated values are shown in Table A3.2.

The log-normal distribution resulted in very high concentration estimates for the means of the highest interval. These very high concentrations were rarely experienced in the Chestnut Ridge area. The small deviation from normality -- which mostly occurred in the tails of the distribution -- has a major effect in this analysis.

The log-normal dosage measure gives more weight to higher concentrations. The fraction of the contribution for each interval may be seen in the Table C.3.

	<u>30 - 80ppb</u>	<u>80-150ppb</u>	<u>150ppb & up</u>	<u>total dosage</u>
Normal Dosage				
ppb-year	11.53	3.57	6.03	21.12
% *	54.5	16.9	28.6	100.
Log-Normal Dosage				
ppb-year	10.23	4.47	14.50	29.21
% *	35.0	15.3	49.6	100.

* contribution in percent

Table C.3. Contribution of the various intervals to the total threshold dose. Averages of 17 monitors for 4 years for sulfur.

Appendix D. Weighting for Statistical Variation

Unlike other types of data, air pollution scores should not be inversely weighted by their variance to account for the certainty of the observation. A pollution score which remains "dirty" relative to other scores of other sites in the cross-sectional analysis, and which varies considerably from year to year, should not be penalized. This situation is, in fact, quite common in the Chestnut Ridge area, as it probably is elsewhere. For example, monitor 17 was consistently dirty and also had a high coefficient of variance (COV). The technique used here assumes 1) a normal distribution for pollutant measures at all sites, and 2) that the mean and variance of all individual sites (as well as the mean of all sites) are known.

Figure D.1 illustrates the technique. A normal distribution is shown

for the concentrations of all sites, using the average concentration Ca , and the mean standard deviation of all the individual sites Csd , where:

$$Ca = \frac{\sum_{\text{years}}^n \sum_{\text{sites}}^m \text{score}}{(n + m)} \quad (D.1)$$

$$Csd = \sum_{\text{sites}}^m \left(\frac{\sum x^2 - (\sum x_i)^2/n}{n - 1} \right) / m \quad (D.2)$$

The interval surrounding the 4 or 5 year mean concentration of each site is simply defined by:

$$I_{\text{upper}} = \text{site mean} + 1 (\text{site}) \text{ standard deviation} \quad (D.3)$$

$$I_{\text{lower}} = \text{site mean} - 1 (\text{site}) \text{ standard deviation} \quad (D.4)$$

These interval boundaries are then normalized to the group's statistics, where:

$$I_{\text{un}} = \frac{I_{\text{upper}} - Ca}{Csd} \quad (D.5)$$

$$I_{\text{ln}} = \frac{I_{\text{lower}} - Ca}{Csd} \quad (D.6)$$

Lastly, the probability of the interval is found, and subtracted from 1:

$$W = 1 - [p(Z > I_{\text{un}}) - p(Z > I_{\text{ln}})] \quad (D.7)$$

Figure D.1. Normal distribution fitted to of averages of site concentrations. Site a lies in the tail of the distribution of values but is not penalized in the weighting scheme. Site b, with the same standard deviation, is penalized because it lies near the center of the distribution, where most observations are located.

School Number	Sulfur		TSP	
	Index 6	Index 9	Index 6	Index 9
1	0.6753	0.0930	0.6232	0.6049
2	0.8215	0.3579	0.8391	0.5054
3	0.0967	0.5025	0.6297	0.3010
4	0.4391	0.0736	0.7806	0.5426
5	0.1095	3.6E-03	0.7210	0.9475
6	0.2377	0.5768	0.6032	0.4998
7	0.0396	0.0841	0.8539	0.5428
8	0.3072	0.4251	0.5706	0.0660
9	0.2434	5.2E-04	0.7754	0.7312
10	0.0540	0.0182	0.6324	0.5801
11	0.2753	0.8579	0.5943	0.7342
12	0.0922	3.4E-03	0.7136	0.9497
13	0.7592	0.9921	0.8209	0.3688
14	0.1882	0.5811	0.5813	0.4515

Table D.1. Weights for statistical variation of indices 6 and 9 at each of the 14 schools.

The result may be defined as the expected porportion of observations that have a 67% chance of being different than the mean of the site being investigated. Sites with either a small standard deviation or "located" the tail of the distribution, will have a weight approaching unity. This is demonstrated in Figure D.1, where the certainty weight is effectively equal to the area not included in shaded area, for two sites, A and B.

Table D.1 shows certainty weights for indices 6 and 9 for sulfur and TSP which were developed for the 14 schools. Index 9 is a high percentile, short averaging time measure, which does not fit the assumption of a normal distribution of values. The range of the weights is far too large. Weights for index 6, the annual mean, are better behaved. For example, school 13 is dirty but with a large COV: its weight is high. Schools 3, 5, 6, and 14 experience more or less average concentrations, but are downwind of several major sources. Thus, their concentrations may tend to vary, depending on the weather. They have been penalized by the weighting scheme.

The text describes some of the limitations of this technique.

Bibliography

American Thoracic Society, "Health Effects of Air Pollution," American Lung Association, New York, NY, 1978.

American Meteorological Society, "Judging Air Quality Model Performance," Workshop on Dispersion Model Performance at Woods Hole, Sept 8-11 1980.

Ashford, N., "Crisis in the Workplace: Occupational Disease and Injury," MIT Press, Cambridge, MA, 1976.

Ashford, N., M. Seminario, "The Chemists' Guide to Occupational Safety and Health," MIT Press, Cambridge, MA, 1977.

Bacow, L.S., "The Technical and Judgemental Dimension of Impact Assessment," MIT Industrial Liason Program, Jan., 1981.

Casarett, J., "Casarett's and Doull's Toxicology: The Basic Science of Poisons, 2nd Edition," Macmillan Publishing Co., New York, 1980.

Center For Policy Alternatives (CPA), "Evaluation of Chemical Regulations: Trade-Off Analysis and Impact Assessment for Environmental Decision Making," Cambridge, MA, March, 1979.

Cole, P., "The Evolving Case Control Study," Journal of Chronic Diseases, vol. 32, p. 15-27, 1979.

Committee on Interstate and Foreign Commerce, U.S. House of Representatives, "Clean Air Act Amendments of 1977 -- Report 95-294," U.S. Government Publishing Office, Washington, D.C., 1977.

Committee on Science and Technology, U.S. House of Representatives, "The EPA's Research Program with Primary Emphasis on the Community Health Surveillance System," U.S. Gov. Printing Office, Washington, D.C. 1976.

Cooper, D.E., W.C. Hamilton, "Atmospheric Sulfates and Mortality -- The Phantom Connection," Continental Oil Co., Ponca City, OK, 1979.

Council on Environmental Quality (CEQ), "Ninth Annual Report," U.S. Government Printing Office, Washington, D.C. December, 1979.

Council on Environmental Quality (CEQ), "Tenth Annual Report," U.S. Government Printing Office, Washington, D.C. December, 1980.

Crump, K., "Dose-Response Problems in Carcinogenesis," Biometrics, v. 35, p. 157-67, March, 1979.

Crump, K., "Fundamental Cancer Processes and their Implications for Low Dose Risk Assessment," Cancer Research, no. 36, p. 2973-79, Sept. 1979.

Bibliography (continued)

Dixon-Davis, D.K., "Mortality in 191 Counties near Energy Facilities: Use of Economic and Social Variables for Standardization," Division of Biological and Medical Research, U.S. Department of Energy, 1979.

Evans, J., et. al., "A Proposal for Air Pollution Research: To Estimate the Benefits of Air Pollution Abatement," Proposal to U.S. EPA, Feb. 1981.

Gruhl, J., "Scientific Background on Probabilistic Air Pollution Dosage Modeling," MIT Energy Lab., MIT-EL-014WP, May, 1976.

Gruhl, J., et. al., "Systematic Methodology for the Comparison of Environmental Control Techniques for Coal-Fired Electric Generation," MIT Energy Lab., EL-76-012, Nov, 1976.

Gruhl, J., et. al., "Evaluation of the Health Effects of Air Pollution in the Chestnut Ridge Area," Report to U.S. DOE," MIT Energy Lab, May, 1979.

Gruhl, J., et. al., "Evaluation of the Health Effects of Air Pollution in the Chestnut Ridge Area," Report to U.S. DOE," MIT Energy Lab, May, 1980.

Hamilton, L.D., "Impacts of Atmospheric Alteration, Summary Discussion and Proposal for Future Work, Long Range Transport, Air Pollution Health Effects," Biomedical and Environmental Assessment Division, Brookhaven National Laboratory, Feb., 1981.

Heimann, H., "Air Pollution," World Health Organization, Columbia University Press, New York, NY, 1961.

Henrion, M., "Expected Value of Information for an Air Pollution Control Policy on Atmospheric Sulfates (Draft)," Department of Engineering and Public Policy, Carnegie Mellon Univ., Pittsburg, PA, April, 1981.

Holland, W.W., "Air Pollution and Respiratory Diseases," Technomic Publishing Co. Inc., Westport, CT, 1972.

Larsen, "A New Mathematical Model of Air Pollution Concentrations," Journal of the Air Pollution Control Association, v.19, p.24-30, 1969.

Lave, L., E. Seskin, "Air Pollution and Human Health," John Hopkins University Press, Baltimore, 1977.

Lee, S.D., ed., "Biomedical Effects of Environmental Pollutants," U.S. Environmental Protection Agency Health Effects Research Lab., Ann Arbor Science Publishing Inc., Ann Arbor, MI, 1977.

Lipfert, F.W., "Differential Mortality and the Environment: The Challenge of Multicollinearity in Cross-Sectional Studies," Energy Systems and Policy, vol. 3, no. 4, 1980.

Bibliography (continued)

Lundy, R.T. et. al., "Predictions of the Effects of Energy Production on Human Health," Proceedings of the Social Statistical Section, American Statistics Association, Chicago, IL, Aug 15-8 , 1980.

Lynn, D.A., "Air Pollution -- Threat and Response," Addison-Wesley Publishing Co., Reading, MA, 1976.

MacMahon, B., "Strengths and Limitations of Epidemiology," Harvard School of Public Health, Boston, MA, 1975.

Morgan, M.G., et. al., "A Probabilistic Methodology for Estimating Air Pollution Effects from Coal Fired Power Plants," Energy Systems and Policy, vol. 2, no. 3, 1978.

Murdoch, "Environment, Resources, Pollution and Society," 1975.

National Commission on Air Quality, "To Breath Clean Air," U.S. Government Printing Office, Washington, D.C., 1981.

Orloff, N., "Rethinking Environmental Law," Sunday New York Times, May 10, 1981.

Paté, E., "Acceptable Decision Making Process Versus Acceptable Risk for Public Policy Decisions," ORSA-TIMS Meeting, Nov. 1980, Colorado Springs, CO, 1980.

Perera, F.P., "Respirable Particulates: Impact of Fine Particulates on Health and the Environment," Ballinger Pub. Co., Cambridge, MA, 1979.

Purdon, P.W., ed., "Environmental Health, Second Edition," Academic Press, New York, NY, 1980.

Ricci, P., et. al., "A Statistical Review of Cross-Sectional Estimates of Mortality," presented at the US DOE Statistical Symposium, Berkeley, CA, Oct. 1980.

Ricci, P., "Modeling Basis for Energy Risk Assessment," presented at the Air Pollution Control Association Conference, Phila, PA, June 21-6, 1980.

Rothman, K.J., "Causes," American Journal of Epidemiology, v.101, no.6, 1976.

Rothman, K.J., "Induction and Latent Periods," Report for the National Institute of Neurological and Communicative Disorders and Strokes (No. NS-8-2399), 1977.

Ruane, M., et. al., "Design and Implementation of a Supplementary Control System," MIT Energy Lab., MIT-EL-80-033, 1980.

Bibliography (continued)

Schenker, M.B., et. al., "Health Effects of Pollutants Associated With Coal-Fired Power Plants in a Rural Area," Channing Lab., Harvard Medical School, June, 1981.

Szolovits, P. "Artificial Intelligence and Clinical Problem Solving," MIT Industrial Liason Paper, Computer Science Lab, Cambridge, MA, Jan. 1980.

U.S. Army Corps of Engineers, "Draft Environmental Impact Statement, Seward Generating Station Unit 7, Indiana and Westmoreland Counties, Pennsylvania," Pittsburg Division, April, 1981.

U.S. Department of Health, Education and Welfare, "Air Quality Criteria for Sulfur Oxides," Washington, D.C., March, 1967.

U.S. E.P.A., Office of Air Quality Planning and Standards, "Guideline on Air Quality Modeling: Proposed Revisions," Research Triangle Park, NC, Oct. 1980.

U.S. E.P.A., Office of Air Quality Planning and Standards, "Guidelines for Air Quality Maintenance, Planning and Analysis: Procedures for Evaluating Air Quality Impacts of Stationary Sources," vol. 10, Research Triangle Park, NC, Oct. 1977.

U.S. E.P.A., "Air Pollution Studies Near a Coal Fired Power Plant: Wisconsin Power Plant Impact Study," EPA-600/3-80-048, May 1980.

U.S. E.P.A., "Pilot Study of Ambient Air Pollution and Survival From Cancer," Epidemiological Branch, Research Triangle Park, NC, Nov. 1979.

U.S. Supreme Court, "Industrial Union Department, AFL-CIO v. American Petroleum Institute et al.," July 2, 1980.

Viren, J.R., "Cross-Sectional Estimates of Mortality due to Fossil Fuel Pollutants: A Case for Spurious Association," Report TR-17900, for Division of Policy Analysis, U.S. Department of Energy, Washington, D.C., Aug. 1978.

Williams, L., M.A. Spence, S.Tideman, "Implications on the Observed Effects of Air Pollution on Birthweight," Social Biology, v. 24, no. 1, Spring, 1977.

Wilson, R., et. al., "Health Effects of Fossil Fuel Burning: Assessment and Mitigation," Ballinger Pub. Co., Cambridge, MA, 1980.