EPA-450/4-74-012
September 1974
(OAQPS No. 1.2-030)

GUIDELINES FOR AIR QUALITY MAINTENANCE PLANNING AND ANALYSIS

VOLUME 11 : AIR QUALITY MONITORING AND DATA ANALYSIS

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

EPA-450/4-74-012
(OAQPS No. 1.2-030)

GUIDELINES FOR AIR QUALITY MAINTENANCE PLANNING AND ANALYSIS VOLUME 11 : AIR QUALITY MONITORING AND DATA ANALYSIS

Prepared by

the GCA Corporation
in partial fulfillment of
Task Order No. 1, Contract No. 68-02-1478

Program Element No. 2AH137

ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711
September 1974

OAQPS GUIDELINE SERIES

The guideline series of reports is being issued by the Office of Air Quality Planning and Standards (OAQPS) to provide information to state and local air pollution control agencies; for example, to provide guidance on the acquisition and processing of air quality data and on the planning and analysis requisite for the maintenance of air quality. Reports published in this series will be available - as supplies permit - from the Air Pollution Technical Information Center, Research Triangle Park, North Carolina 27711; or, for a nominal fee, from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22151.

This report was furnished to the Environmental Protection Agency by the GCA Corporation, Bedford, Massachusetts, in partial fulfillment of Task Order No. 1, Contract Number 68-02-1478. Prior to final preparation, the report underwent extensive review and editing by the Environmental Protection Agency and other concerned organizations. The contents reflect current Agency thinking and are subject to clarification, procedural change, and other minor modification prior to condensation for inclusion in Requirements for Preparation, Adoption, and Submittal of Implementation Plans (40 CFR Part 51).

Publication No. EPA-450/4-74-012
(OAQPS Guideline No. 1.2-030)

FOREWORD

This document is the eleventh in a series comprising Guidelines for Air Quality Maintenance Planning and Analysis. The intent of the series is to provide State and local agencies with information and guidance for the preparation of Air Quality Maintenance Plans required under 40 CFR 51. The volumes in this series are:

Volume 1: Designation of Air Quality Maintenance Areas
Volume 2: Plan Preparation
Volume 3: Control Strategies
Volume 4: Land Use and Transportation Consideration
Volume 5: Case Studies in Plan Development
Volume 6: Overview of Air Quality Maintenance Area Analysis
Volume 7: Projecting County Emissions
Volume 8: Computer-Assisted Area Source Emissions Gridding Procedure
Volume 9: Evaluating Indirect Sources
Volume 10: Reviewing New Stationary Sources
Volume 11: Air Quality Monitoring and Data Analysis
Volume 12: Applying Atmospheric Simulation Models to Air Quality Maintenance Areas

Additional volumes may be issued.

All references to 40 CFR Part 51 in this document are to the regulations as amended through July 1974.

PREFACE

The Monitoring and Data Analysis Division of the Office of Air Quality Planning and Standards has prepared this report entitled "Guidelines for AQM Planning and Analysis, Air Quality Monitoring and Data Analysis (Vol. 11)," for use by the Regional Offices of the Environmental Protection Agency and by State and local air pollution control agencies. This report draws upon the information contained in the previously issued air quality monitoring documents listed below and applies them toward air quality maintenance planning:

Guideline No.	Date	Title
1.2-008	Aug. 1974	Guidelines for the Interpretation of Air Quality Standards
1.2-012	Jan. 1974	Guidance for Air Quality Monitoring Network Design and Instrument Siting (Draft revision)
1.2-013	May 1974	Procedures for Flow and Auditing of Air Quality Data
1.2-014	Feb. 1974	Guidelines for the Evaluation of Air Quality Trends
1.2-015	Feb. 1974	Guidelines for the Evaluation of Air Quality Data
1.2-018	May 1974	Designation of Unacceptable Analytical Methods of Measurement for Criteria Pollutants
1.2-019	July 1974	Air Quality Monitoring Site Description Guideline

Adherence to the guidance presented in this report will, hopefully, lead to acquisition of more useable and mutually compatible data by all States and Regions and will also facilitate the implementation of State air monitoring programs which are compatible with the goal of air quality maintenance.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

SECTION 11.1

INTRODUCTION

The purpose of this document is to provide the states with planning information and preliminary guidance for the preparation and implementation of a monitoring system compatible with the goal of air quality maintenance and the need for the development of Air Quality Maintenance Plans required under 40 CFR 51 as amended on June 18, 1973. In order to satisfy this objective, background information on the air quality maintenance planning process, as well as summary descriptions of current monitoring techniques, is included herein. However, the major thrust of this document is to provide instruction on monitoring to the state and local agencies that are charged with the maintenance of air quality.

General guidelines have been issued which are intended to provide a basic rationale for the development and evaluation of air monitoring networks. These earlier guidelines are largely subjective but do provide a good understanding of the generally applicable considerations. Much of the guidance provided by these earlier publications is repeated herein. Areas where new guidance is presented include: location of samplers, emphasis of the monitoring network, and data analysis.

The development and implementation of a network must by necessity involve a trade-off between what is considered desirable from a strictly technical point of view and what is feasible with the available resources. An ideal network will, in almost all instances, require more resources than are currently available. Since air quality maintenance planning is a continuing, long-range activity, this document encourages the development of the "ideal" network. While the implementation of such a network would not, in most cases, be possible immediately, it is felt that the time scales involved allow for the eventual development of a network which fully meets the needs of air quality maintenance planning. This document does not address the development of interim networks due to the wide variations of situations which occur in maintenance planning; however it is felt that this document provides sufficient background information on the needs of air quality maintenance planning to establish the basis for making judgemental decisions for the most appropriate interim network. The basic difference between the interim monitoring network and the ideal is that the interim network has fewer and perhaps less sophisticated instruments. Designers of the network should attempt to maximize the effectiveness of the interim network through careful selection of sampling sites, scheduling of variable sampling frequencies, and possible use of mechanical (integrated) as well as automatic (continuous) samplers. Careful consideration of the basic requirements of the monitoring network will often allow for the selection of one sampling site which may satisfy more than one need of the maintenance planning activities.

As with any guideline document, it is expected that the procedures presented will aid the goal of continuity and uniformity in data gathering and handling. It is the nature of the scope of this document that this uniformity is applicable to not only the data which is normally collected through a monitoring network, but also the handling and interpretation of data for the use of air quality maintenance planning. Again, since air quality maintenance planning is a continuing process, this document more heavily emphasizes these goals.

11.1.1 MONITORING IN GENERAL

The development of an air quality monitoring network includes determining the number and location of sampling sites, selecting appropriate instrumentation, determining frequency of sampling, and following established instrument siting criteria. Experience and judgment are essential in this process, especially for determining the number and location of sampling sites because mathematical models or other methods may not be entirely reliable or, in some instances, may not be available. In addition, the lack of necessary resources often requires the establishment of a minimum network which is based upon this judgmental expertise to best cover the monitoring needs under these constraints. General guidelines for the development of an appropriate network are given in Guidance for Air Quality Monitoring Network Design and Instrument Siting (OAQPS No. 1.2 - 012).

The selection of the methods for measurement of air pollutants requires less judgement. Officially approved methods (federal reference methods) are described in appendices to 40 CFR Part 50. In addition, any method which can be demonstrated to be "equivalent" to the reference method may be approved. Further guidance is provided through the Designation of Unacceptable Analytical Methods of Measurement for Criteria Pollutants (OAQPS No. 1.2 - 018).

Specific procedures have also been established for the processing of air quality data for input to the National Aerometric Data Bank. The guidelines for data handling, Procedures for Flow and Auditing of Air Quality Data (OAQPS No. 1.2 - 013), provide information and suggestions concerning data flow, editing, validation, correction procedures and certification, verification, and statistical flagging techniques.

The analysis of the data provided by the air quality monitoring network is less structured due, in part, to the varied analytical and statistical techniques which are available for the interpretation of the data. Such techniques range from the basic conventions for handling air quality data, summarizing it, determining characteristic patterns and making inferences from the data, as presented in Guidelines for the Evaluation of Air Quality Data (OAQPS No. 1.2 - 015), to the actual assessing and interpreting of trends based on statistical methodologies, as given in Guidelines for the Evaluation of Air Quality Trends (OAQPS No. 1.2 - 014).

These four basic concerns of any air quality monitoring network - network design, instrument selection, data processing, and data analysis - are separately discussed in this document with regard to the needs for air quality maintenance planning. While these concerns already have some basis due to their prior implementation, it is necessary to have an understanding of the activities involved in air quality maintenance planning and the likely impact these activities will have on monitoring needs. A comprehensive understanding of the considerations involved can be derived from reviewing the primary documents in this series. A summary of these considerations is given below.

11.1.2 MONITORING FOR AIR QUALITY MAINTENANCE

11.1.2.1 The Air Quality Maintenance Planning Process

All states, pursuant to 40 CFR 51.12(e), are required to identify areas that have the potential for exceeding any National Ambient Air Quality Standard (NAAQS) due to the current air quality and/or the projected air quality due to growth over the subsequent 10-year period. Such areas are called Air Quality Maintenance Areas (AQMA's), and may be identical to counties, urban areas, Standard Metropolitan Statistical Areas (SMSA's) or other boundaries. At 5-year intervals, the area identification shall be reassessed to determine if additional areas should be designated as AQMA's. EPA reviews the information supplied by the States and the Administrator issues an official list of the designated AQMA's.

For each area identified by the Administrator, the State must submit an analysis of the impact on air quality of the projected growth and development over the subsequent 10-year period. In addition, the State must prepare and submit an Air Quality Maintenance Plan (AQMP) which demonstrates that, within each AQMA, the national standards will be prevented from being exceeded over the 10-year period from the date of plan submittal. Such a plan is considered a revision to the state implementation plan and includes any control strategy revisions and/or other measures that may be necessary to insure that the projected growth and development will be compatible with the maintenance of the national standards throughout the 10-year period. The States are also required to review and revise the plans where necessary at 5-year intervals.

11.1.2.2 Particular Emphases of Monitoring

Throughout the maintenance planning process the emphasis is on determining and projecting air quality. Therefore, monitoring, which provides hard data on the actual air quality, is important in providing the basis from which projections are made and models are calibrated. In addition, the data from the monitoring network will be used during the preparation and implementation of the AQMP to assess the impact of the control strategies aimed at controlling the growth of emissions and therefore the degradation of the air quality.

With regard to air quality maintenance activities, the first task of the States must be the designation of the AQMA's. As this responsibility will have been performed for the first time as of the publication of this document, these guidelines concentrate on the establishment of a monitoring network useful for the identification of new AQMA's when considering the designation of AQMA's. Basically, such a system is directed at the determination of air quality trends. Since the development of an adequate monitoring network for an area which is not currently an AQMA requires much the same type of analysis as for an AQMA for which a maintenance plan has been developed, and therefore the analysis completed, the order of presentation in this document is first those areas which are not currently AQMA's but may be considered "potential AQMA's" and then those areas which are designated AQMA's. It is felt that this format allows for a better understanding of the considerations involved with minimum repetition.

Once an area has been designated as an AQMA and a control strategy has been implemented for the pollutant(s) of interest, the air quality monitoring network necessarily shifts from establishing the trend, which leads to the plan preparation, to the assessment of the impact of the control strategy. This provides the basis from which the determination of the achievement of the goals of the AQMP can be made. This allows for more stringent control measures to be developed and implemented if the need arises. In addition, the air quality data serves as the basis for the review of new stationary and indirect sources.

11.1.2.3 General Changes in Monitoring Due to Maintenance Planning

Several changes in the monitoring network are assumed to occur as a result of the need for monitoring for air quality maintenance planning. Generally, these changes are directed at the size of the network and the location of the instruments. The extent of these changes will depend upon how well the currently established network fulfills the criteria presented in this document.

The size of the network and location of the instruments for maintenance planning are naturally interrelated. Given the special concern which AQMA's have within the overall implementation planning mechanism, the data required, and hence the monitoring necessary, may be significantly more than is now operational or even planned in a given maintenance area. Air quality maintenance planning requires that certain areas, especially controlled or encouraged growth areas, be monitored to establish the trend of air quality. Other sites which may be of special interest in the maintenance plan deserve monitoring.

The other major change in monitoring due to air quality maintenance activities involves the analysis of the data. More attention needs to be directed at the reasons for variations or trends in the data, especially as they relate to other parameters. Essentially, many of the same techniques as currently employed will continued to be used with either an increase or change in emphasis.

Other areas which are stressed in this document are the result of the long-term nature of air quality maintenance planning. Such planning allows for more standardization of monitoring networks, data handling, and data analysis. Of particular interest is the phasing in of new standard equipment and the more complete use of improved data handling techniques.

SECTION 11.2

NETWORK CONFIGURATION

The configuration of an air quality monitoring network involves two distinct elements: the number of sampling sites of various types, and their geographical location. Under varying circumstances, decisions on the two elements can be made in either order; an overall number of sensors or sites may be selected, based on a criterion such as resource availability, and then distributed geographically, or specific sites may be selected first, based on a criterion such as the need for the data, with the aggregate number of sites then being just the total number of sites selected.

Historically, as the national control effort under the Clean Air Act began, emphasis was on developing not only new networks but also the resources, both human and monetary, to operate them. Consequently, the first approach was necessarily taken; networks were sized directly or indirectly in relation to the resources available, and the sites were distributed with as much consideration of sources, topography, etc., as was possible. These monitoring networks, were then, and still are being expanded or modified in the light of increasing knowledge, experience and funding.

However, in the case of planning an air quality monitoring network for purposes of long-term maintenance of ambient air quality standards, the emphases are necessarily different and the second approach is the more appropriate one. Rather than being primarily concerned with resource limitations, maintenance network planning is involved with a long-term, on-going, network configuration, and resource concerns will affect primarily the length of time required for the network to evolve into its ultimate configuration. Rather than being constrained by knowledge deficiencies, pollution control officials can now understand with some degree of precision what decisions they will be faced with in the future and what air quality data will be needed to assist in making those decisions. Consequently, the emphasis in this section will be on designing a network configuration to meet specific air quality maintenance needs, and the aggregate size of the monitoring effort will be determined as a result of meeting the various specific needs.

11.2.1 SCALES OF SENSOR PLACEMENT

Determining the location of air quality monitoring sites, including the location of the sensor or sampling probe itself, is a decision process that ranges over a very wide range of geographical scales. In the past, with air quality monitoring networks existing primarily in urban areas, the decisions involved could be conveniently summarized by defining three selection processes, each of which operates in a different size scale, and each of which depends on somewhat different criteria:

- Neighborhood Selection - the choice of the areas within the city where sites should be located - operating in a scale of many blocks or a few kilometers, and primarily determined by data needs.
- Site Selection - the choice of a specific site within the neighborhood to be monitored - operating in a scale of tens of meters, and largely determined by questions of accessibility, security, etc.
- Probe location - the selection of the precise location of the sensor or inlet probe - operating in a scale of a few meters or smaller, and determined by concerns of sample representativeness and standardization.

These three categories are derived largely from concerns with urban networks; monitoring for maintenance purposes is of course the same in many respects, but involves the addition of one larger-scale decision category which has been labeled Community Selection. This is meant to involve the selection of, from among all areas in an AQCR, those portions where monitoring is required for air quality maintenance purposes. While in some AQCR's, this may be no more than the control urban area, in others it may, in addition, involve outlying growth and recreational areas, areas where major new industrial or resource development activity is expected, and other areas of non-traditional concern.

These four categories form the framework for the discussion of designing a network configuration for air quality maintenance purposes. Community selection and neighborhood selection, which are primarily dependent on data and information needs, are considered rather thoroughly in the next two subsections. The areas of site selection and probe placement, which are less a function of data needs, are then considered somewhat more generally.

11.2.2 COMMUNITY SELECTION

The selection of communities for increased monitoring attention is basically a function of the growth of the area, whether it exists in a potential or a designated AQMA. The task, however, is much easier for designated AQMA's as most of the analysis of growth has already been done and the development of the maintenance measures has identified those communities which will be the primary target of the AQMP. For the purpose of this discussion, the division between potential and designated AQMA's is made to allow for this different degree of required analysis as well as providing a logical separation of the two activities which may have varying priorities to different states.

11.2.2.1 Potential AQMA's

After the designation of AQMA's at each 5-year interval, a review of those areas within each state which were not so designated should be made. This review would identify those other areas which, while not expected to exceed the NAAQS within the subsequent 10-year period, have the potential for violating one of the standards within 15 years. Such a determination would be based upon similar growth projections as used in designating the AQMA's except that the time period under evaluation would be 15 instead of 10 years. Ample consideration should also be given to those natural resource areas which, while not currently being exploited perhaps due to economic reasons, may be opened up when social or economic constraints are removed.

A region may be determined to be a potential air quality maintenance area if the projections of growth indicate that a national air quality standard will be violated within the area by 1990. For the purpose of this discussion, the term "potential AQMA" refers to not only those areas which have not yet been designated AQMA's for the current 10-year maintenance planning work but also any presently designated AQMA which may exceed a standard by 1990 for which pollutant it is not now designated.

Initial Selection of Potential AQMA's - Criteria for the initial designation of areas as potential AQMA's are whether the recent analysis for AQMA designation indicated that the area will be nearing the standard within 10 years, the area is currently undergoing a high growth rate which is expected to continue, the area is expected to undergo rapid expansion after the current 10-year period, or new and major development, either residential or industrial, is forecasted to occur after 1985. The resource development areas in the western states come under the latter criterion.

In addition, it is possible to establish boundary conditions which can serve to automatically exclude from consideration as a potential AQMA a particular area. Basically, such criteria for exclusion are functions of the current regulations controlling emissions of individual pollutants as well as expectations of normal growth rates. By pollutant, these exclusion criteria are:

1. Particulate matter: any standard metropolitan statistical area (SMSA) undergoing normal growth which is not projected to exceed any of the particulate NAAQ's in 1985.

2. Sulfur dioxide: any SMSA undergoing normal growth, not currently an AQMA for SO_2.

3. Carbon Monoxide: any SMSA undergoing normal growth.

4. Photochemical oxidants: any SMSA undergoing normal growth.

5. Nitrogen dioxide: any SMSA undergoing normal growth.

The SMSA's were chosen above for providing means for the elimination from concern of those areas that are already developed, that is areas with already defined industrial development characteristics. New growth in areas not previously monitored is more likely to cause problems of maintenance than growth in SMSA's where new, controlled sources will be replacing older, less controlled ones. In addition, the emissions from automobiles are generally expected to be leveling off around the 1985 to 1990 period with little if any increase by 1990 under normal growth conditions so that any SMSA which was not already under a maintenance plan for the motor vehicle pollutants for 1975 to 1985 will not be providing a maintenance problem for 1990.

The SMSA, alone or in its entirety may serve as a desirable geographic area for designation as potential AQMA. However, it is not always the area which may be most applicable. For instance, projections of emissions for cities, counties or townships within the SMSA may be possible to calculate, in which case it would be desirable to designate these sub-SMSA areas. In other cases, the projected growth in emissions may be expected to occur around the fringe of the SMSA, in which case the designation may be more desirable if it included that fringe area with or without the SMSA in whole or in part. Many of the potential AQMA's will be found to be areas completely apart from any SMSA's, occurring in the more rural areas where natural resources are to be developed.

In deciding upon the actual area to be designated a potential AQMA some consideration should be given to the difficulty of the eventual management of control programs within an AQMA. It is easier to designate by the names of the existing areas (political or non-political) than to delineate an area by listing roads, rivers, other topographical features, or latitude-longitude coordinates which constitute the boundaries of the area. Designation by currently defined areas, however, does not mean that the subsequent detailed analysis of the potential AQMA and possible monitoring network must apply to the entire AQMA as originally designated--the analysis and network could be restricted to selected problem areas within the AQMA.

In addition, one should be aware of possible relationships between the potential AQMA's and the areas to be chosen under the forthcoming regulations concerning significant deterioration. For instance, if the significant deterioration regulations provide that some (probably urban) areas are permitted to deteriorate up to the secondary national ambient air quality standard, these areas will probably be the same areas as the potential AQMA's. Therefore, it might be appropriate to designate an area large enough to allow for the proper amount of desired growth.

In deciding upon the particular boundaries of a potential AQMA, the following factors should be considered.

1. The potential AQMA should include all of the territory which shares a common air envelope and a common aggregation of sources. This will usually be an urbanized area plus some adjoining areas which are now undeveloped but which are expected to develop in the next 10 years or so. It may include satellite communities which are now separated from the central urbanized area but will, in 10-20 years, become part of the central urbanized area and thus share the air resource.

2. Use of areas previously designated by agencies of various kinds may have merit in that a data base may be available and a proliferation of "regions" can be avoided. Examples are regional planning areas; State designated planning areas; transportation planning areas; etc.

3. Emission control and other air conservation measures necessary to maintain air quality standards in the urbanized and developing parts of major urban centers may be quite stringent. Application of such stringent measures in isolated or undeveloped areas may not be advantageous. Thus, inclusion of large rural areas in a potential AQMA may not be desirable.

4. Design and implementation of air conservation measures will involve certain governmental agencies. Common boundary lines for potential AQMA's and one or some combination of jurisdictional areas of implementing agencies may have merit from an operational point of view.

5. Long-range transport of pollutants is a matter of concern. It is also true that if ambient air standards are maintained near an aggregation of sources, such standards will also usually be maintained at more distant locations. Therefore, it may not be necessary to include those areas on the periphery of an aggregation of sources in order to assure maintenance of standards at locations distant from the aggregation of sources.

6. The influence of topography and geography on dispersion of pollutants and on overall community growth patterns should be considered.

7. When selecting potential AQMA's, preparation of air quality projections and development of any needed monitoring networks will need to be based on presently available land use, transportation and other plans because of time constraints. It may be, however, that new general regional development plans will be prepared in the future because of air quality considerations or other reasons. The potential AQMA selection would desirably be compatable with any such future community planning activity.

A non-exhaustive list of types of areas which might be used for designation include:

AQCR's

SMSA's

Urbanized Areas

Communities

Groupings of: { Cities, Townships, Boroughs }

Potential AQMA Analysis for Community Selection - The decision for community selection within the potential AQMA may be on a best guess basis depending upon the guidelines presented later; however, for a more effective network it is recommended that an indepth investigation of the projected air quality situation be conducted. Descriptive analysis for selecting communities would proceed along the general lines described below concerning analytical procedures. The tasks to be performed here have a different purpose than those outlined above where it was only necessary to identify potential AQMA's on the basis of non-specific designation criteria. These tasks go beyond that and quantitatively evaluate the air pollution problem in each potential AQMA for the period 1975 to 1990. The tasks are:

1. Determine baseline emissions for each pollutant for which the potential AQMA was designated
   a. By source category
   b. By location as required by EPA models.
2. Identify principal sources (baseline and projected to 1990).
3. Acquire all necessary data to determine growth in emissions from 1975 to 1990 by source category and location for each pollutant. This would involve acquiring data on:
   a. Past trends
   b. Planned and projected economic and demographic growth
   c. Projected control technology
   d. Present and future regulations for new and existing sources
   e. Expected industrial and residential development.

4. Project a detailed emission inventory for 1980 to 1990 by source category for each pollutant.

5. Project 1980 to 1990 air quality using calibrated diffusion models as provided by EPA. Use these models to:

   a. Analyze the impact of indirect sources

   b. Analyze the impact of new sources.

6. Determine which parts of the potential AQMA's are problem areas and require increased monitoring attention. (A problem area is any portion of potential AQMA where the above analysis indicates any standard may be violated at any time between 1980 and 1990. If the standard is violated by 1985, the area should be designated as a real AQMA and an air quality maintenance plan should be prepared for it.)

The models, emission factors, growth projection techniques, etc., suitable for performing the needed analysis can be found in the respective guideline documents listed in the bibliography. Guidance on the analysis, as well as much of the data base and methodology applicable to each state, would also be available from the personnel involved in the current activities relating to the designation of AQMA's and the preparation of the air quality maintenance plans therefor. A short review of the considerations involved are presented below.

Emission Baseline - While air quality is the final determining factor for the selection of the community, the emission inventory, and changes therein, is the only real base from which the air quality can be projected. In order to estimate emissions between the time standards are attained and 1990, it is necessary to determine emissions at the time that standards are attained. Some state implementation plans (SIP's) contain these projections of emissions and these can be used where available. If not available, these attainment date emissions can be calculated by the method presented in the Manual for Analysis of State Implementation Plan Progress. Regulations which are currently in existence should be used to project emissions. NEDS emission data should be used as the basis for these emissions projections. Other emissions data that may be locally available may be used in lieu of or in addition to the NEDS data; such data must be entered into the NEDS at the next semiannual update. The update is to be accomplished as prescribed in APTD-1135, Guide for Compiling a Comprehensive Emissions Inventory.

Emission Projections - Using best judgement, constrained by data availability limitations, demographic-economic indicators would be selected as representative of each pollutant-source category combination. Normally these would be population, employment, and earnings by industrial category. Percentage growth rates for these indicators are determined for the interval of interest, baseline year to 1990, for application to the corresponding pollutant-source category. Detailed instructions for the development of growth factors and the projection methodology are contained in the EPA guideline document, Projecting County Emissions (Volume 7 of this series). For the purpose of community selection from potential AQMA's, the Level 1 approach presented in this document would be appropriate. Should resources permit, however, the projections could be refined using the Level 2 or Level 3 approach.

Air Quality Baseline - Several of the models presented below for use in predicting air quality require the use of air quality at the time of implementation of existing regulations. It is the major purpose of this review to define an adequate monitoring network so that later evaluation of the area for actual AQMP development will be on a firmer air quality basis. For current evaluation of potential AQMA's, however, certain assumptions may be necessary. As with emissions, the SIP's may contain projections of air quality at the time of full SIP implementation, and these air quality values can be used. For cases where air quality projections are not contained in the SIP, it may be assumed that the NAAQS will not be exceeded. Alternatively, recent (1972-1973) air quality data may be projected to 1975 and hence to 1990, making proper adjustments for growth and scheduled abatement actions.

Air Quality Projections - The projected emissions are now used to drive an air quality projection model to estimated 1980 and 1985 air quality. The following are types of models which may be used for the pollutant indicated:

- Particulate Matter - Atmospheric Diffusion Model
- Sulfur Oxides - Atmospheric Diffusion Model
- Nitrogen Oxides - Appendix J Relationships (Rollback Model for Texas and California AQMA's)
- Hydrocarbons - Appendix J Relationships (Rollback Model for Texas and California)
- Carbon Monoxide - Supplements to the Indirect Source Review guidelines (Volume 9 of this series) for new sources; rollback model described in the AQMA designation guideline (Volume 1) for existing sources.

Acceptable models for projecting air quality are described in Volume 12 of this guideline document, Applying Atmospheric Simulation Models to Air Quality Maintenance Areas . Locally available models may also be used.

Community Selection - Each portion of a potential AQMA which is projected to exceed one of the national ambient air quality standards by 1990 should be considered as a community in which an evaluation of the monitoring network is needed. Different portions which may have similar causes for the violation and are adjacent may be considered as one community. If, however, the pollutants vary or the causes of the high concentrations are different, they should be individually considered as communities.

If the projected air quality would not be adequately measured and/or represented by the current monitoring network, the network should be revised and/or expanded to provide the coverage needed. The implementation of the revised network should occur at least three years before the development of the next set of AQMP's so that the basis for the air quality projections will be well founded. Considerations appropriate to the review of the network include whether the spatial configuration of the monitors would be appropriate to that of the pollution and whether the type(s) of monitors are compatible with the pollutant(s) under consideration.

The following national ambient air quality standards should be considered in designating communities in which standards may be exceeded:

Pollutant	Primary	Secondary
Particulate matter	(a) 75 $\mu g/m^3$, annual geometric mean (b) 260 $\mu g/m^3$, second highest 24-hr average per year	150 $\mu g/m^3$, second highest 24-hr average per year
Sulfur dioxide	(a) 80 $\mu g/m^3$, annual arithmetic mean (b) 365 $\mu g/m^3$, second highest 24-hr average per year	1300 $\mu g/m^3$, second highest 3-hour average per year
Carbon monoxide	10 mg/m^3, second highest 8-hour average per year	
Photochemical oxidants	160 $\mu g/m^3$, second highest 1-hour average per year	
Nitrogen dioxide	100 $\mu g/m^3$, annual arithmetic average	

For carbon monoxide, assume that the 1-hour standard will be maintained if the 8-hour standard is maintained. As in the original SIP's, a demonstration of achieving the oxidant standard will imply that the hydrocarbon standard is also achieved.

11.2.2.2 Designated AQMA's

As areas currently designated as air quality maintenance areas are already to be receiving extensive analysis of the projected air quality situation for the development of applicable control measures, the degree of review, as discussed above for potential AQMA's, will not be necessary solely for the basis for the establishment of a monitoring network, especially the community selection thereof. In addition, the purpose of the monitoring network for designated AQMA's is different then for the potential AQMA's.

Whereas the monitoring networks in the potential AQMA's are directed at establishing a sufficient data base for future air quality maintenance planning, the design of a monitoring network in a currently designated AQMA is contingent upon the maintenance plan itself. The network must be able to monitor the air quality sufficiently to assure that the control measures implemented in the AQMP are providing the desired degree of control.

With regard to air quality maintenance planning at the community level, there are basically only two types of areas which relate directly to control measures and should therefore receive particular attention in network designing: controlled growth areas and activity centers.

Controlled Growth Areas - In the course of the preparation of the air quality maintenance plan for a currently designated AQMA, certain portions of the AQMA will be found to either be presently exceeding a NAAQS or be projected to do so by 1985. An appropriate control measure for maintenance of air quality, especially for areas presently in violation, would be to discourage or control increased growth in these areas. Growth may be controlled by doing some form of emission density zoning, refusal to provide any increase in services (e.g., sewer systems), or other regional planning measures.

The monitoring network in such an area should be re-evaluated to determine its overall adequacy and reliability. Normally, a properly designed network would already be operative in such an area.

Activity Centers - In a similar vein, many air quality maintenance plans are expected to encourage development in other areas. These "activity centers" are normally proposed either as an alternative area where growth may still occur away from the controlled growth areas or as an integral part of a regional plan which is directed at the concept of de-centralized, sub-compartmentialized services or a defined transportation system.

The selection of such activity centers as communities deserving increased monitoring attention naturally follows from the need to also control growth in these areas. Care must be exercised in promoting and distributing the growth in these communities so that the standards are not similarly exceeded in these areas. The air quality monitored by the community network plays a major role in this control by indicating trends in air quality due to recent activity and by providing a proper base from which the impact of continued development can be projected.

Other Non-Planned Communities - Most of the control measures proposed in the maintenance plan will not be solely applicable to the general growth and development of selected communities. Instead, they will be generally applicable throughout the designated AQMA (e.g., emission standards) or major portions thereof (e.g., transportation control measures for an SMSA). In addition, some measures will apply only to a couple of sources in a defined area, especially in resource development areas.

In the latter case, such defined areas are definitely communities which should be selected for a monitoring network or the evaluation of the currently operating one. The placement of the monitors would depend primarily on the sources involved and meteorological conditions, as discussed under neighborhood selection.

For the other areas throughout the AQMA which do not come under specific growth restrictions, the criteria for selecting a community as one which deserves increased monitoring attention are whether a NAAQS was projected to be exceeded by 1985 before the implementation of the control measures or a NAAQS is currently projected to be exceeded in the area, under the AQMP control measures, by 1990. If the community comes only under the second criterion, it should be monitored as described below for a potential AQMA.

11.2.3 NEIGHBORHOOD SELECTIONS

Neighborhood selection involves the further delineation of the network configuration by selecting those smaller areas within a community which are the most deserving for the locations of the monitors. Normally, this selection process operates on a scale of a quarter to a few square miles (if the community is large or the number of samplers small) and depends primarily on the type of data needed.

For the purpose of monitoring for air quality maintenance, these data needs are again different for potential AQMA's and designated AQMA's. The following discussion highlights these differences and discusses the proper considerations for the selection of a neighborhood based upon maintenance considerations.

11.2.3.1 Potential AQMA's

The primary purpose for establishing a monitoring network in a potential AQMA is the creation of a data base adequate for the eventual analysis required for the designation of the area as an AQMA and, most importantly, the preparation of an air quality maintenance plan. As the standards, and the plans for the maintenance thereof, revolve around the highest concentrations measurable in an area, this becomes the determining factor in neighborhood selection.

The analysis needed for neighborhood selection in potential AQMA's is essentially similar to that done for the selection of the community: determine base line emissions, determine growth factors, project emissions, determine projected air quality. The scale of the analysis, completed for the community selection, especially the modeling aspects, will determine the amount of additional effort that is necessary for the proper selection of the neighborhood. If the analysis was completely on a macroscale basis, it will be necessary to undertake the re-evaluation of the community on a mesoscale basis to provide a sufficient understanding of the air quality situation. If, however, enough detail is accounted for in the preliminary analysis for community selection, it would be possible to use those results for neighborhood selection.

The criterion for such a distinction is whether the analysis provided estimates of concentrations on a scale capable of allowing the selection of several high and low points within the community. This may be determined either directly from the air quality modeling or the emission densities, or indirectly via isopleths derived from one of these two.

Since the eventual strategy options must be chosen on the basis of the control of standard violations, the neighborhoods selected for monitoring sites should, theoretically, be any individual portions of the community which are projected to exceed the standards. This is perhaps even more desirable for potential AQMA's than for existing ones due to the need for eventual strategy formulation. It is expected that, more and more, the control measures selected in future air quality maintenance plans will be directed at individual problem solving. Instead of determining a control strategy which has broad application to the whole AQMA, or even major portions thereof, individual measures will be tailored to the individual problem areas. Therefore, it is necessary to have a thorough understanding of the air quality in the particular problem area of interest.

<u>General Neighborhood Selection Criteria</u> - These considerations, along with the need for an understanding of the total air quality situation in a potential AQMA, provide the basis for the following general criteria under which a neighborhood would be selected for monitoring:

1. The neighborhood is projected to exceed a NAAQS by 1990; or
2. The neighborhood contains the highest projected population of the community; or
3. The neighborhood is projected to undergo the most rapid industrial growth between 1980 and 1990; or
4. The neighborhood is projected to be that portion of the community with the highest air quality level below the standard of interest.

It must be realized that many of the above criteria rely upon knowledge of 15-year projections of occurrences within a neighborhood. While such planning is rarely accomplished at this stage, some best guess estimate of expected activity is possible.

The above criteria require that at least two neighborhoods within a community be selected for monitoring. The minimum network would apply to a community where the standard is projected to be exceeded in only one neighborhood which is also the highest, industrial growth region and the second highest projected level occurs in the highest populated area, or vice versa.

<u>Pollutant Specific Neighborhood Selection Criteria</u> - The development of a larger than minimum network would also follow along these same general guidelines; however, some consideration of the characteristics of the pollutants can be used to provide more pollutant specific guidelines. These are developed and discussed below.

Particulates - Violations of the total suspended particulate standards are most likely to be initially occurring in isolated sections of a community due to one large source or a combination of smaller ones. In addition, these are likely to occur in industrial areas, with a lesser chance of the sources being from the residential sector. As these are individual occurrences with minimal degree of correlation among them, the selection of the neighborhoods for monitoring should parallel that of the general criteria presented above.

Carbon monoxide - Projected occurrences of carbon monoxide standard violations are likely to be either due to a high growth rate in VMT in a business district or a particular type of development activity which would contain large sources of CO. While indirect sources such as interchanges and shopping centers are likely to also allow for the violation of the CO standard, such sources are seldom predictable on a more than 10-year basis.

While the development activities containing new large sources of CO would allow for the same general criteria considerations as for particulates above, the high concentrations of CO as a result of significant increases in VMT are likely to be much more dispersed over the general community and may appropriately enough involve several neighborhoods. Normally one CO monitor placed at the highest projected level of CO would suffice; however, in the few cases where the CO standard may be exceeded in an area greater than 25 square kilometers (about 10 square miles), monitors should be added at rate of $1 + \frac{A}{25}$ where A is the area of the region (in kilometers) which violates the standard. The actual selection of the neighborhoods for the location of the monitors should be done on the basis of optimizing the highest levels which would be measured and the spatial distribution of the monitors.

Photochemical oxidants - The criteria for selection of neighborhoods for monitors of photochemical oxidants would also be similar to that for CO as this is also a pollutant which comes about primarily as a result of motor vehicles. For large areas which exceed the standards, however, the degree of monitoring need not be as detailed. Current guideline documents for the preparation of an air quality maintenance plan allow for the aggregation of emissions on a county-wide basis. Therefore, for the purpose of monitoring oxidants, the number of monitoring sites could be $1 + \frac{A}{50}$. (This relationship, as the one above, is based on judgmental experience derived from the development of a sample air quality maintenance plan.) The large point sources of hydrocarbons would receive the same attention as given under the general criteria.

Nitrogen dioxide - The criteria for selection of neighborhoods for monitors of NO_2 are similar to those for photochemical oxidants. The neighborhoods, and eventually the sites selected, would be generally the same as for photochemical oxidants when monitoring NO_2 from motor vehicles.

11.2.3.2 Designated AQMA's

The selection of neighborhoods for designated AQMA's is much more structured and requires less analysis than for potential AQMA's for the same reasons as under the community selection process. The primary purpose of monitoring in a designated AQMA is to ensure that the control measures, which were deemed necessary for the maintenance of air quality, are having the desired effects. Due to different land uses within the AQMA, the impact of the control measures on the individual communities will vary. In addition, certain communities (e.g., activity centers, controlled growth areas) will have particular measures which are applicable only to the individual community. Despite these variations, certain characteristics of the different types of control measures which may be applied help to provide some delineation of the neighborhoods which should be monitored.

For the purpose of this discussion, the control measures are divided into three basic categories: emission regulations for stationary and indirect sources, transportation control measures, and land-use planning. Generally the discussion of the criteria for neighborhood selection under each of these categories is applicable to all communities; individual applications to a particular type of community are presented as needed. Since a major objective of current monitoring is the evaluation of the progress made in attaining and maintaining the desired air quality, sampling stations are often already strategically situated to facilitate evaluation of the implemented control tactics. The development of a monitoring system for a designated AQMA will often then just require the evaluation of the current monitoring system in light of the air quality maintenance plan with changes and additions made as it appears necessary. As throughout this report, any use of the criteria or discussion thereon must be tempered with the experience and characteristics within the individual regions.

Monitoring for Emission Regulations - The selection of neighborhoods for the monitoring of the impact of emission regulations depend partially upon the types of regulations imposed and the sources to which these regulations apply. The types of emission control measures which are listed in Volume 2 of the Guidelines for Air Quality Maintenance Planning are:

A. New Source Performance Standards
B. Revision of Existing SIP Control Measures
C. Phaseout or Prohibition of Emission Sources
D. Fuel Conversion
E. Energy Conservation and Utilization
F. Combination of Emission Sources
G. Special Operating Conditions
H. Stack Height Regulations
I. Control of Fugitive Dust Sources

Full explanations and discussions of these measures are given in Volume 3, Control Strategies, of the above guidelines.

For the purpose of the discussion of neighborhoods selection, it is useful to group these types of measures into three main categories: controls directed at significantly large contributors (B, C, D, E, G, H, I) and controls which are generally applicable (A, B, E, F, I). There is obviously overlap as similar types of measures can be applied to both large and small sources.

The reason for the above division lies entirely in the need for the selection of a limited number of neighborhoods due to constraints on the extent of the monitoring network. Whereas the state or local air pollution control agency is likely to be charged with the responsibility of general ambient monitoring, many large sources, especially new ones, may be required to provide their own monitoring capable of determining the adequacy of control measures.

Significantly Large Sources - Those individual sources which received special attention in the air quality maintenance plan due to their degree of contribution to the total air quality burden also deserve to have increased attention when reviewing or establishing a monitoring network. As most individual sources make their largest contribution to the ambient concentrations within a reasonably short distance from the stack, the neighborhood for location of sampling sites is likely to be the one in which the source is located or an adjoining one. The final selection of the neighborhood and the number of sites which should be monitored will depend upon the meteorology of the area, the source strength, the other sources in the area, and the population exposure potential.

These factors are best decided on a case by case basis, however, certain general comments can be made to direct the consideration given to the selection of the neighborhood.

1. Monitor(s) should be placed in the neighborhood to record the maximum likely concentration which will occur as a result of the individual source.

2. If another adjoining neighborhood is heavily populated and also receives a not insignificant contribution from the source of interest, a monitor should also be located there.

3. The above selection of the neighborhoods would rely principally upon the use of meteorological data to determine the most frequent wind direction and speed which would account for the high levels. The review of data should include the determination of possible seasonal influences on emission of contaminants and specific meteorological conditions.

4. A monitor should not be located in a neighborhood which would receive significant contributions from other nearby sources so as to confuse the interpretation of the data.

5. The use of maximum short-term (1-24 hour) concentrations for neighborhood selection allows more validity than long-term concentrations due to averaging effects.

6. Some consideration should be given to compromising the optimum location of two monitors for different pollutants to allow for the establishment of a station containing the two monitors.

7. For isolated sources, a minimum of three sampling sites is suggested. Two should be along the most frequent downwind direction(s) and one would be along the direction that is predominantly upwind.

A reasonable estimate of the distance to the maximum concentration from the source can be derived from Figure 1. The distance is a function of the effective source height and atmosphere stability; effective source height is the sum of the physical stack height and the estimate plume rise. The Briggs plume rise or other appropriate plume rise equations may be used.

It should be remembered that certain sources, i.e., those under supplementary control strategies, are required to perform their own monitoring. In addition, many other sources, especially new indirect sources, may be required to perform sufficient monitoring before and after construction to demonstrate that standards are not being exceeded as a result of the installation.

General Emission Regulation Monitoring - The selection of neighborhoods for the location of monitors for the review of the progress of the maintenance plan involves decisions regarding the distribution of samplers within the region and expected impact of the general emission regulations inherent in the plan. Such a network should consist of stations that are situated primarily to document the trends in the highest pollution levels, to measure the exposure to the population, to measure the pollution generated by the specific types of sources which are controlled under the maintenance plan, and also to allow for the monitoring of air quality in basically uncontrolled areas to provide a baseline from which to understand the fluctuations in measurements.

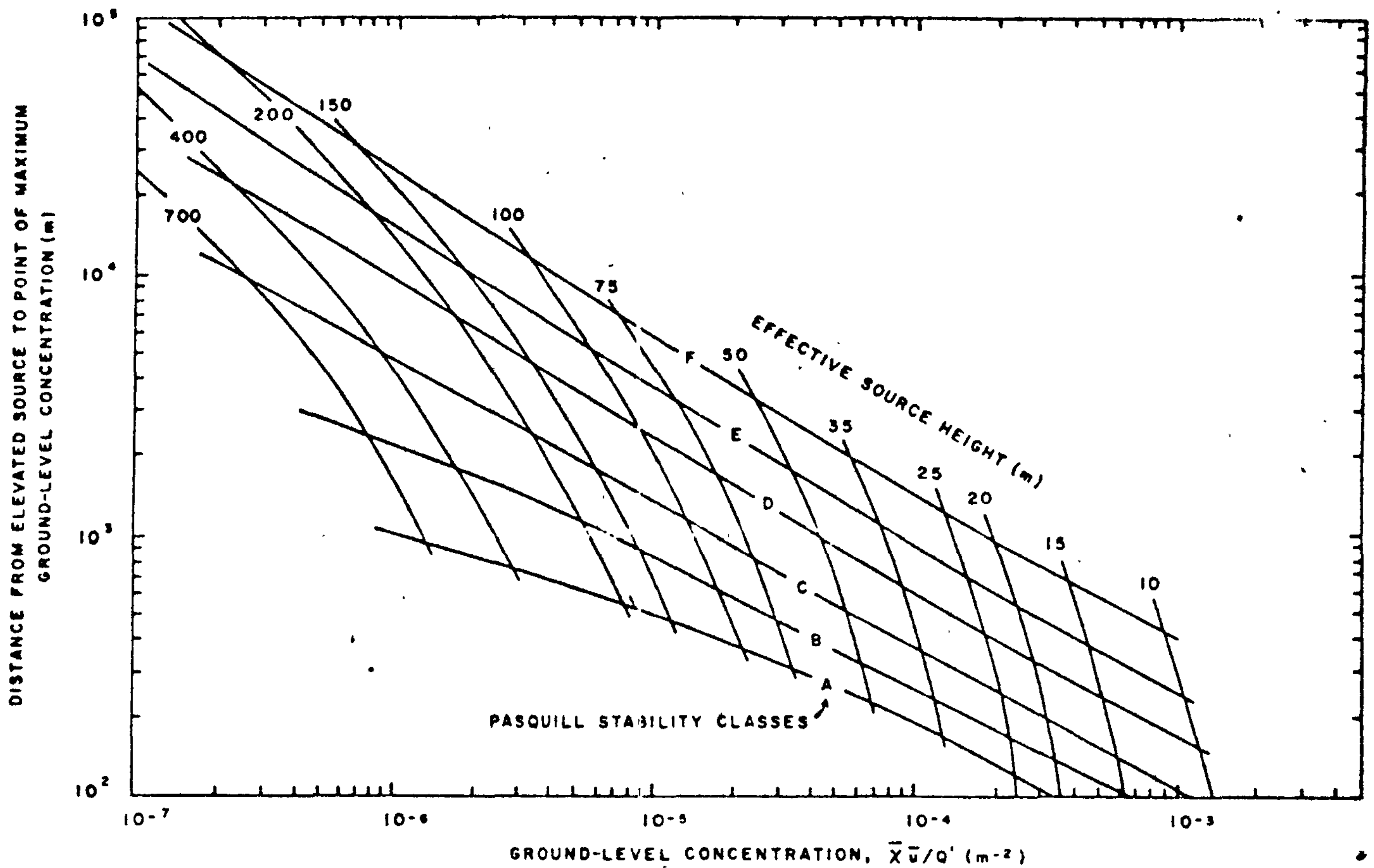


Figure 1. Estimated distance from an elevated point source to the maximum ground level concentration

The selection of the neighborhoods which takes into account the above consideration requires a fairly thorough understanding of the air quality throughout the area. It is assumed that much of this understanding will come about as a result of the re-inventorying of sources and modeling of air quality needed for the development of an air quality maintenance plan. As the plan must also provide an estimate of the expected impact of the measures formulated therein, much of the analysis for neighborhood selection will be started. Working from this basis, it is possible to develop a sampling network applicable to the monitoring of the progress of the maintenance plan.

The extra degree of analysis to be applied to the definition of the network depends upon the funding and time constraints for the establishment of the network. The extent of the network will also be contingent upon the funding constraints. Working within these constraints, the demonstration of the establishment of an adequate network should quantitatively or qualitatively be able to show that the following conditions are met:

1. Current areas with highest concentrations are being monitored.

2. Areas projected to have the highest concentrations under the AQMP are being monitored.

3. Areas which are expected to be undergoing the most rapid growth are being monitored.

4. Sections of the AQMA which have the highest population density and/or total population are being monitored.

5. At least a few areas which are not expected to have concentrations significantly affected by the regulations are being monitored. This should include at least one area with a currently low concentration (background) and one area with a moderate to high concentration of the pollutant of interest.

6. The monitors are sufficiently dispersed to provide overall coverage of the AQMA.

Monitoring for Transportation Measures - If the air quality maintenance plan for the designated AQMA contains any measures for controlling the total emissions from the transportation sector or relies on those in the state implementation plan or a subsequent transportation control plan, then the sampling network should also be geared to monitor their impact. However, where the controls on the emissions are directly on the sources themselves (e.g., FMVCP) or assume a generally applicable reduction in VMT (e.g., gas rationing, carpooling) then the monitoring of the impact of the measures would be done under the same conditions discussed above for other generally applicable emission regulations.

Those measures which, on the other hand, call for reductions and increases in VMT in particular corridors of the AQMA deserve to be monitored on an individual basis. Such measures include the establishment of a highway network which may redirect the flow of traffic or a mass transit system which would reduce the travel in particular corridors or sections of the AQMA. The monitoring of these controls should include not only air quality but also travel related parameters, such as VMT, to provide an adequate measure of the actual effect the various systems are having. A description of the network adequate to determine the changes in these transportation related parameters is beyond the scope of this guideline.

The number of samplers which are needed in any particular corridor or area depends upon the size of that area and its expected air quality. Normally it would suffice to have one monitor which is measuring the air quality in a particular corridor or area due to the broad scale type of emissions. For corridors which are axial to an area of high concentrations (e.g., a central business district), a monitor located in each corridor and one at the hub would be appropriate. Where a highway network may be directed at rerouting the traffic away from areas of high concentration (e.g., a business district by-pass), monitoring should occur both along the by-pass and in the area which is being routed around. If individual corridors are part of the transportation system only as a convenience to residents or for the control of the air quality in an area to which that corridor leads or by-passes and there is no potential problem with the air quality within the corridor itself, it would not be necessary to monitor that corridor.

The actual selection of the neighborhood for the placement of air quality samplers which will be monitoring the changes due to these transportation controls depends primarily on the neighborhoods in which these controls will be operating. For central areas which are the object of the control measures, it would, in most cases, be sufficient to monitor in the neighborhood where the concentrations are projected to be the highest. The selection of the neighborhood for monitoring of the air quality due to the corridor control measures would normally best be consistent with the midpoint of the corridor. If this midpoint places the neighborhood within three miles of the neighborhood selected for monitoring the central area due to a short corridor length, then it may be better to monitor the concentrations which correspond to the emissions at the beginning of the corridor. Monitoring at the outermost point of the corridor would also be appropriate if significant activity was expected to be generated at this point due to a mass transit system.

The monitors for measuring the air quality which results from a bypass should be in any neighborhood where the standard has the potential to be exceeded as a result of the bypass. If the standard could be violated along the bypass, then those neighborhoods with the highest potentials should be monitored. Normally the concentrations in these areas will correspond to the emissions from interchanges where maximum traffic levels will occur.

Modeling activities need not be part of the analysis in determining the neighborhoods for the location of these monitors. For carbon monoxide, the neighborhoods selected will be one of the neighborhoods in the area or corridor of interest due to the local nature of CO (e.g., the midpoint of the corridor would exactly determine the neighborhood). For nitrogen oxides and photochemical oxidants, the neighborhoods selected would be located a distance away from the actual corridor or area. The exact distance and direction would depend upon the most frequent wind direction and the average wind speed.

Monitoring for Land-Use Planning - If the air quality maintenance plan for the AQMA also incorporates certain land-use planning measures as means of maintaining the air quality, then these measures should receive some attention from the monitoring network. Land-use and planning measures as listed in Volume 2 of the Guidelines for Air Quality Maintenance Planning are:

A. Emission Allocation Procedures
B. Regional Development Planning
C. Emission Density Zoning
D. Zoning Approvals and Other Indirect Regulatory Controls
E. Transportation Controls
F. Emission Charges
G. Transfer of Emission Source Location
H. Indirect Source Review
I. Environmental Inpact Statements (EIS's)

Full explanations and discussions of these measures are given in Volume III, Control Strategies, of the above guidelines.

Transportation Controls have already been discussed above as a separate program for monitoring due to their importance and broad applicability to AQMA's. Of the remaining measures, two main categories appear: Those measures which are real planning measures (A, B, C, D, G) and those measures which are review processes for ensuring the standards will not be violated (H, I). Emission charges are not currently expected to be an acceptable measure for implementation and is therefore not considered.

This division allows for the discussion of these two separate categories each of which have different monitoring requirements. The primary difference is that the review measures are individual source oriented while the general planning measures are more community or neighborhood oriented.

General Planning Measures - Of the measures presented above which are considered actual planning measures, regional development planning is the only one which is necessarily a wide-ranging planning measure. While emission allocation procedures may well be on a community basis, this is much closer to the neighborhood type scale of the other planning measures. For these reasons, the selection of neighborhoods for the latter planning measures is much better defined than under regional development planning.

Theoretically, measures which are neighborhood oriented need to be monitored in that neighborhood or one downwind from it. Practically, however, this would require a prohibitively large network with monitors in every neighborhood in an area in which zoning was a primary control measure. This problem may be alleviated by monitoring in selected neighborhoods for a period of time and then moving the location of the monitors to other neighborhoods, or by having elaborate modeling activities which are adequate to interpolate concentrations between monitoring sites based on known air quality, meteorological conditions, and source strengths. In addition, certain areas which may be zoned for low emissions or none at all (e.g., green belt) may be excluded from the monitoring requirement though the impact of green belts would be nice to know.

For those areas for which regional development planning is used to maintain the air quality, certain activity centers and controlled growth areas, as well as a possible transportation system, are likely to be established. These measures have already been discussed under community selection and transportation controls. The selection of the neighborhoods for monitoring growth in these communities should basically be a function of the degree of growth incentives or restraints within each neighborhood.

Review Processes - Where indirect source review and environmental impact statements are used as a means of ensuring that a particular project will not cause a violation of the standard, the monitoring required is obviously source oriented. The considerations are then the same as discussed above under the monitoring for regulations which apply to significantly large sources. Particularly in this case, the monitoring can be expected to be a short term operation where the ambient air quality around the proposed source is monitored for a period before and after the construction of the source. The monitoring may also be required to be performed by the source itself.

11.2.4 SITE SELECTION AND PROBE PLACEMENT

The final selection of the site for a sampler and the exact placement of the probe rely upon the microscale influences of buildings which may exist in the neighborhood and the need for power and security for the equipment. These are problems which are independent of the considerations of monitoring needs for maintenance planning and therefore come under the same guidelines as for a regular sampling network. Such guidelines are given in Section 4 of Guidance for Air Quality Monitoring Network Design and Instrument Siting (OAQPS No. 1.2 - 012). The following discussion is basically drawn from this report.

11.2.4.1 General Considerations

In the selection of a particular site for a single sampler or a complex station, it is essential that the sampler(s) be situated to yield data representative of the location and not be unduly influenced by the immediate surroundings. Little definitive information is available concerning how air quality measurements are affected by the nearness of buildings, height from ground and the like. There are, however, general guidelines that can be provided based on operational experience:

1. Avoid sites where there are restrictions to air flow in the vicinity of the air inlet--such as adjacent to buildings, parapets, trees, etc.

2. Avoid sampling sites that are unduly influenced by down-wash from a minor local source or by reentrainment of ground dust, such as a stack located on the roof of a building where the air inlet is located close to ground level near an unpaved road. In the latter case, either elevate the sampler intake above the level of maximum ground turbulence effect or place the sampler intake away from the source of ground dust.

3. The monitoring site should be generally inaccessible to the public and should have adequate security, electricity, and plumbing.

4. Uniformity in height above ground level is desirable. Roof-top samplers should be utilized in moderate to high density areas (in terms of structures). Ground-level samplers should be utilized in low or sparse density areas (in terms of structures).

5. For CO or NO_2 monitoring, samplers should not be located in the median area of multilane highways.

11.2.4.2 Pollutant Specific Considerations

Sulfur dioxide can be considered to be rather well mixed near the ground at receptors not overly affected by specific point sources. Therefore, either ground or roof-top (one to three stories) sampling is recommended.

Similarly, TSP is usually well mixed within the first few hundred feet above the ground, but only roof-top sampling is recommended to avoid the influence of possible reentrainment of particulates close to ground level.

In contrast, the distribution of CO across a neighborhood consists of many more areas of peak levels, one at each street or major traffic center, with areas of quite low levels between. On a microscale level, the variability means that one must consider site locations for CO in scales of plus or minus a few meters, rather than plus or minus thousands of meters as might be the case with SO_2 and particulates. For the same reason, height from the ground of the air inlet is more restrictive than for the other pollutants. It is desirable, however, to sample as close as possible to the breathing zone within practical considerations (i.e., proper exposure, security from vandalism, minimizing surface effects, etc.).

The strong dependence of carbon monoxide concentration upon distance from its source was illustrated in a recent field survey. In this study it was found that the concentrations experienced by pedestrians exceeded those measured by the fixed air monitoring station, while concentrations at randomly selected locations throughout the survey grid were less than those at the monitoring station. More specifically, the data indicated that average concentrations determined by the monitor would be reduced to near the urban background level by moving the monitor approximately 200 feet farther back from the street. For peak CO sampling within street canyons, the side of the street which is opposite the side facing the roof-top-level winds is more likely to experience the highest concentrations.

These variations basically require that two alternative locations be selected, one for the 1-hour standard to which the pedestrian is exposed and another for the 8-hour standard which better represents the exposure of most individuals. In most cities, the 1-hour standard is not violated or likely to be if the 8-hour standard is maintained; therefore, such double monitoring is not normally necessary as long as the urban background level is being adequately monitored. The urban background site for CO should be utilized to measure the maximum area-wide concentrations to which the general population is exposed. Thus, either roof top or ground-level sampling in urban or suburban areas is recommended. This station should not be adjacent to major thoroughfares (not closer than 50 feet from the street curb) to rule out the influence of localized peaks due to roadway traffic.

In the case of the reactive secondary pollutants (oxidants and NO_2), the best sampling locations are, in most cases, away from the sources which emit the necessary precursors (and contribute to the reaction processes). Thus, the use of emission density and land use maps are not always helpful in determining sampling site locations. They can, however, be used in conjunction with information on the direction and magnitude of prevailing mid-morning winds to provide approximate sampling locations. There are no well established meteorological dispersion models presently available for selecting areas of expected maximum concentration for the secondary pollutants. Probably high concentration areas for these pollutants are based on: (1) available information on the reaction kinetics of atmospheric photochemical reactions involving hydrocarbons, nitrogen oxides, and oxidants; (2) the diurnal variation in pollutant concentrations; (3) the distribution of primary mobile sources of pollution; and, (4) meteorological factors.

In general, the maximum concentrations are indicated to occur between 5 and 15 miles downwind from the downtown or area of heavy traffic density. However, if the winds are light and variable, high levels may occur in the vicinity of the pollutants emissions such as the center city. The location of good NO_2 and oxidant sampling sites is a difficult process and in many cases is based largely on intuition or trial and error. The use of mobile NO_2 and oxidant samplers could be helpful in locating areas of concentration.

A minimum distance away from major traffic arteries and parking areas of 100 meters is specified for the oxidant monitoring site because NO emissions from motor vehicles consume atmospheric ozone. NO_2 is considered both as a primary stationary source pollutant and as a secondary pollutant and air monitoring stations for this pollutant should be located consistent with the respective station location guidelines. Differences in horizontal and vertical clearance distances are based on increased probability of reaction between reactive gases and vertical surfaces.

11.2.5 SITE DESCRIPTIONS

A complete and accurate description of a monitoring station, including the monitoring instruments, is extremely useful in validating, editing and interpreting ambient air quality data. The location of the station helps to establish what kind of levels are expected and how these are supposed to change over time. The identification of the instruments provides the basis upon which the reliability of the data, especially at the end points of the instrument's sensitivity, can be determined. The type of information which should be supplied for the stations and the instruments is outlined in the Air Quality Monitoring Site Description Guideline (OAQPS No. 1.2-019). The type of information called for and the specific items of the description which relate particularly to maintenance activities are discussed below.

The general description of the station, for all pollutants, should include the objective for monitoring each pollutant at that site, the type of monitoring station (mobile or stationary), the location of the station, nearby pollutant sources, and the heating and air conditioning requirements. This later information is needed because many heating and cooling systems generate one or more of the criteria pollutants. Of special interest to maintenance planning is the objective for monitoring at a particular site. Special attention needs to be paid to those monitors placed with the objective of trend analyses and of assessing the maintenance of the NAAQS. It would be helpful if the specific objective, e.g., monitoring area of controlled growth or activity center, were indicated. Some of this information may be provided by giving the location of the station.

Special information concerning the stations for continuous monitors includes the manifold design, manifold composition, and the electrical requirements. The former information helps to indicate whether the manifold will react with the constituents of the air sample so that the composition of the air in the maniford will change. The latter information is simply for determining if an adequate power supply is being received by all instruments. Other special information that should be provided concerns the actual siting positions for primary and secondary pollutant stations and the air inlets. None of this special information is felt to be especially applicable to monitoring for air quality maintenance planning.

Descriptions of the instruments are also no different for maintenance monitoring than for any other reason. The basic information which should always be provided include the identification of the manufacturer, trade name, and model of the instrument, the pollutant being monitored, the SAROAD codes used by NADB to store the data from each monitoring instrument used at each site, the measuring principle (e.g. colorimetric, nondispersive infrared, etc.), whether it is manual or instrumental, and the actual techniques used in monitoring. For continuous instruments, the performance specifications should also be given.

Once a site description for a station has been prepared and submitted to the EPA Regional Offices, it should not be considered to be invariable. Instead, care should be taken to provide periodic updates of the information included in the site description. Under normal circumstances, this should occur once a year; however, if significant changes occur (e.g. new sources or new samplers), these should be recorded immediately so that any variations in the data will not be misunderstood.

With regard to maintenance planning, changes which reflect implementation of a component of the air quality maintenance plan should be recorded. In most cases these changes occur on a scheduled basis and may be projected in the yearly site description update. Estimates, qualitative or quantitative, of the impact these control measures would have on a particular site would help in understanding the trend of the data at that site. In areas which are being monitored as potential AQMA's, it will probably be necessary to update on a yearly basis only, unless significant changes occur.

SECTION 11.3

SAMPLING INSTRUMENTATION SELECTION

11.3.1 NEED FOR CONTINUITY

Selecting the proper sampling instruments for a monitoring network which is directed toward the air quality maintenance planning activities involves the same considerations as for any other air quality monitoring network. One consideration which deserves more attention in this area, however, is the incorporation of reference methods.

Since air quality maintenance planning is a continuing process and is not a special case-study type of project which has a defined beginning and end, it is important to establish a proper sampling methodology and procedure which may be continued with few minor interruptions or variations over time. Therefore care should be taken in the selection of instruments which are to be used in the monitoring network. Only reference methods or approved alternatives, as discussed below, should be used.

Where it is necessary to use unapproved methods, due perhaps to previous acquisition of the equipment, they should be confined as much as possible to networks in potential AQMA's. It is much more likely that the stations themselves will be changed in a potential AQMA than in a designated AQMA for which an air quality maintenance plan establishes the need for monitoring for at least 10 years; in addition, the unapproved methods will probably be sufficient to establish the needed trend data for the potential AQMA. If an unapproved method is used as part of the AQMP monitoring network in a designated AQMA, it should be noted and care paid to replacing it at the earliest possible convenience.

11.3.2 APPROVED SAMPLING METHODS

A good summary of the sampling methods which may be used and discussion of those methods which are not acceptable is given in the guideline document on the Designation of Unacceptable Analytical Methods of Measurement for Criteria Pollutants (OAQPS No. 1.2-018). The following is extracted from that document to provide an understanding of the categories of sampling methods. Reference should be made to the document itself for a more thorough understanding.

11.3.2.1 Categories of Analytical Methods

Methods for measuring air pollutants fall into one of three categories: (1) approved, (2) unacceptable, and (3) those methods which are neither approved nor unacceptable (unapproved). At present, the only officially methods are the federal reference methods described in appendices to 40 CFR Part 50, originally promulgated on April 30, 1971 (36FR8186) with the National Ambient Air Quality Standards (NAAQS). This Federal Register also introduced the concept of an "equivalent method," which is any method which can be demonstrated to be "equivalent" to the reference method. Thus, unapproved methods may become approved only by demonstrating equivalence to the reference method.

Those methods designated as unacceptable are not equivalent to the reference methods because they are known to yield measurements of poor accuracy and reliability. They are considered to be obsolete. In each case, suitable analytical methods which produce measurements of greater reliability are available to replace the unacceptable methods.

11.3.2.2 Reference and Equivalency Regulations

Regulations governing the procedures and criteria by which unapproved methods may be determined to be equivalent have been proposed in the Federal Register on October 12, 1973 (38 FR 28438) as a new Part 53. Pending revision based on comments from interested persons, the new regulations, when finally promulgated, will require that a method must be tested according to prescribed procedures and meet certain prescribed specifications to be approved as an equivalent method. In essence, manual methods must demonstrate a consistent relationship to the reference method in side-by-side measurements of ambient air. Automated methods (automatic air analyzers) must demonstrate such a consistent relationship as well as meet certain performance specifications. The new regulations will also cover reference methods which are automated methods (i.e. CO and Oxidants). An analyzer must meet prescribed performance specifications before it can be determined to be approved as a reference method.

Unapproved methods must be tested according to the prescribed procedures and submitted with an application for approval to the Quality Assurance and Environmental Monitoring Laboratory (QAEML) of the National Environmental Research Center (NERC) in Research Traingle Park, North Carolina. Approved methods are to be published in the Federal Register. The regulations will apply to SO_2, CO and Oxidants corrected for SO_2 and NO_2.

11.3.2.3 Acceptability of Analytical Methods

Table 1 lists those analytical methods for which data were submitted by the States in 1972. The individual methods have been listed as "approved," "unapproved" and "unacceptable." Use of methods designated "unacceptable" should be discontinued as soon as possible. Data derived from those methods will not be accepted or used by the NADB after September 1974.

Table 1. 1973 POLLUTANT-METHOD-STATIONS SUMMARY

Pollutant	code		Method	No. of stations	Percent of total	Approved	Unapproved	Unacceptable
TSP	11101	91	Hi-Vol (FRM)[a]	3602	100	X[a]		
CO	42101	11	NDIR (FRM)	278	96	X		
		12	Coulometric	2	0			X
		21	Flame ionization	10	4		X	
				290	100			
SO_2	42401	11	Colorimetric	89	5		X	
		13	Conductimetric	108	6		X	
		14	Coulometric	172	9		X	
		15	Autometer[b]	1	0		X[b]	
		16	Flame Photometric	29	1		X	
		31	Hydrogen Peroxide[b]	38	2		X[b]	
		33	Sequential Conductimetric	6	0		X	
		91	West Gaeke-sulfamic acid (FRM)	1510	77	X		
		92	West-Gaeke bubbler	11	0			X
		93	Conductimetric bubbler	0	0			X
				1964	100			
NO_2	42602	11	Colorimetric	136	8		X	
		12	Colorimetric	14	1		X	
		13	Coulometric	10	1		X	
		14	Chemiluminescence	8	0		X	
		71	J-H bubbler (orifice)	14	1			X
		72	Saltzman	5	0			X
		84	Sodium Arsenite (orifice)	26	1		X	
		91	J-H bubbler (frit)	995	60			X
		94	Sodium Arsenite (frit)	456	28		X	
		95	TEA				X	
		96	TGS				X	
				1664	100			

Table 1 (continued). 1973 POLLUTANT-METHOD-STATIONS SUMMARY

Pollutant code			Method	No. of stations	Percent of total	Approved	Unapproved	Unacceptable
Photochemical								
O_x	44101	11	Alkaline KI instrumental	10	2			X
(Ozone)		13	Coulometric[c]	10	2		X[c]	
		14	Neut KI colorimetric	89	21		X	
		15	Coulometric	22	5		X	
		51	Phenolphthalin	3	1			X
		81	Alkaline KI bubbler	79	18			X
		82	Ferrous oxidation	91	21			X
	44201	11	Chemiluminescence (FMR)	131	30	X		
		13	Coulometric[c]	1	0		X[c]	
				436	100			

[a]FRM = Federal Reference Method.
[b]These methods should be reported under method code 42401 13.
[c]These methods should be under method code 44101 15.

For SO_2, CO and Oxidants corrected for SO_2 and NO_2, unapproved methods may be used until the Equivalency Regulations are promulgated. After promulgation of those regulations and additional approved methods become available, unapproved methods may be used only until they can be replaced with approved methods, and not later than 5 years after promulgation after which time only approved methods are to be used. For NO_2 and hydrocarbons corrected for methane, guidance for selecting adequate automated methods may be found in the forthcoming EPA Environmental Monitoring Series Document (EPA-650/4-74), Guidelines for Determining Performance Characteristics of Automated Methods for Measuring Nitrogen Dioxide and Hydrocarbons Corrected for Methane in Ambient Air.

Until these regulations and guidelines become available, the following guidance should be considered:

- TSP - The hi-vol method is the federal reference method for total suspended particulates. Since the air quality standard is defined by the method, the hi-volume sampler is the only acceptable method. No procedures for determining equivalency of alternate methods have been developed, so all other methods are to be considered unacceptable.

- Carbon Monoxide - The nondispersive infrared (NDIR) is the federal reference method for CO. Automated analyzers based on other principles have not yet been tested with respect to equivalency, and are therefore unapproved.

- Sulfur Dioxide - The manual West-Gaeke - sulfamic acid (24-hour bubbler) method is the federal reference method for SO_2. The other manual methods listed are unacceptable. The similarly named "West-Gaeke" method (SAROAD method code 42401 92) is not equivalent to the reference method (SAROAD method code 42401 91). Since no continuous method has yet been tested for equivalency, they are classified as unapproved.

- Nitrogen Dioxide - The manual NASN bubbler method is the federal reference method for NO_2. However, in the June 8, 1973 issue of the Federal Register (38 FR 15174), it was proposed that the NASN method be withdrawn as the reference method and a new one designated after testing of proposed candidate methods. Although the method was not officially withdrawn, the problems with variable collection efficiency and NO interferences are such that it must be considered unacceptable. All other methods, both manual and continuous, have been classified as unapproved.

- Photochemical Oxidants (Ozone) - The reference method for photochemical oxidants is a continuous chemiluminescent method based on the gas-phase reaction of ozone with ethylene. This method is specific for ozone. All other methods listed in Table 1 are total oxidant methods. Six of these methods for total oxidants are being designated unacceptable. While the remaining automated methods are not being designated unacceptable, strong consideration should be given to replacing them with the reference method.

- Total Hydrocarbons Corrected for Methane - This category is unique in that, while hydrocarbons corrected for methane is a criteria pollutant, the Ambient Air Quality Standard is only a guide for achieving the oxidant standard. A gas chromatographic flame ionization technique is the federal reference method for hydrocarbons corrected for methane, but this method is difficult and expensive to use. Other methods are now becoming available and, as mentioned before, guidance for selection of adequate automated methods may be obtained in the EPA Environmental Monitoring Series document (EPA-650/4/74), Guidelines for Determining Performance Characteristics of Automated Methods for Measuring NO_2 and Hydrocarbons Corrected for Methane in the Ambient Air.

SECTION 11.4

DATA PROCESSING AND SUMMARIZATION

11.4.1 DATA PROCESSING EVALUATION

Data processing, as discussed in this section, involves the basic procedures needed for ensuring a smooth flow of air quality data from the point of primary recording of the data, by an instrument or in a lab, to its ultimate storage in some format or system from which it may be later retrieved for analysis. While this is directed at the handling of the data and not the eventual analysis, which is considered in the following section, some analysis is essential to provide validation of the data entries into the storage system. As with the selection of instruments, the actual processing of the data from a monitoring network that is directed toward the air quality maintenance planning activities involves the same considerations as for any other air quality monitoring network.

The evaluation of current monitoring networks which must occur in light of the maintenance activities provides the opportunity for the similar evaluation of the data processing procedures presently being employed. The establishment or expansion of monitoring networks may be increasing the amount of data which must be handled, and the data processing systems must be able to adequately cope with any increases. In addition to evaluating the capacity of the current data processing system, the adequacy of the validation processes in the system should be determined.

This latter point, the validation of data, deserves special attention when considering air quality maintenance planning. One of the major reasons for this document is to aid in the establishment of a monitoring network which will provide air quality data adequate for a basis from which projections and long-range control strategies can be formulated. If the data is not reported accurately, due either to equipment malfunction or human error, then much of the effort expended in establishing and operating the network is wasted.

The discussion of data processing which follows, attempts to provide guidance as to how a data processing system should operate. It is directed more at the establishment of proper procedures in handling and validation than at current practices, which are considered in the Guideline "Procedures for Flow and Auditing of Air Quality Data" (OAQPS No. 1.2-013). It is recommended that these procedures be followed by the states, as they are the bodies which are responsible for the designation of AQMA's and preparation of AQMP's based on this data.

11.4.2 DATA PROCESSING AT THE STATE LEVEL

11.4.2.1 Data Handling

Data handling refers to the basic mechanics of processing monitoring data from the sensor to the data storage system and the preparation of routine summaries in one or more standard formats. Included in this category is the activity of preparing, on appropriate SAROAD data forms or in other SAROAD-compatible format, the quarterly air quality data reports to EPA that are required by regulation. The actual characteristics of the data handling system depend upon the requirements of the state. Generally, the more data that needs to be processed, the more computerization and automation will be employed. Therefore, the data handling systems can be categorized as either totally manual, semi-automatic (partially computerized), or totally automated (including telemetering of data).

It is not the intent of this document to present guidelines on the mechanics of a data handling system, especially as optimum systems are most likely different for each state. However, some generally applicable comments regarding the type of air quality data which should be handled can be made.

a. All air quality data obtained by the State, which satisfies the criteria established for monitoring network adequacy, should be processed. This includes data from all operational stations in the maintenance planning network, any other state or local network, special studies, industrial monitoring, or private citizen monitoring.

b. It is desirable that the data be representative of a consecutive 3-month period. For continuous 24-hour data, it is desirable that there be at least five data points in the quarter, with at least 2 months being reported and a minimum of two data values in the month with the least number of data values reported.

c. Data must represent an interval of 1 hour or greater - shorter interval data must be averaged over a clock hour

d. Note should be made of whether the data is coming from a site for which an updated description has been made.

e. Data must be representative of the conditions of the site for the period of time specified; modification of the environment in which the site is located must be reported on the site description forms.

f. The number of significant figures that are meaningful for a particular air quality measurement is limited by the methodology employed. To use more significant figures than is warranted by the sensitivity of the analytical procedure adds no real information and can often be misleading. Table 2 presents the suggested reporting accuracy for raw data for various pollutants. While the conventions apply to the raw data it is also useful to specify the accuracy of geometric and annual means. For simplicity, the general convention is that all means be reported to one more significant digit than the raw data.

Table 2. SUGGESTED REPORTING ACCURACY FOR RAW DATA

Pollutant	Number of decimal places	
	ug/m^3	ppm
Suspended particulate matter	0	-
Benzene soluble organic matter	1	-
Sulfates	1	-
Nitrates	1	-
Ammonium	1	-
Sulfur dioxide	0	2
Nitrogen dioxide	0	2
Nitric oxide	0	2
Carbon monoxide	1	0
Total oxidants	0	2
Total hydrocarbons	1	1
Ozone	0	3
Methane	1	1

11.4.2.2 Data Validation

The type of analysis that occurs in data processing is directed at the individual pieces of data as opposed to a set of data as discussed in the following section. Data auditing is that portion of the data processing procedure that attempts to detect and correct errors. This is not considered to occur at a distinct step in the flow of data but rather at many small steps. Data auditing has two major components: editing and validation. Editing refers primarily to the handling of the data where it is important to make sure that the proper data is recorded on the proper forms. This has already been briefly mentioned under data handling.

On the other hand, data validation refers to the actual examination of the individual data at specific steps in the processing system to determine where anomalous values may have entered. It is not generally a major aspect of manual data processing operations, due in part to the small volume of data being handled and because the data will often be processed entirely by a single individual. However, data validation is of much greater significance with semi-automatic and automatic data handling systems which involve a complex series of steps involving both human and computer operations. There are major laboratory or field operational errors, data entry or copying errors made by field or clerical personnel, bad data from malfunctioning instrumentation; all of these will likely proceed through the processing beyond the point at which they were made, to be caught and fixed at some later point, or else not to be fixed at all, leaving a number which is not what it ought to have been.

The only validation system which would have the slightest chance of catching the vast majority of mistakes of such diverse nature would be a system of 100 percent double checking of every operation. Such a system, is, however, not only prohibitively expensive, but also really impossible in the sense that the many operations in the field are not susceptible to repetition for the sake of checking.

The types of errors which are likely to occur may be divided into two groups, even though they are really just the opposite ends of the same continuum. There are small errors in data that are detectable only if the true value is known, because they are really quite close to the true value and do not seem "out of line". Then there are the gross errors, which are detectable because they produce data which deviates very much from what is expected.

An error in the former group is the more difficult to detect. Because it is never really known what the "true value" should be, a system capable of detecting small errors, e.g., a 28 in the midst of the data which should really be a 31, cannot be established. One must approach this type of error purely by prevention, by designing the sampling, analysis, and data processing procedures to minimize the chance for error, spot-checking each operation to avoid systematic errors, and then relying on the knowledge that a few random errors of small magnitude will generally "average out." Systematic errors will of course not average out.

The primary concern of procedural data validation must be those errors which are relatively large in magnitude, i.e., those that produce numbers that are "outlying" in some sense, either extremely large, extremely small, or perhaps extremely different from their neighbors. It is important to remember that the data cannot be judged by what the true value is but rather only by what it is expected to be.

What is most easily done is to set an upper limit above which all data is marked by the computer (or data handler) as possibly in error; for particulate data and some gases a lower limit is also wise. For continuous data recorded as hourly averages or less, a criterion of change between adjacent data values is also appropriate. These boundary limits, beyond which data are singled out, may be based upon the past data for that pollutant and that station, etc., setting aside a fixed percentage of the data, or they may be standard for the entire system. The former is by far preferable, being a more even-handed, rational, way of doing this operation and is discussed later, while the latter, being easier to program into a computer system, is by far more common.

Minimum Validation Process - For performing upper limit validation checks on a total network basis, the selection of the upper limits must be such that legitimate excursions above the expected are not unnecessarily flagged out. The following list illustrates some computerized hourly validation checks under consideration:

CO	100	ppm
SO_2	2	ppm
Ozone (Total Oxidant)	0.7	ppm
Total Hydrocarbons	10	ppm
Non-methane Hydrocarbons	5	ppm
NO_2	2	ppm
NO	3	ppm
NO_x	5	ppm
Total Suspended Particulate	2000	$\mu g/m^3$

In cases where the reported pollutant measurements are below the limit of detection for the analytical procedure, the reported number should be viewed as representing a range from zero to the minimum detectable. However, in order to use such data in computing annual summary statistics such as geometric means it is convenient to have a convention indicating what value should be substituted for a measurement below the minimum detectable. As a general rule, each value below the minimum detectable is replaced by a value approximately equal to one-half the minimum detectable. Table 3 indicates selected minimum detectable limits used by the National Aerometric Data Bank (NADB) for various analytical methods. A complete listing may be obtained from the

Table 3. MINIMUM DETECTABLE LIMITS FOR SELECTED MEASUREMENT TECHNIQUES

Pollutant	Collection method	Analysis method	Units	Minimum detectable
Suspended particulate	Hi-vol	Gravimetric	ug/m^3	1.0
Nitrate	Hi-vol	Reduction-diazo coupling	ug/m^3	0.05
Sulfate	Hi-vol	Colorimetric	ug/m^3	0.5
Carbon monoxide	Instrumental	Nondispensive infra-red	ug/m^3	0.575
Sulfur dioxide	Gas bubbler	West-Gaeke sulfamic acid	ug/m^3	5.0
Total oxidants	Instrumental	Colorimetric neutral KI	ug/m^3	19.6

National Air Data Branch, EPA, Research Triangle Park, N.C. 27711. The mid-point substitution was selected after examining the statistical distribution of the data. It should be noted that in comparing data over several years, a standard minimum detectable should be used unless it has changed by an order of magnitude.

In preparing summary statistics, if more than 25 percent of the observations are less than the minimum detectable no statistics are computed from the data.

Having selected certain data for further investigation (by maximum upper limit validation), the field and laboratory personnel (in the case of data from mechanized sampling equipment) or the instrumentation staff (in the case of data from automatic instrumentation) then would go back to the record cards, strip charts, or other records to either supply the correct value, if in fact an error has been detected, or to offer an explanation of why the value, though extreme, is probably correct. This second look by the sampling personnel is more often rewarding in the case of continuous instruments, where automatically-recorded data has probably been processed from the instrument to the first preliminary computer tabulation with little human scrutiny beyond the station operator's maintenance of the instruments themselves. In the case of intermittent data, the laboratory personnel have probably used their experience to verify the numbers, at least subconsciously, while they were recording the data, so that second looks are less frequently fruitful here.

A system of this type, combining computer flagging of extremes with manual verification of the flagged values, will catch only the most blatant instrument errors, laboratory errors, and clerical misrecordings. Since it is desirable to perform validation on many data points between these upper and lower extremes, which may be inapplicable to particular sites, further sophistication in this area is needed.

Desirable Validation Process - A desirable sophistication in the data validation process is the use of comparison values based on the air quality history of the specific site in question. This is generally not done at present, but should increasingly become the general practice as monitoring networks assume a permanent form and accumulate larger bodies of historical data, as is the case for maintenance planning.

Such a technique simply involves using a set of comparison values that reflect the actual air quality at the site in question, rather than values that reflect only the extremes of all possible data from all possible types and locations of sites. The comparison values listed previously, for instance, are quite extreme values, far beyond what typical urban stations would ever experience. They need to be this high, however, in order to be unquestionably above the maximum levels of those very few sites in the most extremely polluted locations. Thus, for most stations, their use would be appropriate only for automatically discarding values as obvious extreme errors, and not for selecting out the few highest values for study and validation.

This difficulty is avoided by the use of a set of comparison values specific to each site. These can take into consideration not only the general level of the air quality at the site, but also seasonal variation patterns and, in the case of continuous stations, daily and weekly variation patterns. This ability to take into consideration the regular variations of air quality over time is very important in computerizing such processes. As an example, a human data analyst would react quite differently to an abrupt change in CO levels if it occurred between 6 and 7 a.m. at the time of the morning rush hour than if it occurred between 2 and 3 a.m. when CO levels are typically stable. In designing a system to computerize these judgments, the comparison values must either be adjustable for different hours of the day, or else must be so large to accommodate the morning peak that they could miss significant smaller changes at other times.

Examples of the types of tests then can be used to identify potentially anomalous values as a step in the addition of the new data to the file are given below:

- Values that are larger than the arithmetic mean of the data by some preassigned factor (such as 2).

- Values that are some factor, say 1.5 times larger than the 99th percentile of the observed data

- Hourly values that differ from adjacent values by more than some preassigned ratio, suggesting some abrupt change in baseline or a transient interference.

- Chebyschev type tests, wherein values that are more than four standard deviations away from the mean are to be considered suspect.

- Detection of any values that are larger by some factor (e.g., 1.5) than the theoretical expected value of the 99th percentile of the distribution under question.

- The finding that the average of $K \geq 5$ successive values falls outside the $(\mu \pm \frac{3\sigma}{\sqrt{K}})$ limit, where μ and σ^2 are the mean and variance, respectively, of the distribution under question.

The difference in these types of tests should be noted. In the first four, the assigned percentile is estimated from the data, whereas in the latter two it is theoretically obtained. The sensitivity of these latter tests can be determined analytically from the frequency distribution.

11.4.2.3 <u>Data Summarization</u>

<u>Characteristic Data Patterns</u> - Before summarizing any data, some thought must be given to the characteristics of the raw data. This is particularly true for pollutants that have strong seasonal and diurnal patterns that will affect the interpretation of the data. For example, the maximum hourly oxidant value for a year based on 4,000 observations could have completely different meanings, depending upon whether the

observations were made primarily during the winter or the summer. This section presents examples of some of these patterns. The study of these patterns can frequently be an important analysis in itself, since they usually provide important insight into the behavior of the pollutant. An awareness of these patterns also provides a means for screening the data for anomalous values. It should be noted that while the following discussion is general in nature, the characteristic pattern at a given site is a function of local emissions and meteorology and, as a consequence, characteristic patterns may be specific to that site or locality.

Seasonal patterns - Figure 2 displays graphs of monthly averages for various pollutants at a particular site. Superimposed on these graphs is a smoothed curve representing the seasonal patterns in the data. Although the intensity of the seasonal pattern for a particular pollutant may vary from site to site depending upon local factors, the qualitative patterns are generally consistent. Oxidants tend to have seasonal maxima in the summer, while the other gaseous pollutants usually have maxima in the winter. SO_2 usually has a more distinct pattern than in this example, though particulate levels are frequently as erratic as here. A knowledge of the seasonality of a pollutant can provide useful information for interpreting the data since it suggests the season in which maximum concentrations would be expected.

Diurnal patterns - In addition to seasonal patterns some pollutants also have pronounced diurnal patterns. These patterns may be due to factors such as solar radiation, traffic density, etc., which influence pollution levels.

Table 4 summarizes the 1971 oxidant data for the Downtown Los Angeles site operated by Los Angeles County Air Pollution Control District. The number of times that the national oxidant standard was exceeded is presented by month and hour of the day. The marginal totals indicate both the diurnal pattern and the seasonal pattern.

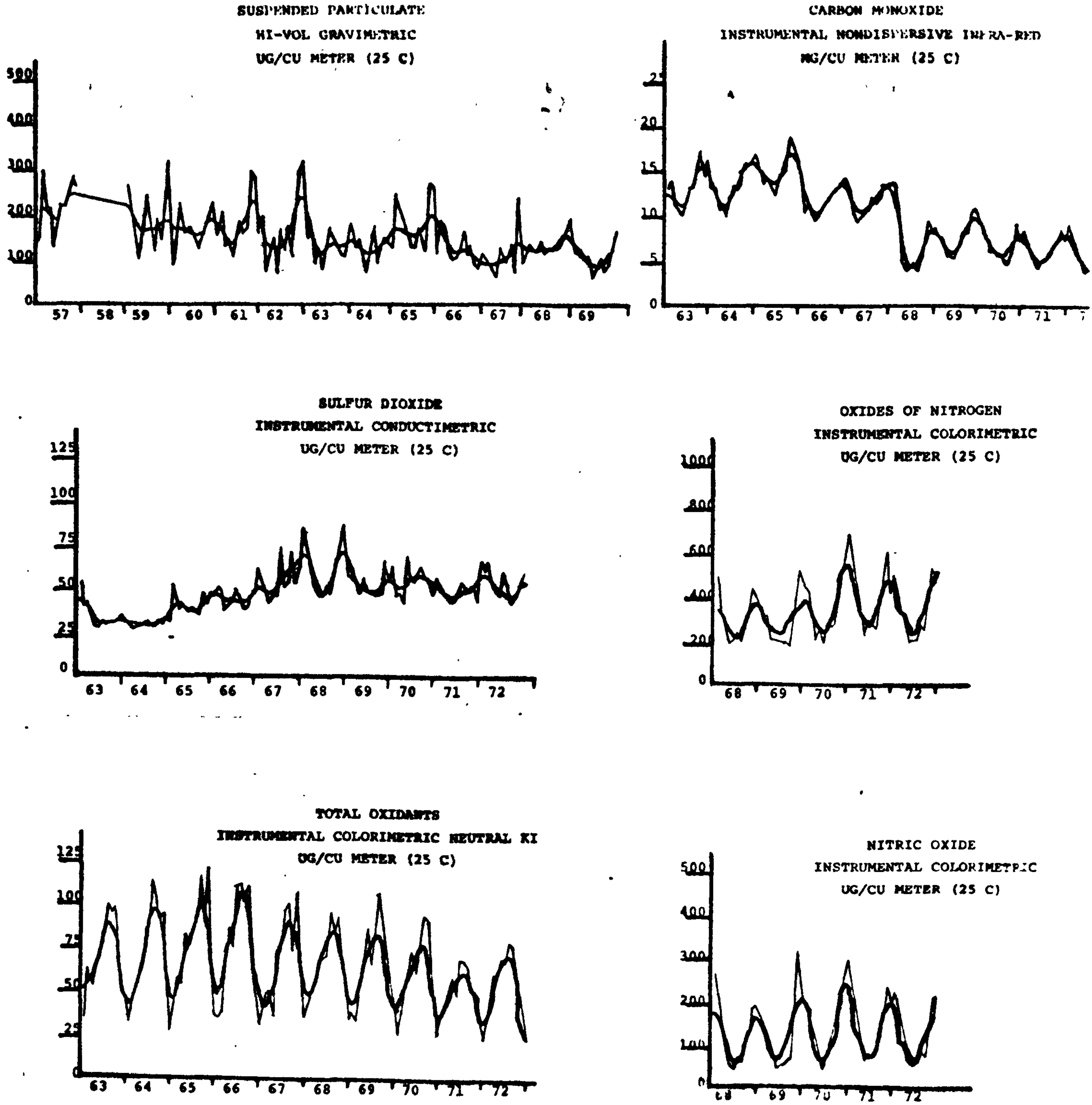


Figure 2. Graphs of seasonal patterns for various pollutants at a particular site

Table 4. NUMBER OF HOURS ABOVE OXIDANT STANDARD BY MONTH AND TIME OF DAY (1971 DATA)

	M	1	2	3	4	5	6	7	8	9	10	11	N	1	2	3	4	5	6	7	8	9	10	11	Total by month
Jan													1	2	2	3									8
Feb												1	4	4	4	3									16
Mar										1	1	1	3	3	2	1									12
Apr											4	6	8	8	7	7	3	1							44
May												3	4	4	3	1	1								16
Jun									1	2	9	9	12	12	11	6	2	1							65
Jul										2	13	19	18	15	11	4		1							83
Aug										2	8	17	16	16	7	3	1								70
Sept										3	6	10	10	10	6	1									46
Oct											2	7	5	9	6	2									31
Nov													1	1											2
Dec																									0
Total by hour									1	10	43	73	82	84	59	31	7	3							393

Frequency distributions - One characteristic pattern of air quality data that is particularly important becomes apparent after examining some frequency distributions. Many quantities are assumed to have a distribution that is symmetric about the average, such as the normal distribution. Figure 3 shows the frequency distribution of total suspended particulate data from Philadelphia. It is apparent that this distribution is not symmetric. However, Figure 4 shows the frequency distribution for the logs of this same data. The distribution is more symmetric and can be better approximated by a normal curve. Data having this property is said to be log-normally distributed, and this is a common assumption regarding air quality data. [1,2]

Preparing Data Summaries - In planning or preparing a summary of air quality data, perhaps the most important step is to first define the purpose of the summary. The usual use of these summaries is to simply describe typical and peak levels. This section discusses several basic statistics that can be used for this purpose. The first two subsections discuss the treatment of typical and peak values. The third discusses the range of the data.

Indicating typical values - This section discusses the arithmetic mean, the median, and the geometric mean as indicators of typical values. The geometric mean and the median are frequently used in air pollution studies because of certain properties of the log-normal distribution. In choosing the appropriate statistic, the purpose of the summary must be considered. While all three may indicate typical values, if the purpose of the summary is to compare the data to the National Ambient Air Quality Standards, then the standard suggests the appropriate statistic. Another statistic commonly used to indicate typical values is the mode, which is the value that occurs most frequently. The use of the mode is not discussed here since it is usually of little value in summarizing air quality data. For example, the mode for oxidant would be near the minimum detectable due to the large number of very low values throughout the night.

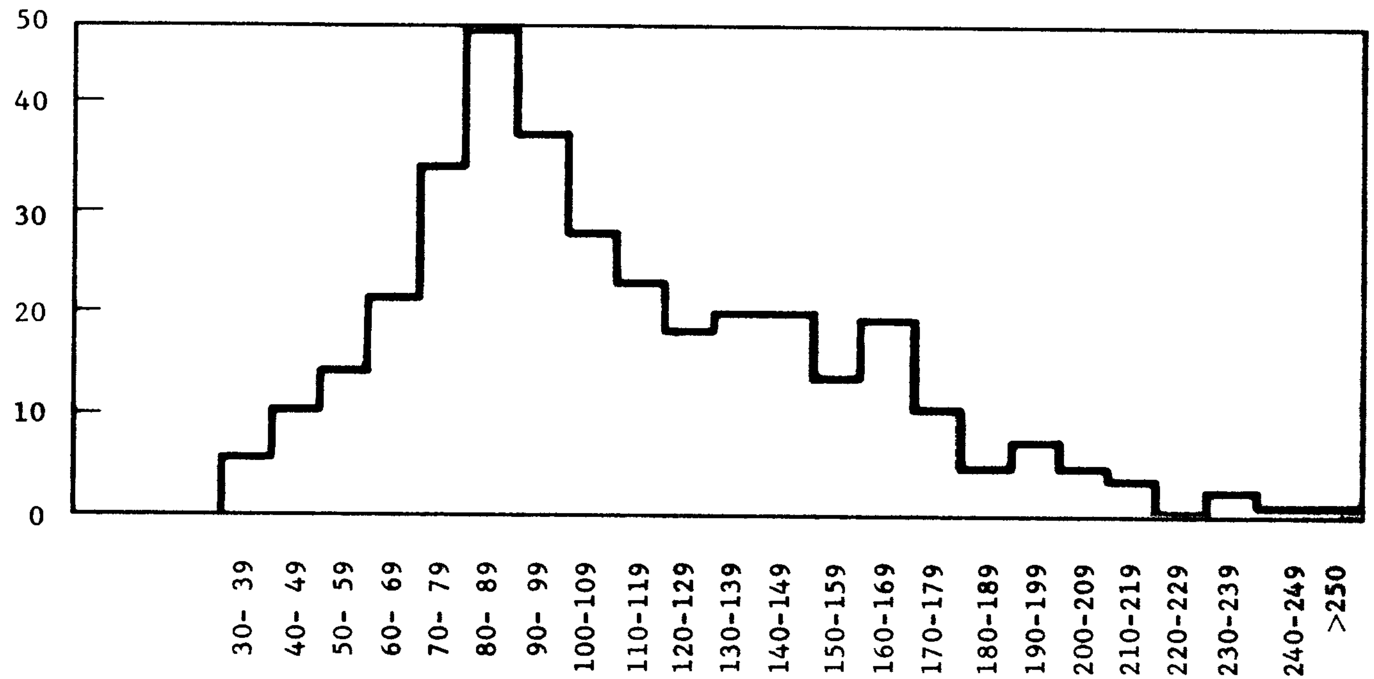


Figure 3. Frequency distribution - TSP (Philadelphia-1969)

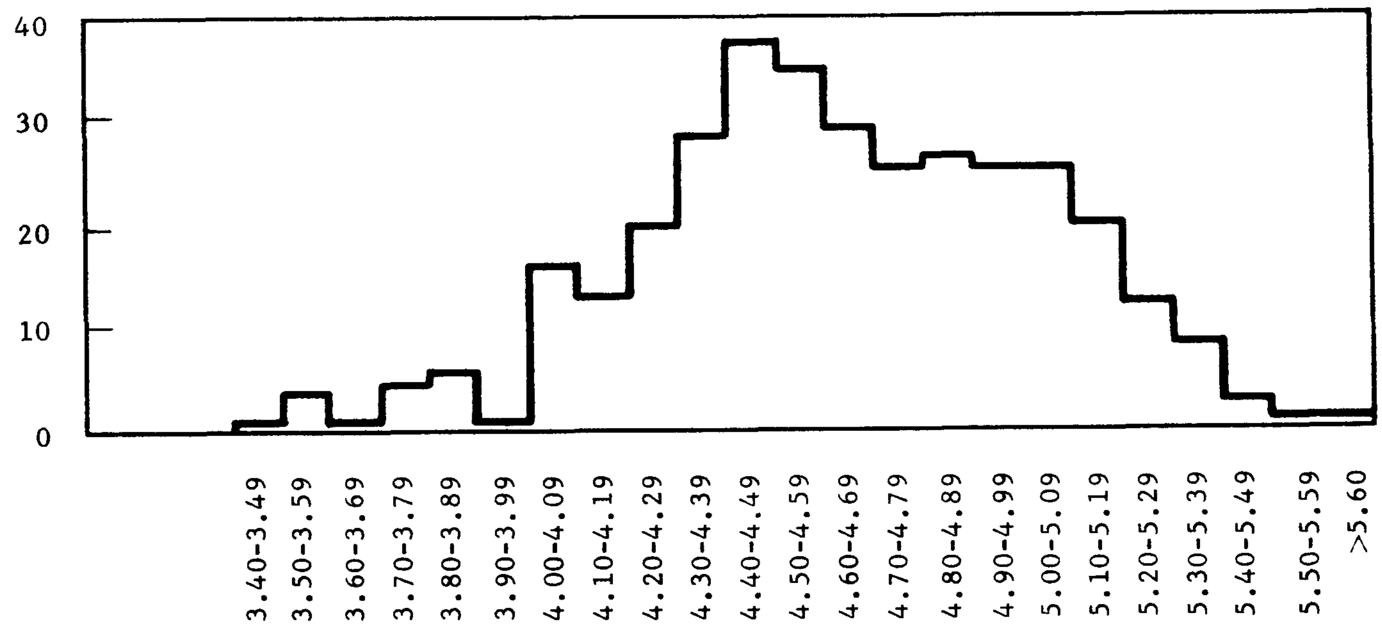


Figure 4. Frequency distribution - log of TSP data (Philadelphia-1969)

Arithmetic Mean

Given a set of n observations, say $X_1, X_2, \ldots, X_n$, the arithmetic mean is simply

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

When the term "average" is used the arithmetic mean is usually what is meant.

Median

The median is the middle value of the data, that is, the value that has half the data above and half below. If the data is ranked in order of magnitude so that

$X_1 \le X_2 \cdots \le X_n$, then the median is $X_{\frac{n+1}{2}}$ if n is odd, and

$\left(\frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}\right)$ if n is even.

The median is a convenient statistic that is not influenced by changes in the extremely high or low values of the distributions, as would be the arithmetic mean.

Geometric Mean

Given a set of n observations, say $X_1, X_2, \ldots, X_n$, the geometric mean is $g = (X_1 \cdot X_2 \ldots X_n)^{1/n}$. Since this probably is the least intuitive of the statistics presented, it is worthwhile to discuss it in more detail.

If a distribution is symmetric, such as the normal distribution, then the expected value of the arithmetic mean and median are identical. However, for a log-normally distributed variable, it is the expected value of the geometric mean that approximates the expected value of the median. Therefore, since air pollutants often have a distribution that is approximately log-normal, the geometric mean has become commonly used as a convenient method of summarizing the data; for total suspended particulate matter, the annual standards are expressed as geometric means.

As an alternate computational formula, it should be noted that

$$\log g = \frac{1}{n} \sum_{i=1}^{M} \log x_i \text{ or } g = \text{EXP}\left\{\frac{1}{n} \sum_{i=1}^{n} \log x_i\right\}.$$

Indicating maximum values - As in the previous section, the purpose of the summary is the critical factor in determining the appropriate statistic. Maximum values may be indicated by listing the maximum and/or the second highest value. The second highest value is important because compliance with the short-term air quality standards is determined by this value. However, there are other statistics that are useful for indicating maximum values. The principal difficulty in using the second highest value is that it does not allow for differences in sample sizes. For example, if two monitoring devices are side by side and one operates every day of the year, while the other operates only every 6th day, it would be expected that the second highest value for the everyday sampler would be higher than that for the other, even though they both monitored the same air. Table 5 illustrates how the second high value may vary depending upon different sampling frequencies based upon total suspended particulate data from a Philadelphia site that sampled daily.

To allow for this dependence upon sample size, various percentiles are sometimes used to indicate maximum values. For example, the 99th percentile might be used for hourly data, while the 90th might be appropriate for daily measurements. By using a percentile value rather than an absolute count of samples, allowance is made for sampling frequencies that differ from site to site and year to year. Table 6 indicates the 90th percentile for the sampling schedules used in Table 5.

Table 5. MAXIMUM AND SECOND HIGH VALUES (PHILADELPHIA-1969) FOR VARIOUS SAMPLING SCHEMES

Sampling Schedule	Observations	Maximum	Second Highest
Everyday	365	325	244
Every Sixth Day	61	219	215
"	61	195	171
"	61	244	238
"	61	215	211
"	61	325	234
"	60	239	205
Every Fifteenth Day	25	205	176
"	25	325	207
"	25	239	191
"	25	219	196
"	25	234	165
"	24	201	198
"	24	215	211
"	24	195	183
"	24	188	173
"	24	195	169
"	24	160	154
"	24	244	199
"	24	215	201
"	24	179	171
"	24	238	205

Table 6. GEOMETRIC MEANS, MEDIANS, AND 90TH PERCENTILE VALUES FOR SAMPLING DATA OF TABLE 5.

Sampling Schedule	Observations	Geometric Mean	Median	90th Percentile
Everyday	365	102.6	97	171
Every Sixth Day	61	99.8	105	162
"	61	95.2	93	155
"	61	113.6	113	188
"	61	107.2	101	177
"	61	106.4	105	171
"	60	94.7	94	158
Every Fifteenth Day	25	100.2	111	175
"	25	114.6	121	178
"	25	125.0	130	189
"	25	104.9	96	192
"	25	100.8	105	148
"	24	99.8	90	190
"	24	104.4	98	177
"	24	102.4	99	171
"	24	92.1	95	143
"	24	100.8	96	162
"	24	92.0	88	140
"	24	104.6	97	186
"	24	107.2	109	173
"	24	94.1	94	162
"	24	99.6	98	165

Indicators of spread - In addition to an indication of typical and peak values, it is also frequently desirable to have a measure of how variable the data is - did it fluctuate widely or were all values fairly uniform? The customary statistics for this purpose are the arithmetic standard deviation and the geometric standard deviation, used in conjunction with the arithmetic and geometric means respectively. Ranges or percentiles could also be used depending upon the desired use of the summary, but they are not discussed. The basic formulas for the arithmetic and the geometric standard deviations are given below.

Let $X_1, X_2, \ldots, X_n$ be a set of n observations.

Then the arithmetic standard deviation is:

$$s = \left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right]^{1/2} \quad \text{where } \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

and the geometric standard deviation is

$$s_g = \text{EXP} \left[\frac{1}{n} \sum_{i=1}^{n} (\ln X_i - \ln g)^2 \right]^{1/2}$$

where g is the geometric mean.

11.4.3. SUBMITTING DATA TO THE NADB

Beyond the data processing and validation the states conduct for their own purposes, they have a responsibility to provide the data to EPA for entry into the National Aerometric Data Bank. This subsection briefly describes the procedures used by EPA to accept and process this data. Although these procedures include some data screening and validation, it must be emphasized that these efforts can <u>in no way</u> be viewed as supplanting any portion of the data validation effort required of the states. The major contributions to quality control of air quality data handling must necessarily be preventative, and there is no choice other than the states' conducting these efforts. The

EPA procedures are designed only to protect the integrity of the data bank; they cannot really contribute to the quality of the data for the states' use.

The procedures used in entering data are at present (mid-1974) still settling into ultimate form, as the new data reporting requirements are implemented the first few times. During this transition, it is expected that the EPA Regional Offices will be assuming increasing responsibility with respect to the screening and validation of data before it is entered into the data bank.

The initial steps involved in submitting data are:

1. The State agency submits air quality data to the appropriate EPA Regional Office as part of the State Implementation Plan reporting procedures. These reports, which are forwarded on a quarterly basis, contain both the air quality data and any new site descriptions for the State's air monitoring station The data may be sent in more frequently than quarterly if desired, but must be submitted to the Regional Office in SAROAD format on either coding forms, punched cards, or magnetic tape. Data for all operational stations as described in the SIP's, beginning with that used in plan preparation, must be submitted. It is strongly encouraged that all reliable data satisfying the criteria established for monitoring network adequacy also be submitted, including data from the stations established for air quality maintenance purposes.

2. The NEDS/SAROAD contact in the Regional Office arranges for keypunching of forms, if necessary, and then mails the data to the MDAD's National Air Data Branch in card or tape form.

3. Air Quality data submitted to the National Air Data Branch should have the following characteristics:

    a. Data must be coded in SAROAD format.

    b. Data values less than the monitoring minimum detectable sensitivity should be reported as a "zero" value.

c. It is desirable that the data be representative of a consecutive 3-month period for which at least 75 percent of the data values are valid. (Values below the minimum detectable sensitivity are considered valid.) However, if the validity criteria are not met, the data should still be submitted, particularly for evaluation of maximum value standards. For intermittent 24-hour data, there should be at least five data points in the quarter, with at least 2 months being reported and a minimum of two data points in the month with the least number of data value reported.

d. Data must represent an interval of 1-hour or greater -- shorter interval data must be averaged over a clock hour.

e. Data must be representative of the conditions of the site for the period of time specified; modification of the environment in which the site is located must be reported to the MDAD by the State and/or the Regional Office.

4. Data will then be processed using the SAROAD edit program, and investigation and correction of potential errors accomplished by the Regional Office in conjunction with the state involved.

5. Corrected data are submitted to the NADB for file updating.

11.4.4 REFERENCES

1. Larsen, R. I., "A Mathematical Model for Relating Air Quality Measurements to Air Quality Standards," U. S. Environmental Protection Agency, Office of Air Programs, Research Triangle Park, North Carolina, OAP Publication No. AP-89, 1971.

2. Hunt, W. F., Jr., "The Precision Associated with the Frequency of Log-Normally Distributed Air Pollutant Measurements," J. Air Poll. Control Assoc., Vol. 22, No. 9, p.687, 1972.

SECTION 11.5

DATA ANALYSIS

The ultimate purpose of monitoring is of course the use of the data gathered. The analysis of data gathered for air quality maintenance purposes utilizes the same general approaches and statistical techniques that are applicable to analysis of any data, although the emphases and some of the interpretive conventions are necessarily different. This section considers the conventions involved in interpreting data with respect to the NAAQS in the context of maintenance planning and analysis, the general thrusts and emphases of data analysis required by the maintenance planning process, and finally, some of the analysis techniques of primary utility in such efforts.

11.5.1 DATA INTERPRETATION CONVENTIONS

Because the National Ambient Air Quality Standards (NAAQS) and other promulgated rules and regulations are usually expressed in results-oriented phrasing, there have arisen specific questions about the detailed interpretation of data with respect to the NAAQS or other regulations. Many of these questions have been addressed in the EPA Guideline Document "Guidelines for the Interpretation of Air Quality Standards," (OAQPS No. 1.2-008), and in a recent publication by Curran and Hunt.[1] Some of these raise substantive policy issues, while others are more a matter of simply assuring uniformity, and some arise primarily in programming computerized analysis routines, where questions that are normally a matter of human judgment must be quantified in advance. Guidance on these issues has been provided in the past for use in an implementation planning framework; the discussion herein concerns these issues in an air quality maintenance context.

11.5.1.1 Geographic Aggregation of Air Quality Data

In using and studying air quality data, it is possible to either consider each monitoring site separately or to average the data over a specific geographic area. When considering the data in relation to the NAAQS, each monitoring site in an AQCR must be considered individually in determining whether or not the AQCR is in violation of the standards. This policy applies not only in the process of implementation planning for the achievement of the standard, but also in evaluating the maintenance of the standards over the years.

The basic reason behind this policy is simply that the NAAQS were defined to protect human health and welfare. The presence of one monitoring site within an AQCR violating any given standard indicates that receptors are being exposed to possibly harmful pollutant concentrations. This question arises most often when the concentrations in excess of the standards values at a single monitoring station result from the effect of a small, nearby source that is insignificant in terms of the total emission inventory, or when the station in violation is so located that the probability of individuals being exposed for prolonged periods is negligible. Such circumstances do not mitigate the stated interpretation, since NAAQS are generally interpreted as being set to protect health and welfare regardless of the population density. Although air quality improvement should be most stressed in areas of maximum concentrations and areas of highest population exposure, the goal must be ultimately achieving and maintaining the standards in all locales. Data from monitoring sites are the only available measure of air quality and must be accepted at face value. Attention is thus focused on the selection of monitoring sites in terms of the representativeness of the air they sample. This is discussed in more detail in the guideline series document entitled "Guidance for Air Quality Monitoring Network Design and Instrument Siting," (OAQPS No. 1.2-012). Good station siting should minimize this type of problem, and consideration should be given to the relocation of monitoring stations that do not meet the guideline criteria.

In the case of projecting air quality for purposes of designing an air quality maintenance plan, however, the aggregation of data within homogeneous geographical areas is frequently necessary and desirable. Since emissions are necessarily aggregated and projected for areas as large as counties or larger, perhaps for an entire AQMA or potential AQMA, it is only sensible to project the air quality on a scale neither much finer nor much coarser than the emissions data with which it will be related, so long as the air quality is sufficiently uniform that an average of several stations represents the area with sufficient accuracy for 10- and 15-year projection purposes.

This degree of uniformity can be recognized in situations where no single site is consistently, year after year, significantly higher than the others, but rather where various sites are roughly similar in level, with the relative magnitude of the levels at the various sites changing randomly from year to year or season to season because of minor meteorological and source activity factors. This would be experienced, for instance, in an AQMA or potential AQMA that consists of a number of towns or small cities within a large relatively undeveloped county.

If, in contrast, one of the towns or cities were consistently higher than the others, its data should not be averaged in with other towns; rather, the reason for its higher levels should be sought, separate emission projections should be made, and a control strategy tailored to its particular problem should be developed. The extreme example of this latter situation is of course an AQMA including both a major urban area and its outlying areas. In this case, the air quality data should not be aggregated for the center city and the outlying areas any more than one should expect that the same control strategy could be uniformly applied over such disparities of conditions. In such cases, fortunately, it is often possible to make separate projections of emissions, etc., for such sub-areas because more finely-divided data bases are available.

11.5.1.2 Frequency of Violation of 24-hour Maximum Standards

The chances of detecting violations of 24-hour maximum standards depend considerably upon the frequency with which the samplers are operated. In view of this, there are questions about how data obtained from intermittent monitoring should be interpreted.

Ideally, continuous monitoring of all pollutants would be conducted. However, except for those pollutants specified in Federal regulations, EPA does not currently require continuous monitoring. Thus, one is left with either (1) predictive equations employing data from partial annual coverage, or (2) the data collected through partial annual coverage. It has been EPA policy that noncompliance will not be declared based on predicted frequencies, because the accuracy of predictive equations is not well established. However, the detection of excursions over the short-term standards should be a major consideration in determining the sampling frequency. For assistance in making these judgements, the following table gives the probabilities of sampling on two or more days on which excursions have occurred for different numbers of actual excursions above the standard and different sampling frequencies. The underlying assumption in determining these probabilities is that excursions above the standard occur randomly over the days of the year. This is, of course, an oversimplification, but is sufficient for the purposes of this discussion.

11.5.1.3 The Use of Running Averages for Short-Term Standards

The NAAQS for CO and SO_2 include 8-hour and 3-hour averages, respectively. With continuous monitoring data, there are various ways these averages could be defined and the second highest average chosen. For

Table 7. PROBABILITY OF SELECTING TWO OR MORE DAYS WHEN SITE IS ABOVE STANDARD

Actual number of excursions	Sampling frequency - days per year 61/365	122/365	183/365
2	0.03	0.11	0.25
4	0.13	0.41	0.69
6	0.26	0.65	0.89
8	0.40	0.81	0.96
10	0.52	0.90	0.99
12	0.62	0.95	0.99
14	0.71	0.97	0.99
16	0.78	0.98	0.99
18	0.83	0.99	0.99
20	0.87	0.99	0.99
22	0.91	0.99	0.99
24	0.93	0.99	0.99
26	0.95	0.99	0.99

maintenance planning purposes as well as implementation planning, compliance with these standards should be judged on the basis of running averages starting at each clock-hour. The second highest average should be determined so that there is one other non-overlapping value that is at least as high as the second highest value. Although this seems relatively straightforward, the following discussion indicates some of the subtleties involved.

The use of running averages to determine compliance with specific air quality standards necessitates that the number of values above the standard be evaluated on the basis of non-overlapping time periods. That is, any two values above the standard must be distinct and not have any common hours. This can be achieved by a relatively straightforward counting procedure. For example, in the case of CO, an 8-hour average can be associated with each clock hour of the calendar year. Then values above the 8-hour standard are counted sequentially beginning with the first 8-hour average of the year. Each time a violation is counted, the next seven 8-hour values are ignored, and the counting procedure resumes with the eighth 8-hour average. This counting procedure results in the maximum number of non-overlapping violations of the 8-hour standard.

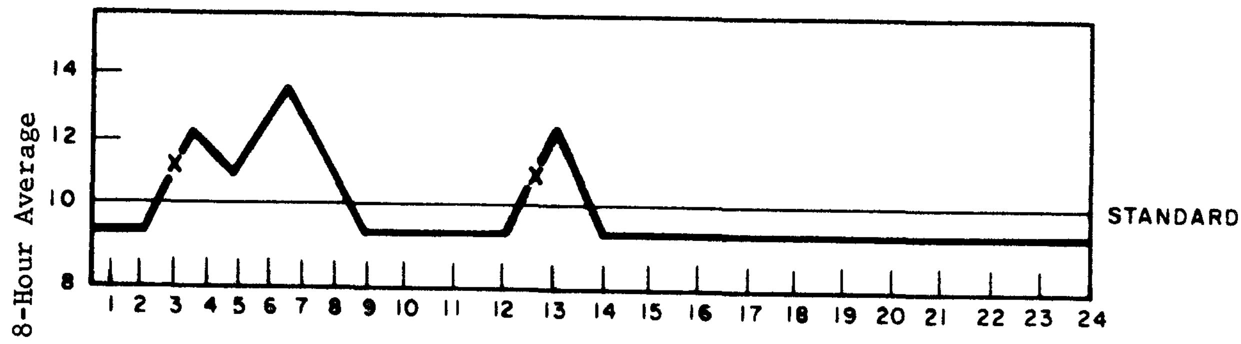
This count is all that is needed to evaluate compliance with the 8-hour standard because the standard is not to be exceeded more than once per year; and, therefore, any count value greater than one is sufficient to indicate non-compliance. However, for maintenance planning purposes, it is also desirable to employ the second highest 8-hour average to indicate the magnitude of the problem. There are several ways to define the second highest value, and three possible definitions will be indicated here in order to briefly discuss their consistency with the counting procedure described above. The three definitions considered for the second highest value are: (1) the second highest 8-hour value of those counted as being above the standard, (2) the second highest 8-hour value that does not overlap the maximum 8-hour value, and (3) the maximum second highest non-overlapping 8-hour average.

Annotated graphs of 8-hour CO are used to facilitate the discussion of the consequences of each definition. For example, Figure 5 illustrates that the first definition underestimates the magnitude of the problem because the counting procedure may count the first time the standard is exceeded and bypass the peak values. Therefore this definition is inadequate.

Although the second definition is intuitively appealing, Figure 6 illustrates that in some cases there could be two violations of the standard, and yet the second highest value that does not overlap the maximum is less than the standard. This can only occur in marginal cases in which the standard is only exceeded during one 15-hour period in the year and that the maximum value occurs in the middle of this interval. Figure 7 shows another case in which this definition produces the peculiarity that a higher CO value may lower the second highest value.

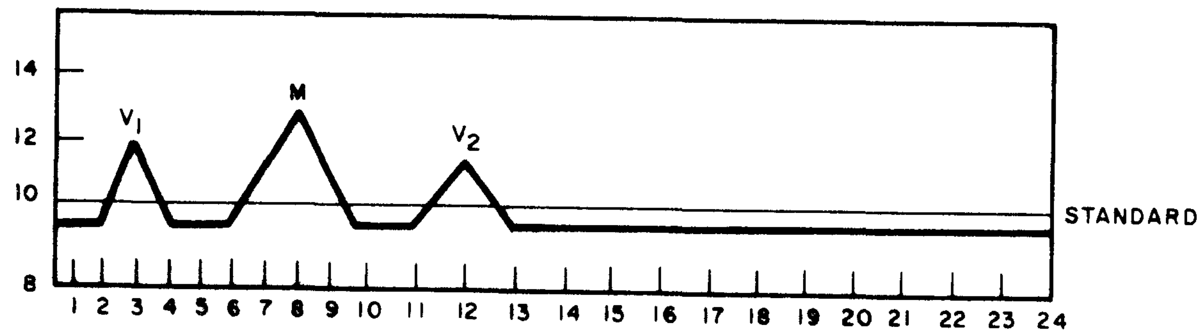
In order to avoid these inconsistencies, it becomes necessary to define the second highest value as the maximum second highest non-overlapping value. What this means is that there is one 8-hour value that is greater than or equal to the maximum second highest value and that these two values are not overlapping. It is important to recognize that the maximum second highest value may overlap the maximum 8-hour value. However, as shown in Figure 6, there is still one other 8-hour non-overlapping value that exceeds the maximum second high.

With these subtleties in mind, it is considered appropriate to use the maximum second highest non-overlapping value as the second high. In this way, the magnitude of the problem is properly assessed; for quantitative planning the second high value is also always consistent with the number of violations. This definition of the second highest value is also consistent with the approach used in determining control strategies on the basis of the roll-back equation. It is this maximum



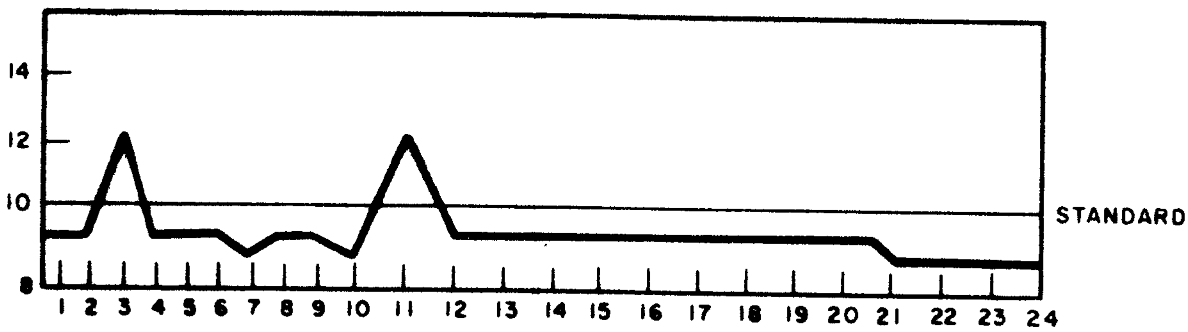
Using the counting procedure, the violations are counted at hours 3 and 12, as indicated by the x's. Note that the peak values do not occur at these points.

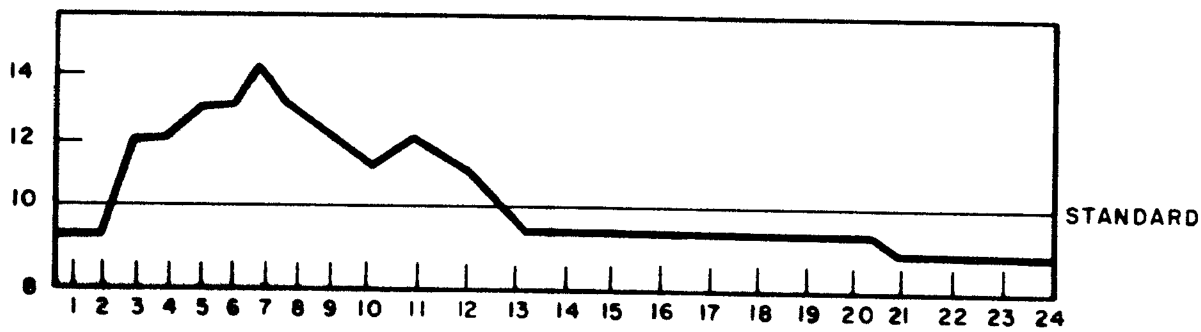
Figure 5. Eight-hour average violations determined by counting procedure



There are two non-overlapping violations at hours 3 and 12, and these are detected by the counting procedure. However, the true maximum occurs at hour 8 and the second is below the standard. However, in this case the maximum second highest would be V_2, which is above the standard. Although V_2 overlaps the maximum, M, there is one 8-hour average, namely V_1, that is at least as high as V_2 and the two time periods are disjoint.

Figure 6. Eight-hour average violations as highest non-overlapping values





In Figure 7a the maximum value is 12, as is the second highest value. However, in Figure 7b the maximum is now 14, and the second highest non-overlapping is below the standard. Thus, the second highest non-overlapping value can actually be lowered by having more high values. It should be noted that in both of the above cases the maximum second highest non-overlapping value is 12.

Figure 7a,b. Subtleties involved in using non-overlapping values second highest value that must be maintained below the standard in order to satisfy the requirement that the standard not be exceeded more than once per year.

11.5.2 DATA ANALYSIS NEEDS FOR MAINTENANCE

The discussion of the considerations for the selection of communities and neighborhoods for monitoring networks, as presented in Section 11.2 relied primarily on the type of data which was desired to be assimilated. Obviously, the type of data which was wanted was a direct result of the use to which it was to be put. Therefore a basic understanding of the concepts behind the establishment of a data gathering system has already been established.

This discussion summarizes these concepts and provides a review of how the data will actually be used. Just as the differences in the needs for data assimilation for potential AQMA's and designated AQMA's required the separate presentation of these two categories, the parallel need for different analysis of this data also requires their separate presentation at this point.

11.5.2.1 Designation of AQMA's

The monitoring network for potential AQMA's, described in Section 11.2 is one which allows for a better analysis of the area under question for the next iteration of AQMA designations in 1980, and the subsequent maintenance plan preparation if warranted, by the establishment of an appropriate data base. To fulfill both of these objectives it is necessary to analyze the air quality and other data to determine what the trend is in air quality and how this compares with the changes in other parameters of interest. This level of comparison may either be on a visual trend-versus-trend basis or may actually be directed at the determination of a quantitative relationship.

Trend Comparison - In order to project air quality on a 10-year basis, it is necessary to have a good understanding of how the air quality can be expected to change due to the growth of certain sectors of the economy and the emission reduction planned under the state implementation plans. At this time, such an understanding is insufficient. A review of the trends of air quality and the parameters of interest may at least provide an estimate of the changes which are likely to occur and may even provide a quantitative methodology which is applicable to the region of interest. The parameters of interest which are suggested for consideration are population, employment, and industry earnings. The more finely these parameters are divided to correspond to the pollution being measured, the more likely good relationships can be determined.

If the emission levels are also being monitored, then it is likely that the air quality data would be compared to the emission levels and the emission levels would be compared to the parameters of interest. This type of review leads the way to the calibration of modeling techniques while providing a more appropriate quantity (i.e., emissions) to be projected from the socioeconomic data.

Special attention in the review could be paid to the impact on air quality of any new sources, regulations, controlling emissions, or special meteorological conditions which occurred during the period of interest. Such events could significantly change the air quality without affecting the socioeconomic data thereby leading to confusion when trying to make comparisons. Consideration of these impacts also provides a good basis from which the impact of possible maintenance measures could be projected.

Other Analyses - As mentioned above, the data provided by the monitoring network may provide sufficient information to allow for a qualitative comparison of the air quality trends with the appropriate parameters which could allow for a semi-quantitative interpretation of future air quality based on the projections of the parameters. Such a trend comparison may also lead the way to a more quantitative determination of an air quality or emission growth projection methodology which is applicable to the area of interest.

11.5.2.2 Monitoring of the AQMP

The monitoring network described for the designated AQMA's was primarily formulated around the need for reviewing the progress of its maintenance plan. This network provides the data necessary for the establishment of a trend from which the general success of the program can be determined and for the continuing review processes which may be inherent in the plan.

Establishment of Trend - By reviewing the air quality data provided by the monitoring network, it is possible to establish a trend in the air quality. For the purpose of determining the real success of the plan, it is most helpful to do trend analysis on the individual stations as well as the overall air quality situation.

The trend in the air quality at each station should be compared with that projected to occur under the AQMP. Where the trend is not that which was projected to occur, whether it is more advantageous or not, the reasons for the differences should be investigated. The review of both situations allows for not only the identification of problem areas where new, more stringent or different controls are needed but also for the better understanding of those measures which were implemented and are having a greater than expected effect. Either of these reviews may provide a better basis for estimating the projected impact of any new measures which may need to be applied.

Review Processes - The air quality data and the trends thereof also serve as the basis for any review processes which may be required in the maintenance plan. This review may be needed for new source construction (both direct and indirect), for special operating conditions (supplementary control strategies), or for emergency actions needed due to periods of air quality alerts. The latter two of these require an immediate response type of network where minimum analysis is possible which the review for new source construction would rely more heavily on the trends in the area and air quality modeling.

11.5.3 TREND ANALYSIS TECHNIQUES AND APPLICABILITY

The detection, verification, and assessment of trends in the ambient air quality, as measured by the monitoring network, is the single most important data analysis requirement, being significant not only for monitoring the maintenance of the standards in designated AQMA's, but also for obtaining guidance relative to the possible designation of additional areas. To fulfill this need, there are a variety of techniques available, ranging from the simple to the sophisticated, and in some cases applying best in slightly different situations. In this section, a variety of techniques will be described, and the purpose and applicability of each discussed.

11.5.3.1 Visual Techniques

When performing a trend analysis, it is extremely desirable to actually look at the data in a graphic form. Since the raw data are commonly quite variable, plots are usually made of quarterly or annual averages. If there is a clear trend in the air quality, its determination may be intuitively obvious in such a plot, and no more may be needed.

If there is still significant variability in the statistics in such a plot, however, the trend may not be obvious; in this case, it is desirable to have a smoothed trend line through the data, and it is important that such a trend line be determined in an objective way. This can be most simply done by calculating a moving average of the observations. This will provide a smoother and simpler representation of the original data, as in Figure 8. For quarterly averages, a moving annual average consisting of four quarterly averages will eliminate the seasonal fluctuations and remove much of the random variation as well. When considering annual values over several years, a 3-year moving average will smooth out much of the year-to-year variation. In specific instances other averaging schemes may be considered; the selection of the appropriate moving average is somewhat subject to personal judgment and experience. When employing the moving average, the first and last values at the beginning and end of the data series are usually omitted.

At least one much more elaborate smoothing technique, the Whittaker-Henderson smoothing formula, has been previously applied in displaying trends in air quality data, but this approach is sufficiently more laborious to apply and generally requires the use of a computer.

The Whittaker-Henderson formula is a finite-difference formula drawn from the field of actuarial science. It calculates the smoothed curve as a balance between the smoothest possible line (a straight line) and the best-fitting line (a line through each data point); by changing a weighting parameter in the formula, the balance between smoothness and good fit can be modified to provide plots that indicate seasonal patterns as well as plots that demonstrate long-term trends. These smoothed plots make excellent, objective, diagnostic tools for examining data in search of trends. They do not, however, provide a means of objectively demonstrating the significance of the trends that may be found; for that, a statistical significance test, such as described in the next subsection, is required.

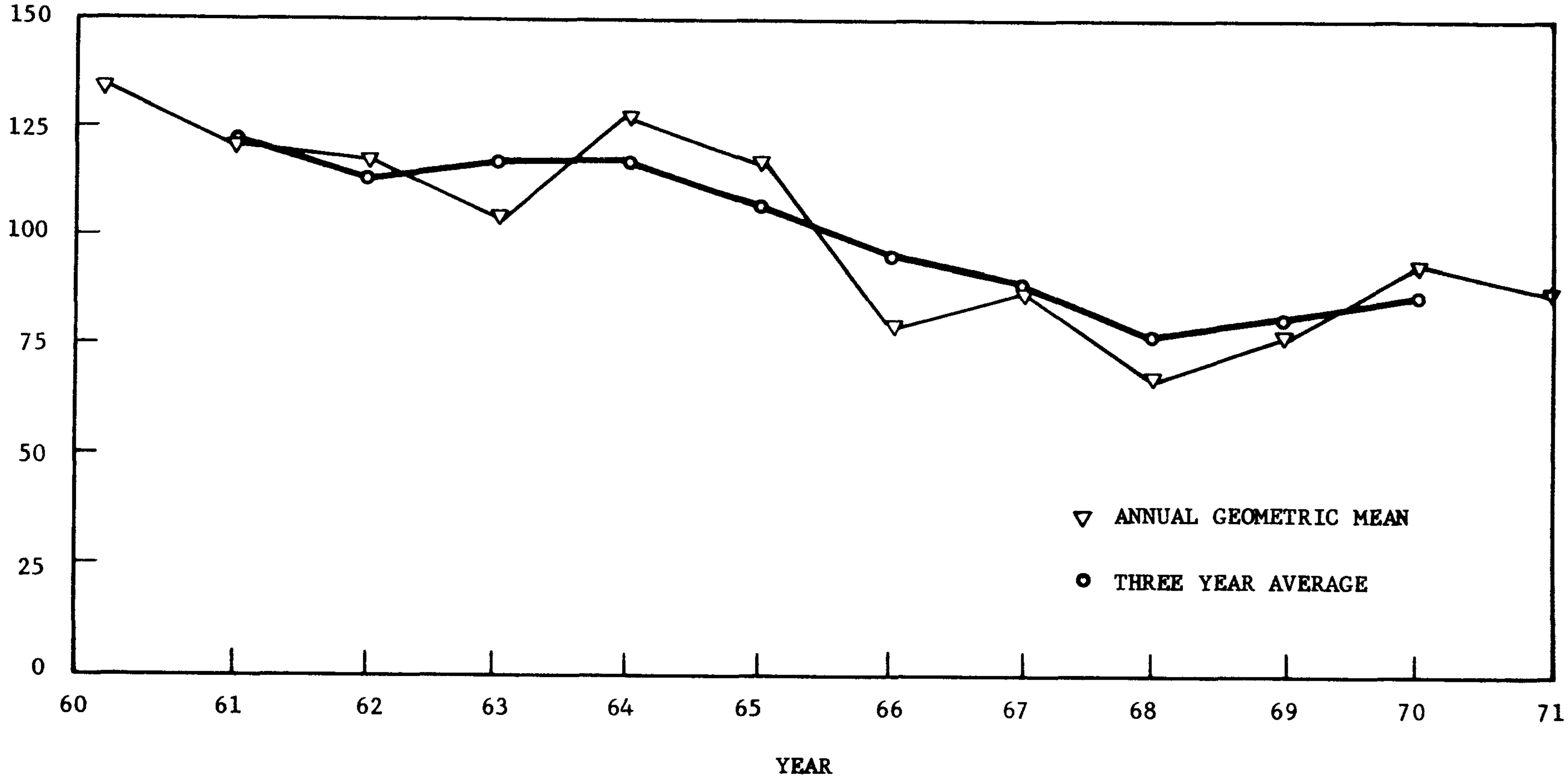


Figure 8. Suspended particulate data from Tucson, Arizona

11.5.3.2 <u>Significance Techniques</u>

These techniques consider the statistical significance of the correlation coefficient between a series of pollutant observations or summary statistics with the sequence in which they were observed. If there is an overall trend upward or downward, this type of technique assesses the probability that it may be a sampling anomaly from a real-life situation which does not have a real trend. This is done by judging the trend in the light of the variability of the data, so that a real trend will be much less likely to be found significant if the series of data show great variability than if the data series is rather smooth. Because of this, it is desirable to use a data series that has no extraneous variability; for a pollutant with strong seasonal patterns, for instance, one should either use annual average values or use the averages for each quarter as a separate series.

Two different types of these correlation significance procedures are presented. The first is nonparametric, meaning no further mathematical assumptions are necessary. It seeks a consistently changing series, either up or down. The second is parametric, requiring the additional assumption, frequently encountered, that either the data or their logarithms are normally distributed. It is sensitive to a constant absolute or percentage change. In both approaches, the time interval between observations is not considered, so that missing observations can be ignored.

<u>Daniel's Test for Trend</u> - In order to utilize this procedure, at least four observations should be available. Given observations $X_1, \ldots, X_n$ and their corresponding relative ranks $R(X_1), \ldots, R(X_n)$, the test statistic is the Spearman Rank Correlation Coefficient:

$$\rho = 1 - \frac{6T}{n(n^2-1)} \quad ,$$

where $T = \Sigma[R(X_i)-i]^2$, that is, the summed squares of the differences between each value's rank and its sequential order, i, in the series of n observations. The absolute value of ρ is compared with a critical value w_p in a table, if $n < 30$, or with $w_p = X_p / \sqrt{n-1}$, if $n \geq 30$, where X_p is the p quantile of a standard normal random variable. If $|\rho| > w_p$, then a trend is declared significant at the $\alpha = 2p$ significance level. A positive value of ρ indicates an upward trend, while a negative value of ρ indicates a downward trend. It can be noted that the estimate of the Spearman rank correlation coefficient ρ is merely the usual product moment correlation of the ranks of the observations with the order in which the observations were taken.

Example applications of Daniel's test - The following table provides the annual geometric means, their relative values, and the index over time, for part of the Tucson TSP data from Figure 8.

X_i	128	118	80	89	70	78	96	88
$R(X_i)$	8	7	3	5	1	2	6	4
i	1	2	3	5	5	6	7	8

If ties had occurred, the ranks could be determined by averaging the ranks among the tied observations, or preferably by utilizing the data estimates to the next available place, even if it is not a significant digit.

$$T = \Sigma\,[R(X_i)-i]^2$$
$$= (8-1)^2 + (7-2)^2 + (3-3)^2 + (5-4)^2 + (1-5)^2 + (2-6)^2 + (6-7)^2 + (4-8)^2$$
$$= 49 + 25 + 0 + 1 + 16 + 16 + 1 + 16$$
$$= 124$$
$$\rho = 1 - \frac{6T}{n(n^2-1)}$$
$$= -0.476$$

The 0.90 quantile of the Spearman test statistic is 0.5000 for n = 8. Since the absolute value of the calculated rank correlation coefficient (0.476) is less than this value, the correlation would not be accepted as being significantly different from zero, even at the $\alpha = 0.20$ level. This means that, based on probability theory, there is at least a 20 percent chance that the observed downward trend in the 8 years' data is a sampling anomaly from a situation where there is really no trend; i.e., no correlation between the position of a year in the series and the relative geometric mean particulate level for that year.

Daniel's test is, like other such tests, sensitive to the number of data points involved. When the test is applied to the 12 years of data 1960 to 1971, the calculated coefficient is $\rho = -0.769$. The absolute value 0.769 is greater than the 0.995 quantile for n = 12. Therefore, the 12-year trend would be clearly classified as significantly different from zero (with 1 percent chance of error) and, because of the negative correlation, as being downward. When the test is applied to the 4-year period 1968 to 1971, the coefficient is + 0.80, which is significant only at the 0.20 level. Although the upward pattern for the 4 years seems quite clear, the test is just not very sensitive for such short periods of data.

The nonparametric correlation technique is primarily useful for classifying the temporal pattern as upward or downward and for indicating the consistency of the pattern by the statistical significance level. It is not very useful for studying the data to find a trend because of its lack of sensitivity, although it could be used in a computerized screening situation to select out only the clearest and most obvious trends. Recent publications by Faoro and Frank provide excellent examples of the use of this methodology.[2,3]

<u>Student t-Test for Linear Trend in a Sequence of Normal Variables</u> - In contract with the previous technique, the theory underlying this approach requires assumptions about the distributional form of the data,

specifically that the data or their logarithms are normally distributed. With particulate data, the latter assumption is usually chosen.

Let X_i, $i = 1,n$ be a sequence of observations or their logarithms. Then the test statistic is

$$T = \frac{\sqrt{n-2}\,\sqrt{c}\,\hat{\beta}}{\sqrt{\sigma^2 - c\,\hat{\beta}^2}}, \quad \text{where} \quad \begin{aligned} c &= \frac{1}{12}(n^2-1) \\ \hat{\beta} &= \frac{1}{nc}\sum\left(i-\frac{n+1}{2}\right) X_i \\ \sigma^2 &= \frac{1}{n}\sum (X_i-\bar{X})^2 \end{aligned}$$

The calculated T is compared to the p^{th} quantile of Student's t statistic with n-2 degrees of freedom provided in Table A-3. If $|T| > t$, then the trend is declared significant at the $\alpha = 2\,(1-p)$ significance level. A positive value of T indicates an upward trend, while a negative value of T indicates a downward trend.

Example Application of Student t-Test - Again using as an example the particulate data from Tucson, Arizona for 1964 to 1971, we have:

$$T = \sqrt{n-2}\,\sqrt{c}\;\hat{\beta}\Big/\sqrt{\sigma^2 - c\beta^2}$$

$$c = \frac{1}{12}(n^2-1) = \frac{1}{12}(64-1) = 5.25$$

$$\hat{\beta} = \frac{1}{nc}\sum\left(i-\frac{n+1}{2}\right) X_i$$

$$= \frac{1}{8(5.25)}\,[-3.5 \ln (128) - 2.5 \ln (118) - 1.5 \ln (80) - 0.5 \ln (89) + 0.5 \ln (70) + 1.5 \ln (78) + 2.5 \ln (96) + 3.5 \ln (88)]$$

$$= -1.985/42 = -.047$$

$$\sigma^2 = \frac{1}{8}\sum (X_i - X)^2 = \frac{1}{8}\sum X_i^{\;2} - \frac{\sum X_i}{8} = 0.03715$$

$$\text{then } T = \sqrt{6}\,\sqrt{5.25}\left(-0.047 \Big/ \sqrt{.037 - 5.25\ (0.002)}\right)$$

$$= -1.65$$

This value lies between the 0.90 and 0.95 quantile of the student's t statistic. Therefore, the trend can be considered significant at the 0.20 level but not the 0.10 level. Thus both the Spearman and the parametric correlation techniques failed to detect a trend during 1964 to 1971 because of the year-to-year variability in the annual estimates.

Considering the entire 12-year period, the value of the nonparametric test statistic T is -3.810. This is significant at the 0.01 level, and the trend can therefore be classified as significantly downward. The corresponding value of the test statistic T for the 4-year interval 1968 to 1971 is + 2.15. This is only significant at the 0.20 level. Again, note the similarity between these results and those obtained by using the simpler non-parametric analogue.

Testing for Proportion of Standards Violations - This technique is useful to test for a trend in the occurrence of extreme values or other short-term statistics. A chi-square test compares the percent of observations above a given threshold concentration, such as a 1-hour standard, between two time periods. It is desirable to consider independent observations. Therefore, for hourly data one should consider at most one observation per day, e.g., the maximum observation per day or the observation of a particular hour. In general, observations derived by intermittent sampling can be considered independent.

The data are arrayed in a table and labeled as shown below; the test should not be used if there are less than five observations in any of the four cells.

	No. Obs $\leq$ Standard	No. Obs > Standard	
TIME PERIOD I	a	b	n_1
TIME PERIOD II	c	d	n_2
			N

Let $p_1 = b/n_1$ be the proportion of observations in time period I that are above the standard. Similarly, $p_2 = d/n_2$ for time period II.

One can test (i) for any change between the two time periods, disregarding whether it is an increase or a decrease; i.e., $p_1 = p_2$, or (ii) for a specific direction of change between the two time periods, say $p_1 \leq p_2$.

The test statistic T is defined as:

$$T = \frac{(n_1+n_2)\ (ad-bc)^2}{n_1 n_2\ (a+c)\ (b+d)}$$

Consider a change to have occurred if T exceeds the chi-square statistic of Table A-4 at the $1 - \alpha$ quantile with 1 degree of freedom.

If only one direction of change is of interest, for example "has there been improvement", then consider improvement to have occurred if T exceeds the chi-square statistic at the $(1 - 2\alpha)$ quantile. In either case the significance level is approximately α.

Example Chi-Square Test - The following table represents the number of days on which the maximum 1-hour oxidant concentration exceeded the 1-hour standard at a particular location during two periods, 1964 to 1967 and 1968 to 1971.

	< standard	> standard	
1964-1967	662	154	n_1 = 816
1968-1971	714	111	n_2 = 825
TOTALS	1376	265	1641

$$T = \frac{(n_1+n_2)\ (ad-bc)^2}{n_1 n_2\ (a+c)\ (b+d)}$$

$$= \frac{(1641)\ [(662)\ (111)\ -\ (154)\ (714)]^2}{(816)\ (825)\ (1376)\ (265)}$$

$$= 8.9$$

At the level of significance of 0.05, the calculated value of T exceeds the tabulated statistic at 0.90 quantile = 2.706. It can therefore be concluded that short-term oxidant levels have significantly decreased in recent years at the particular site.

11.5.3.3 Quantitative Techniques

Although the significance techniques just discussed provide an objective probability judgment about whether an air quality trend is real, they do not provide any quantitative estimate of how much the pollutant levels are increasing or decreasing, say in terms of $\mu g/m^3$ per year. Such information could be gained, for example, by scaling a plot of the data. However, this introduces again an element of subjectivity, as well as being impractical for extensive application. Preferably, objective estimates of the slope of the trend should be made using regression techniques.

Regression techniques involve choosing a simple model to represent the trend and then estimating the parameters of the model to fit the data as well as possible. The two models to be discussed here are a simple linear model, corresponding to a constant absolute change from year to year, and an exponential model, corresponding to a constant percentage change from year to year.

<u>Simple Linear Model</u> - To estimate a constant absolute change, b, corresponding to the model X = a + bT, use as an estimate of the slope b:

$$\hat{b} = \frac{\sum (T_i - \bar{T})\, X_i}{\sqrt{\sum (T_i - T)^2}\sqrt{\sum (X_i - \bar{X})^2}} \quad ,$$

where pollutant concentration X_i exists at time T_i.

The estimate of "a" is $\hat{a} = \bar{X} - \hat{b}\bar{T}$.

The use of this algebra will of course produce an estimate of a slope even if the trend would not be seen as real by one of the significance tests. If there is little real trend, the estimate of the slope will be small, but only by chance will it be precisely zero. Thus there is a need to test, in the manner of the significance techniques, whether the estimated slope is significantly different from zero, given the variability of the original data.

The statistical significance of the estimate of b as compared with an assumed value b. can be tested by computing

$$B = (\hat{b} - b.) \sqrt{s^2 / \Sigma\, (T-\bar{T})^2}$$

where $s^2 = \{\Sigma\, (X-\bar{X})^2 - [\Sigma\, (T-\bar{T})\, X]^2 \,/\, [\Sigma\, (T-\bar{T})^2]\} \,/\, (n-2)$

and comparing B with the Student's t statistic, t, at the p quantile, with n - 2 degrees of freedom. If $|B| > t$ then the rate of change is significantly different than b. at the $\alpha = 2p$ signficance level. If b. is chosen as zero, the test amounts to a test of whether there is any real trend. In a similar manner, a confidence interval can be created about the estimate $\hat{b}$. The interval is defined as

$\hat{b} \pm t \sqrt{s^2/\Sigma (T-\bar{T})^2}$. This interval contains the "true" rate of change, b, with probability $1 - \alpha$.

Exponential Model - To estimate the percent rate of change, r, corresponding to the model $X = ar^T$, calculate and test the significance of log(r) by substituting $\log(X_i)$ for X_i in the formulae of the previous section. The rate of change is usually presented as a change of (r-1) x 100 percent per unit of time; e.g., r = 1.17 would mean 17 percent per year.

Examples of Regression Application - The above regression techniques are applied to the TSP data for New Haven, Connecticut, in Figure 9.

The estimates of absolute and percentage rates of change are presented for the time intervals 1960 to 1971, 1964 to 1971, and 1968 to 1971.

Rates of Change

Absolute ($\mu g/m^3$)		Percent
0.26	1960-1971	+0.27
-2.24	1964-1971	-3.46
+7.00	1968-1971	+9.26

This again demonstrates that the choice of time interval can play an important role in the determination of an estimated rate of change.

11.5.3.4 Quality Control Approach

Another approach to trend analysis that is really a significance type technique applied sequentially, has been adopted from the field of quality control; it is utilized in EPA's computerized Plan Revision Management System (PRMS), and is particularly suited to the situation of air quality maintenance.[4] This quality control approach abstractly

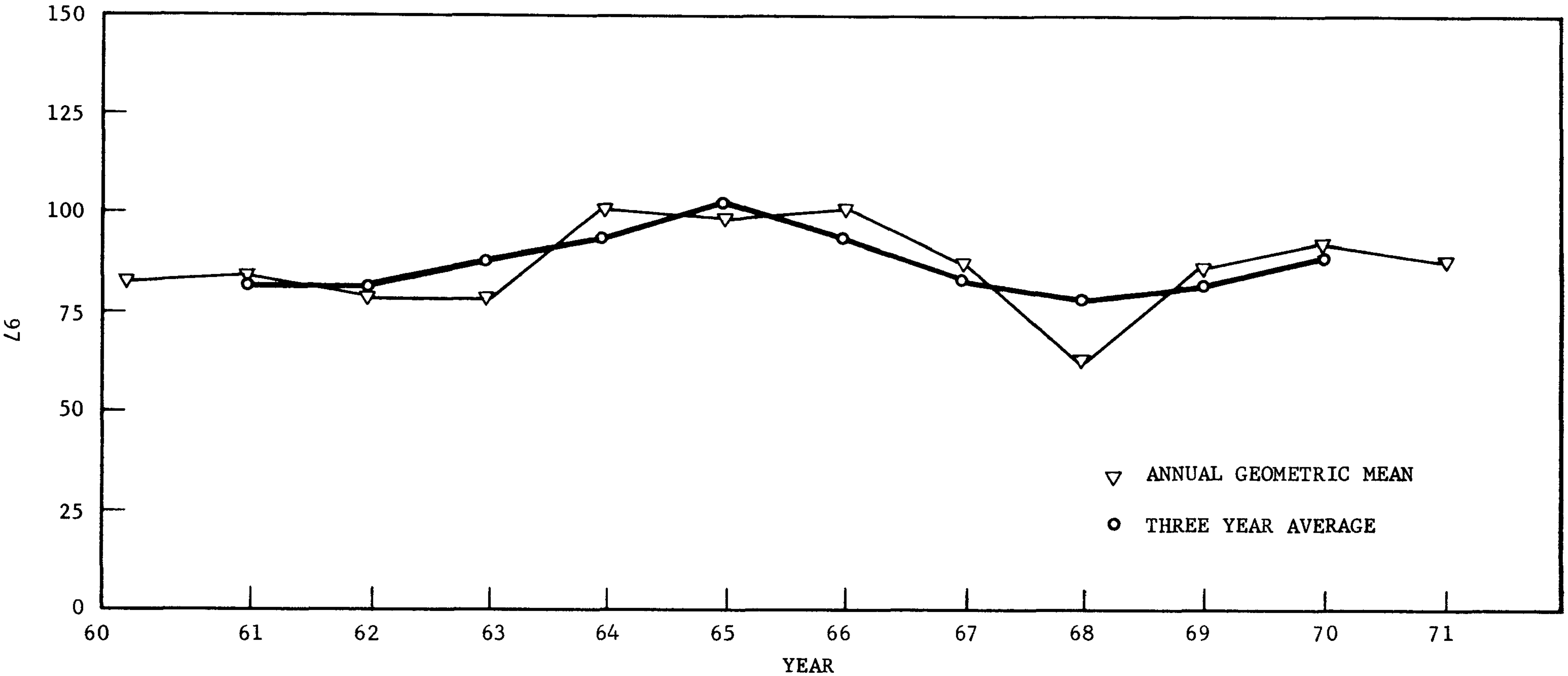


Figure 9. Suspended particulate data from New Haven, Connecticut

involves making a significance test at each period of time in a series, testing the observed air quality value against a prediction. This involves first calculating t-test confidence intervals about the predicted air quality value, based on the observed standard deviation, and using them to test whether the measured air quality is significantly different.

The primary value of this approach is visual; as seen in Figure 10, plotting the predicted and observed air quality and the confidence limits together permit the visualization of trends toward or away from the predicted value even before they are confirmed by exceeding the confidence limits. It is precisely this "advance warning" feature of these techniques that make them attractive in the field of process quality control monitoring. In Figure 6, the line labelled "predicted geometric mean" was calculated based on the Colorado State Implementation Plan. The fourth quarter of 1970 is the starting point for both the observed and predicted annual geometric means because 1970 was used as the base year in the implementation plan. From that point on, both the observed and predicted geometric means increase through the first quarter of 1973; then the predicted geometric mean decreases until it achieves the NAAQS in the first quarter of 1976. However, the observed geometric mean exceeds the upper bound in the first quarter of 1972 and continues to exceed the upper bound through the third quarter of 1973. The observed value exceeded the upper confidence limit as early as the first quarter of 1972, and the trend would have been apparent even before that, although the analysis system was not available at that time.

The quality control approach has obvious applications to air quality maintenance planning. The observed ambient data in an AQMA can be monitored against the NAAQS, for example, or the trends in the air quality of potential AQMA's can be observed relative to their possible

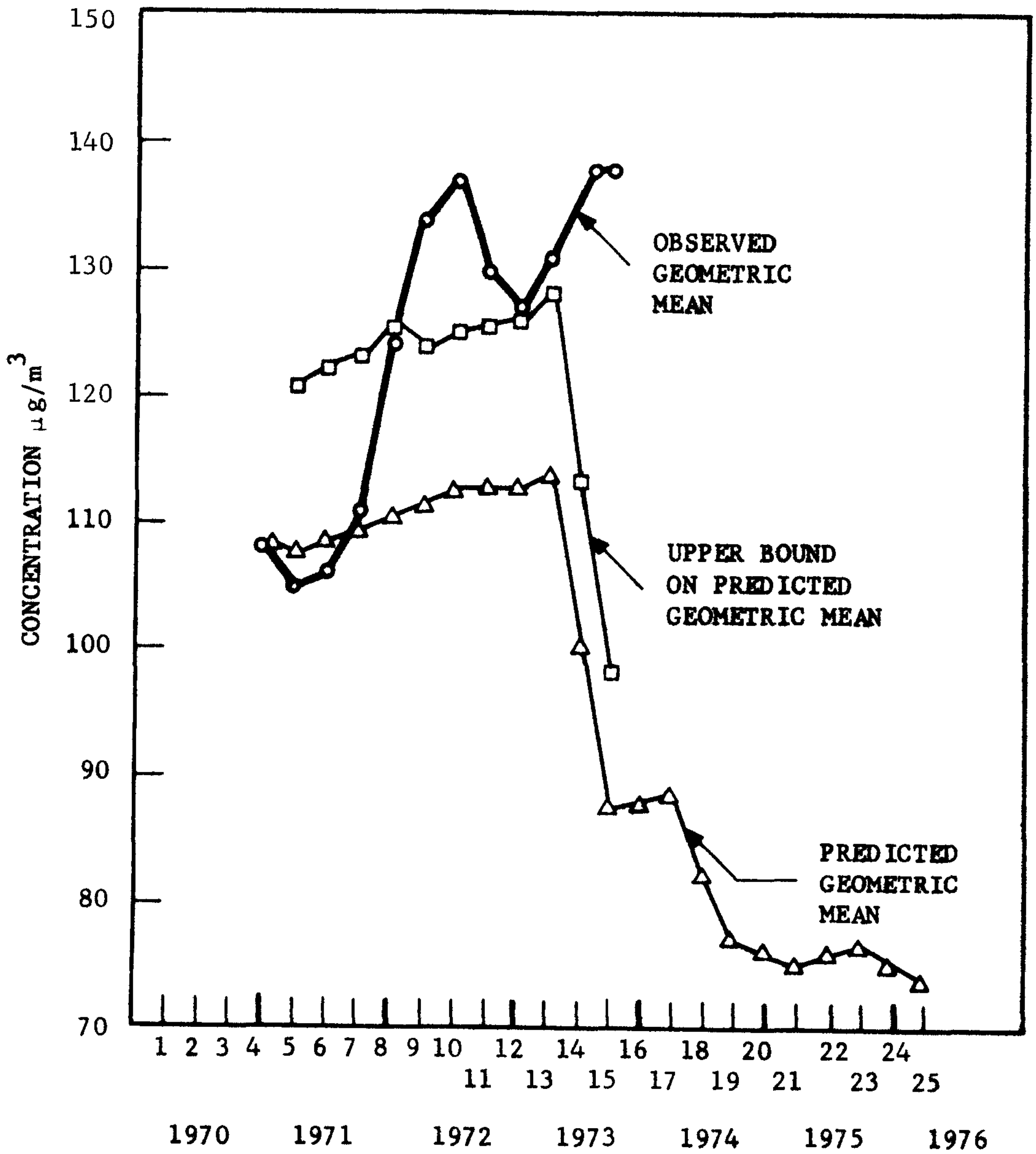


Figure 10. Quality control approach applied to particulate data from Denver, Colorado.

designation. For best utilization, the approach does require computerization, preferably with graphic output, so that it is a significant effort. However, as air quality maintenance planning becomes more institutionalized over the years, some such capability should be obtained as soon as possible.

11.5.4. REFERENCES

1. Curran, T. C. and Hunt, W. F., Jr., "Interpretation of Air Quality Data with Respect to the National Ambient Air Quality Standards," Presented at the 67th Annual Meeting of the Air Pollution Control Association, Denver, Colo., 1974. (Accepted for publication J.A.P.C.A.)

2. Faoro, R., "Trends in Concentrations of Benzene-Soluble Suspended Particulate Fraction and Benzo(a) Pyrene Determined by Data from the National Air Surveillance Network," Presented at the 67th Annual Meeting of the Air Pollution Control Association, Denver, Colo., 1974. (Accepted for publication J.A.P.C.A.)

3. Frank, N., "Temporal and Spatial Relationships of Sulfates, Total Suspended Particulates and Sulfur Dioxide," Presented at the 67th Annual Meeting of the Air Pollution Control Association, Denver, Colo., 1974.

4. Hunt, W. F., Jr. and Curran, T. C., "An Application of Statistical Quality Control Procedures to Determine Progress in Achieving the 1975 National Ambient Air Quality Standards," 28th Annual Technical Conference Transactions, American Society for Quality Control, Boston, Mass., 1974.

APPENDIX A

STATISTICAL TABLES

Table A-1. QUANTILES OF THE SPEARMAN TEST STATISTIC[a]

n	$p = .900$	.950	.975	.990	.995	.999
4	.8000	.8000				
5	.7000	.8000	.9000	.9000		
6	.6000	.7714	.8286	.8857	.9429	
7	.5357	.6786	.7450	.8571	.8929	.9643
8	.5000	.6190	.7143	.8095	.8571	.9286
9	.4667	.5833	.6833	.7667	.8167	.9000
10	.4424	.5515	.6364	.7333	.7818	.8667
11	.4182	.5273	.6091	.7000	.7455	.8364
12	.3986	.4965	.5804	.6713	.7273	.8182
13	.3791	.4780	.5549	.6429	.6978	.7912
14	.3626	.4593	.5341	.6220	.6747	.7670
15	.3500	.4429	.5179	.6000	.6536	.7464
16	.3382	.4265	.5000	.5824	.6324	.7265
17	.3260	.4118	.4853	.5637	.6152	.7083
18	.3148	.3994	.4716	.5480	.5975	.6904
19	.3070	.3895	.4579	.5333	.5825	.6737
20	.2977	.3789	.4451	.5203	.5684	.6586
21	.2909	.3688	.4351	.5078	.5545	.6455
22	.2829	.3597	.4241	.4963	.5426	.6318
23	.2767	.3518	.4150	.4852	.5306	.6186
24	.2704	.3435	.4061	.4748	.5200	.6070
25	.2646	.3362	.3977	.4654	.5100	.5962
26	.2588	.3299	.3894	.4564	.5002	.5856
27	.2540	.3236	.3822	.4481	.4915	.5757
28	.2490	.3175	.3749	.4401	.4828	.5660
29	.2443	.3113	.3685	.4320	.4744	.5567
30	.2400	.3059	.3620	.4251	.4665	.5479

For n greater than 30 the approximate quantiles of ρ may be obtained from

$$w_p \simeq \frac{x_p}{\sqrt{n-1}}$$

where x_p is the p quantile of a standard normal random variable obtained from Table 1.

SOURCE. Adapted from Glasser and Winter (1961), with corrections.

[a] The entries in this table are selected quantiles w_p of the Spearman rank correlation coefficient ρ when used as a test statistic. The lower quantiles may be obtained from the equation

$$w_p = -w_{1-p}$$

The critical region corresponds to values of ρ smaller than (or greater than) but not including the appropriate quantile. Note that the median of ρ is 0.

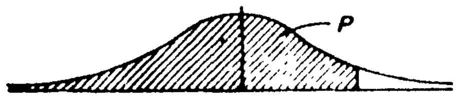
Table A-2. QUANTILES OF THE STANDARD NORMAL DISTRIBUTION[a]

w_p	p	w_p	p	w_p	p
−3.7190	.0001	−.4677	.32	.5244	.70
−3.2905	.0005	−.4399	.33	.5534	.71
−3.0902	.001	−.4125	.34	.5828	.72
−2.5758	.005	−.3853	.35	.6128	.73
−2.3263	.01	−.3585	.36	.6433	.74
−2.1701	.015	−.3319	.37	.6745	.75
−2.0537	.02	−.3055	.38	.7063	.76
−1.9600	.025	−.2793	.39	.7388	.77
−1.8808	.03	−.2533	.40	.7722	.78
−1.7507	.04	−.2275	.41	.8064	.79
−1.6449	.05	−.2019	.42	.8416	.80
−1.5548	.06	−.1764	.43	.8779	.81
−1.4758	.07	−.1510	.44	.9154	.82
−1.4395	.075	−.1257	.45	.9542	.83
−1.4051	.08	−.1004	.46	.9945	.84
−1.3408	.09	−.0753	.47	1.0364	.85
−1.2816	.10	−.0502	.48	1.0803	.86
−1.2265	.11	−.0251	.49	1.1264	.87
−1.1750	.12	.0000	.50	1.1750	.88
−1.1264	.13	.0251	.51	1.2265	.89
−1.0803	.14	.0502	.52	1.2816	.90
−1.0364	.15	.0753	.53	1.3408	.91
−.9945	.16	.1004	.54	1.4051	.92
−.9542	.17	.1257	.55	1.4395	.925
−.9154	.18	.1510	.56	1.4758	.93
−.8779	.19	.1764	.57	1.5548	.94
−.8416	.20	.2019	.58	1.6449	.95
−.8064	.21	.2275	.59	1.7507	.96
−.7722	.22	.2533	.60	1.8808	.97
−.7388	.23	.2793	.61	1.9600	.975
−.7063	.24	.3055	.62	2.0537	.98
−.6745	.25	.3319	.63	2.1701	.985
−.6433	.26	.3585	.64	2.3263	.99
−.6128	.27	.3853	.65	2.5758	.995
−.5828	.28	.4125	.66	3.0902	.999
−.5534	.29	.4399	.67	3.2905	.9995
−.5244	.30	.4677	.68	3.7190	.9999
−.4959	.31	.4959	.69		

SOURCE. Abridged from Tables 3 and 4, pp. 111–112, Pearson and Hartley (1962).

[a] The entries in this table are quantiles w_p of the standard normal random variable W, selected so $P(W \leq w_p) = p$ and $P(W > w_p) = 1 - p$.

Table A-3. QUANTILES OF THE STUDENT'S t DISTRIBUTION



df	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.841
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.086	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576

Table A-4. QUANTILES OF A CHI SQUARE DISTRIBUTION WITH ONE DEGREE OF FREEDOM

Quantile, p	w_p
.750	1.323
.900	2.706
.950	3.841
.975	5.024
.990	6.635
.995	7.879
.999	10.830

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

1. REPORT NO. EPA 450/4-74-012	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE Guidelines for AQM Planning & Analysis Air Quality Monitoring & Data Analysis (Vol 11)		5. REPORT DATE September 1974
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS GCA/Technology Division Burlington Road Bedford, Massachusetts 01730		10. PROGRAM ELEMENT NO. 2AH137
		11. CONTRACT/GRANT NO. 68-02-1478
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Protection Agency Office of Air Quality Planning & Standards Monitoring and Data Analysis Division Research Triangle Park, N.C. 27711		13. TYPE OF REPORT AND PERIOD COVERED Rpt. for Task 1; 7/74-9/74
		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

This report contains guidance concerning monitoring and air quality data analysis related to air quality maintenance. Topical areas covered include: network design, instrument siting, acceptable instrumentation, monitoring site description, air quality trend evaluation, air quality data evaluation, interpretation of air quality as it relates national ambient air quality standards and procedures for validating, editing, and screening air quality data.

The quantity, type, temporal, and geographical distribution of air quality data necessary for establishing baseline air quality levels is presented.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Design & Instrument Siting of Air Quality Monitoring Networks Air Quality Data Evaluation Air Quality Data Validation Baseline Air Quality Data	Air Quality Maintenance Guidance Federally Required Air Monitoring Monitoring Pursuant to State Implementation Plans	

18. DISTRIBUTION STATEMENT Release Unlimited	19. SECURITY CLASS *(This Report)* None	21. NO. OF PAGES 113
	20. SECURITY CLASS *(This page)* None	22. PRICE

EPA Form 2220-1 (9-73)

ENVIRONMENTAL PROTECTION AGENCY
Technical Publications Branch
Office of Administration
Research Triangle Park, N.C. 27711

OFFICIAL BUSINESS

AN EQUAL OPPORTUNITY EMPLOYER

POSTAGE AND FEES PAID
ENVIRONMENTAL PROTECTION AGENCY
EPA - 335

THIRD CLASS BULK RATE

Return this sheet if you do NOT wish to receive this material [ ], or if change of address is needed [ ]. (Indicate change, including ZIP code.)

PUBLICATION NO. EPA-450/4-74-012
(OAQPS NO. 1.2-030)