

United States  
Environmental Protection  
Agency

Environmental Monitoring Systems  
Laboratory  
Research Triangle Park NC 27711

---

Research and Development

EPA-600/9-76-005 Dec. 1984

---



# **Quality Assurance Handbook for Air Pollution Measurement Systems: Volume I. Principles**

QUALITY ASSURANCE HANDBOOK  
FOR  
AIR POLLUTION MEASUREMENT SYSTEMS  
Volume I - Principles

U.S. ENVIRONMENTAL PROTECTION AGENCY  
Office of Research and Development  
Quality Assurance Division  
Environmental Monitoring Systems Laboratory  
Research Triangle Park, North Carolina 27711

## ACKNOWLEDGMENTS

This volume of the Quality Assurance Handbook has been prepared by the Quality Assurance Division of Environmental Monitoring Systems Laboratory in cooperation with PEDCo Environmental, Inc., of Cincinnati, Ohio.

## DISCLAIMER

Mention of trade names or commercial products does not constitute EPA endorsement or recommendation for use.

VOLUME I  
TABLE OF CONTENTS

<u>Section</u>	<u>Pages</u>	<u>Revision</u>	<u>Date</u>
1.1 PURPOSE OF THE QUALITY ASSURANCE HANDBOOK	1	1	1-9-84
1.2 OVERVIEW OF THE QUALITY ASSURANCE HANDBOOK	5	1	1-9-84
1.3 DEFINITION OF QUALITY ASSURANCE	2	1	1-9-84
1.4 ELEMENTS OF QUALITY ASSURANCE	3	1	1-9-84
1.4.1 Document Control and Revisions	3	1	1-9-84
1.4.2 Quality Assurance Policy and Objectives	4	1	1-9-84
1.4.3 Organization	7	1	1-9-84
1.4.4 Quality Planning	5	1	1-9-84
1.4.5 Training	8	1	1-9-84
1.4.6 Pretest Preparation	8	1	1-9-84
1.4.7 Preventive Maintenance	6	1	1-9-84
1.4.8 Sample Collection	2	1	1-9-84
1.4.9 Sample Analysis	4	1	1-9-84
1.4.10 Data Reporting Errors	6	1	1-9-84
1.4.11 Procurement Quality Control	5	1	1-9-84
1.4.12 Calibration	14	1	1-9-84
1.4.13 Corrective Action	3	1	1-9-84
1.4.14 Quality Costs	12	1	1-9-84
1.4.15 Interlaboratory and Intralaboratory Testing	15	1	1-9-84

<u>Section</u>	<u>Pages</u>	<u>Revision</u>	<u>Date</u>
1.4.16 Audit Procedures	7	1	1-9-84
1.4.17 Data Validation	15	1	1-9-84
1.4.18 Statistical Analysis of Data	3	1	1-9-84
1.4.19 Configuration Control	5	1	1-9-84
1.4.20 Reliability	5	1	1-9-84
1.4.21 Quality Reports to Management	4	1	1-9-84
1.4.22 Quality Assurance Program Plan	4	1	1-9-84
1.4.23 Quality Assurance Project Plan	2	1	1-9-84

#### APPENDICES

A	GLOSSARY OF TERMS	23	1	1-9-84
B	NOMENCLATURE	7	1	1-9-84
C	COMPUTATIONAL EXAMPLES OF DESCRIPTIVE STATISTICS	18	1	1-9-84

Data presentation, frequency distribution; measures of central tendency (arithmetic mean, median, geometric mean) measures of dispersion (range, variance and standard deviation, geometric standard deviation, use of range to estimate standard deviation, relative standard deviation, absolute percent difference), number of places to be retained in computation and presentation of data

D	PROBABILITY DISTRIBUTIONS	18	1	1-9-84
---	---------------------------	----	---	--------

Normal distribution, log-normal distribution, Weibull distribution, distribution of sample means, use of probability graph paper

<u>Section</u>	<u>Pages</u>	<u>Revision</u>	<u>Date</u>	
<u>APPENDICES</u> (continued)				
E	ESTIMATION PROCEDURES	10	1	1-9-84
	Estimation (point, estimate, confidence interval, t distribution), determination of sample size			
F	OUTLIERS	14	1	1-9-84
	Procedure(s) for identifying outliers (Dixon ratio, Grubbs test, control chart)			
G	TREATMENT OF AUDIT DATA	11	1	1-9-84
	Audit data, data quality assessment, EPA audit performance, analysis of audit data, sign test, paired t-test			
H	CONTROL CHARTS	32	1	1-9-84
	Description and theory, applications and limitations, construction of control charts (precision and/or variability, averages, percent of defective measurements, individual results, moving averages and ranges, other charts), interpretation of control charts; steps in developing and using a control chart system			
I	STATISTICAL SAMPLING	9	1	1-9-84
	Concepts, how to select a random sample and stratified random sample, acceptance sampling by attributes, acceptance sampling by variables			

<u>Section</u>	<u>Pages</u>	<u>Revision</u>	<u>Date</u>
<u>APPENDICES (continued)</u>			
J CALIBRATION	23	1	1-9-84
Concepts, multipoint calibration, zero-span calibration checks, suggested calibration procedure, other regression problems			
K INTERLABORATORY TESTING	8	1	1-9-84
Concepts, interlaboratory tests (analysis of interlaboratory test - summary analysis, analysis for a particular lab, feedback of information on methods) collaborative test			
L RELIABILITY AND MAINTAINABILITY	8	1	1-9-84
Definitions of availability and reliability, life testing			
M INTERIM GUIDELINES AND SPECIFICATIONS FOR PREPARING QUALITY ASSURANCE PROJECT PLANS	34	1	1-9-84
Sixteen items to consider in a QA Project Plan, example QA Project Plan			

## 1.1 PURPOSE OF THE QUALITY ASSURANCE HANDBOOK

The purpose of this Quality Assurance Handbook for Air Pollution Measurement Systems is to provide guidelines and procedures for achieving quality assurance in air pollution measurement systems. It is intended to serve as a resource document for the design of quality assurance programs and to provide detailed operational procedures for certain measurement processes. This Handbook should be particularly beneficial to operators, project officers, and program managers responsible for implementing, designing and coordinating air pollution monitoring projects.

The Handbook is a compilation of quality assurance principles, practices, guidelines, and procedures that are applicable to air pollution measurement systems.

What is presented in the Handbook is an "ultimate" or "ideal" quality assurance program for air pollution measurement systems. All specific measurement systems will not be amenable to all the principles and guidelines contained in the Handbook. A quality assurance program for air pollution measurement systems should consider a number of areas or elements. These elements are shown in Figure 1.4.1 (Section 1.4) in a "Quality Assurance Wheel." The wheel arrangement illustrates the need for a quality assurance system that addresses all elements and at the same time allows program managers the flexibility to emphasize those elements that are most applicable to their particular program.



## 1.2 OVERVIEW OF THE QUALITY ASSURANCE HANDBOOK

This Handbook includes guiding principles and recommended procedures for achieving quality assurance in air pollution measurement systems. It provides general guidelines applicable to most measurement systems as well as specific guidelines applicable to particular measurement processes.

Volume I contains brief discussions of the elements of quality assurance. Expanded discussions of technical points and sample calculations are included in the Appendices. The discussion of each element is structured to be brief and to highlight its most important features. Organizations developing and implementing their own quality assurance programs will find Volume I and the references contained therein useful for general guidance.

Volume II contains quality assurance guidelines for ambient air quality measurement systems. Regardless of the scope and magnitude of ambient air measurement systems, there are a number of common considerations pertinent to the production of quality data. These considerations are discussed in Section 2.0 of Volume II and include quality assurance guidelines in the areas of:

1. Sampling network design and site selection - monitoring objectives and spatial scales; representative sampling; meteorological and topographical constraints; and sampling schedules.
2. Sampling considerations - environmental controls; probe and manifold design; maintenance; and support services.
3. Data handling and reporting considerations - data recording systems, data validation, and systematic data management.
4. Reference and equivalent methods.
5. Recommended quality assurance program for ambient air measurements.

6. Chain-of-custody procedure for ambient air samples - sample collection; sample handling; analysis of the sample; field notes; and report as evidence.

7. Traceability protocol for establishing true concentrations of gases used for calibrations and audits - establishing traceability of commercial gas cylinders and of permeation tubes.

8. Calculation procedures for estimating precision and accuracy of data from SLAMS and PSD automated analyzers and manual methods.

9. Specific guidance for a quality control program for SLAMS and PSD automated analyzers and manual methods - analyzer selection, calibration, zero and span checks; data validation and reporting; quality control program for gaseous standards and flow measurement devices.

10. EPA national performance audit program.

11. System audit criteria and procedures for ambient air monitoring programs.

12. Audit procedures for use by State and local air monitoring agencies.

The remainder of Volume II contains method and/or principle description and quality assurance guidelines for specific pollutants. Each pollutant-specific section contains the following information.

1. Procedures for procurement of equipment and supplies.
2. Calibration procedures.
3. Step-by-step descriptions of sampling, reagent preparation, and analysis procedures, as appropriate, depending upon the method or principle in the case of equivalencies.
4. Method of calculation and data processing checks.
5. Maintenance procedures.
6. Recommended auditing procedures to be performed during the sampling, analysis, and data processing.

7. Recommended procedures for routine assessment of accuracy and precision.
8. Recommended standards for establishing traceability.
9. Pertinent references.
10. Blank data forms for the convenience of the Handbook user (data forms are partially filled in within the text for illustration purposes).

Matrix tables at the ends of appropriate sections summarize the quality assurance functions therein. Each matrix includes the activities, the acceptance limits, the method and frequency of each quality assurance check, and the recommended action if the acceptance limits are not satisfied.

Volume II contains quality assurance guidelines for pollutant method-specific measurement systems. The measurement methods currently in Volume II include:

1. Reference Method for the Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method).
2. Reference Method for the Determination of Suspended Particulates in the Atmosphere (High-Volume Method).
3. Reference Method for the Determination of Nitrogen Dioxide in the Atmosphere (Chemiluminescence).
4. Equivalent Method for the Determination of Nitrogen Dioxide in the Atmosphere (Sodium Arsenite).
5. Equivalent Method for the Determination of Sulfur Dioxide in the Atmosphere (Flame Photometric Detector).
6. Reference Method for the Determination of Carbon Monoxide in the Atmosphere (Nondispersive Infrared Spectrometry).
7. Reference Method for the Determination of Ozone in the Atmosphere (Chemiluminescence).
8. Reference Method for the Determination of Lead in Suspended Particulate Matter Collected from Ambient Air (Atomic Absorption Spectrometry).
9. Equivalent Method for the Determination of Sulfur Dioxide in the Atmosphere (Fluorescence).

As methods are added to Volume II, these will be sent to Handbook users through the document control system, as described in Section 1.4.1 of Volume I of this Handbook.

### Volume III - Stationary-Source-Specific Methods

Volume III contains quality assurance guidelines on stationary-source-specific methods. The format for Volume III is patterned after that of Volume II.

Regardless of the scope and purpose of the emissions-testing plan, there are a number of general considerations pertinent to the production of quality data. These considerations are discussed in Section 3.0 of Volume III and include quality assurance guidelines in the areas of:

1. Planning the test program - preliminary plant survey; process information; stack data; location of sampling points; and cyclonic gas flow.
2. General factors involved in stationary source testing - tools and equipment; standard data forms; and identification of samples.
3. Chain-of-custody procedures for source sampling - sample collection; sample analysis; field notes; and report as evidence.
4. Traceability protocol for establishing true concentrations of gases used for calibrations and audits of air pollution analyzers - establishing traceability of commercial gas cylinders.

The remainder of Volume III contains quality assurance guidelines for specific measurement methods. The measurement systems currently in Volume III include:

Method 2 - Determination of Stack Gas Velocity and Volumetric Flow Rate (Type S Pitot Tube).

Method 3 - Determination of Carbon Dioxide, Oxygen, Excess Air, and Dry Molecular Weight.

- Method 4 - Determination of Moisture in Stack Gases.
- Method 5 - Determination of Particulate Emissions from Stationary Sources.
- Method 6 - Determination of Sulfur Dioxide Emissions from Stationary Sources.
- Method 7 - Determination of Nitrogen Oxide Emissions from Stationary Sources.
- Method 8 - Determination of Sulfuric Acid Mist and Sulfur Dioxide from Stationary Sources.
- Method 9 - Visible Determination of the Opacity of Emissions from Stationary Sources.
- Method 10 - Determination of Carbon Monoxide Emissions from Stationary Sources.
- Method 13A - Determination of Total Fluoride Emissions from and 13B Stationary Sources (SPADNS and Specific Ion Electrode).
- Method 17 - Determination of particulate emissions from stationary sources (in-stack filtration method).

As methods are added to Volume III, these will be sent to the users through the document control system used for the Handbook.

A separate volume in this series has been issued in bound format to provide guidance concerning quality assurance for meteorological measurements: EPA-600/4-82-060, February 1983, "Quality Assurance Handbook for Air Pollution Measurement Systems: Volume IV, Meteorological Measurements." This Volume IV is available from the National Technical Information Service, Springfield, Virginia.



### 1.3 DEFINITION OF QUALITY ASSURANCE<sup>1-6</sup>

Quality assurance and quality control have been defined and interpreted in many ways. Some authoritative sources differentiate between the two terms by stating that quality control is "the operational techniques and the activities which sustain a quality of product or service (in this case, good quality data) that meets the needs; also the use of such techniques and activities," whereas quality assurance is "all those planned or systematic actions necessary to provide adequate confidence that a product or service will satisfy given needs."<sup>1</sup>

Quality control may also be understood as "internal quality control;" namely, routine checks included in normal internal procedures; for example, periodic calibrations, duplicate checks, split samples, and spiked samples. Quality assurance may also be viewed as "external quality control," those activities that are performed on a more occasional basis, usually by a person outside of the normal routine operations; for example, on-site system surveys, independent performance audits, inter-laboratory comparisons, and periodic evaluation of internal quality control data. In this Handbook, the term quality assurance is used collectively to include all of the above meanings of both quality assurance and quality control.

While the objective of EPA's air programs is to improve the quality of the air, the objective of quality assurance for air programs is to improve or assure the quality of measured data, such as pollutant concentrations, meteorological measurements, and stack variables (e.g., gas velocity and mass emissions). Thus the "product" with which quality assurance is concerned is data.

Since air pollution measurements are made by numerous agencies and private organizations at a large number of field stations and laboratories, quality assurance is also concerned

with establishing and assessing comparability of data quality among organizations contributing to data bases.

#### 1.3.1 REFERENCES

1. Juran, J. M. Quality Control Handbook, 3rd Ed. McGraw-Hill, 1974. Section 2.
2. ASTM. Designation E548-79, "Recommended Practice for Generic Criteria for Use in the Evaluation of Testing and Inspection Agencies."
3. ANSI/ASQC. Standard A3-1978. "Quality Systems Terminology."
4. ANSI/ASQC. Standard Z1.15-1980. "Generic Guidelines for Quality Systems."
5. Feigenbaum, A. V. Total Quality Control, Engineering and Management. McGraw-Hill, 1961.
6. Canadian Standards Association. CSA Standard Z299,1-1978. Quality Assurance Program Requirements.

## 1.4 ELEMENTS OF QUALITY ASSURANCE

A quality assurance program for air pollution measurement systems should cover a number of areas or elements. These elements are shown in Figure 1.4.1 in a "Quality Assurance Wheel." The wheel arrangement illustrates the need for a quality assurance system that addresses all elements and at the same time allows program managers the flexibility to emphasize those elements that are most applicable to their particular program. Quality assurance elements are grouped on the wheel according to the organization level to which responsibility is normally assigned. These organizational levels are the quality assurance coordinator (normally a staff function), supervisor (a line function), and the operator. Together the supervisor and quality assurance coordinator must see that all these elements form a complete and integrated system and are working to achieve the desired program objectives.

The three-digit numbers shown on the wheel show the location in Section 1.4 where a description of the element is provided. Each element is described in three subsections as follows:

1. ABSTRACT - A brief summary that allows the program manager to review the section at a glance.
2. DISCUSSION - Detailed text that expands on items summarized in the ABSTRACT.
3. REFERENCES - List of resource documents used in preparation of the discussion. In addition, where applicable, a list of resource documents for recommended reading is shown under BIBLIOGRAPHY.

The DISCUSSION subsection is designed to be relatively brief. In those cases where a topic would require considerable detailed discussion, the reader is referred to the appropriate

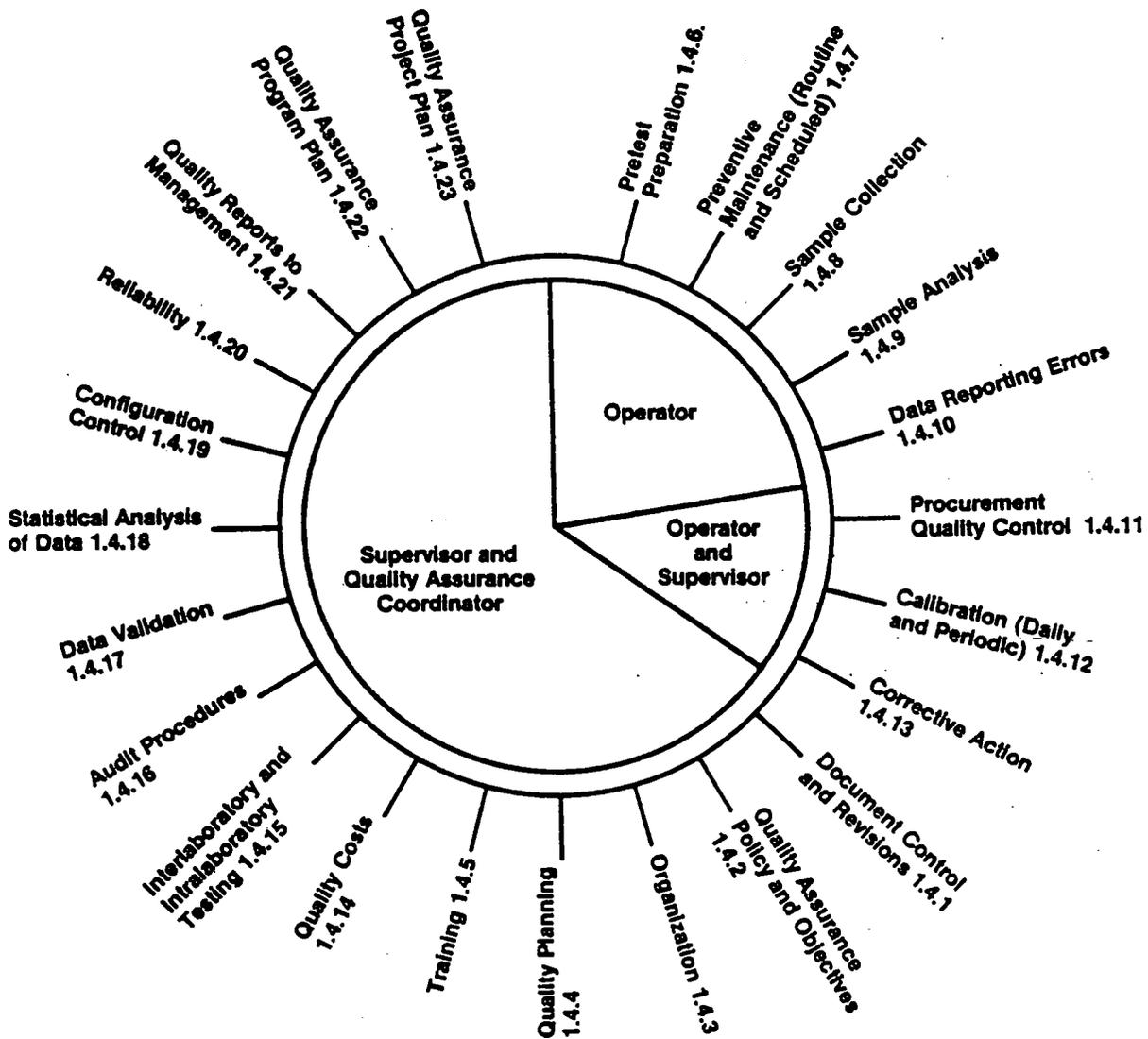


Figure 1.4.1. Quality assurance elements and responsibilities (the quality assurance wheel).

APPENDIX. A case in point is Section 1.4.18 on Statistical Analysis of Data. In this section the statistical methods are briefly summarized. For more details on the methods the reader is referred to the appropriate APPENDICES. For example, for statistical treatment of audit data the reader is referred to Appendix G.



#### 1.4.1 DOCUMENT CONTROL AND REVISIONS

##### 1.4.1.1 ABSTRACT

A quality assurance program should include a system for documenting operating procedures and subsequent revisions. The system used for this Handbook is described and is recommended.

##### 1.4.1.2 DISCUSSION

A quality assurance program should include a system for updating the formal documentation of operating procedures. The suggested system is the one used in this Handbook and described herein. This system uses a standardized indexing format and provides for convenient replacement of pages that may be changed within the technical procedure descriptions.

The indexing format includes, at the top of each page, the following information:

Section No.

Date

Page

A digital numbering system identifies sections within the text. The "Section No." at the top of each page identifies major three-digit or two-digit sections, where applicable. Almost all of the references in the text are to the section number, which can be found easily by scanning the top of the pages. References to subsections are used within a section. For example, Section 1.4.4 represents "Quality Planning" and Section 1.4.5 represents "Training." "Date" represents the date of the latest revision. "Page No." is the specific page in the section. The total number of pages in the section is shown in the "Table of Contents." An example of the page label for the first page of "Quality Planning" in Section 1.4.4 follows:

Section No. 1.4.4  
Date January 9, 1984  
Page 1

For each new three-digit level, the text begins on a new page. This format groups the pages together to allow convenient revision by three-digit section.

The Table of Contents follows the same structure as the text. It contains a space for total number of pages within each section. This allows the Handbook user to know how many pages are supposed to be in each section. When a revision to the text is made, the Table of Contents page must be updated. For example, the Table of Contents page detailing Section 1.4 might appear as follows:

	<u>Pages</u>	<u>Date</u>
1.4.1 Document Control and Revisions	5	1-9-84
1.4.2 Quality Assurance Policy and Objectives	4	1-9-84
1.4.3 Organization	7	1-9-84

A revision to "Organization" would change the Table of Contents to appear as follows:

	<u>Pages</u>	<u>Date</u>
1.4.1 Document Control and Revisions	5	1-9-84
1.4.2 Quality Assurance Policy and Objectives	4	1-9-84
1.4.3 Organization	9	6-2-88

A Handbook distribution record has been established and will be maintained up to date so that future versions of existing Handbook sections and the addition of new sections may be distributed to Handbook users. In order to enter the user's name and address in the distribution record system, the "Distribution Record Card" in the front of Volume I of this Handbook must be filled out and mailed to the EPA address shown. (Note:

A separate card must be filled out for each volume of the Handbook). Any future change in name and/or address should be sent to the following:

U.S. Environmental Protection Agency  
ORD Publications  
26 West St. Clair Street  
Cincinnati, Ohio 45268

Attn: Distribution Record System

Changes may be made by the issuance of (1) an entirely new document or (2) replacement of complete sections. The recipient of these changes should remove and destroy all revised sections from his/her copy.

The document control system described herein applies to this Handbook and it can be used, with minor revisions, to maintain control of quality assurance procedures developed by users of this Handbook and quality assurance coordinators. The most important elements of the quality assurance program to which document control should be applied include:

1. Sampling procedures.
2. Calibration procedures.
3. Analytical procedures.
4. Data analysis, validation, and reporting procedures.
5. Performance and system audit procedures.
6. Preventive maintenance.
7. Quality assurance program plans.
8. Quality assurance project plans.



## 1.4.2 QUALITY ASSURANCE POLICY AND OBJECTIVES

### 1.4.2.1 ABSTRACT

1. Each organization should have a written quality assurance policy that should be made known to all organization personnel.

2. The objectives of quality assurance are to produce data that meet the users' requirements measured in terms of completeness, precision, accuracy, representativeness and comparability and at the same time reduce quality costs.

### 1.4.2.2 DISCUSSION

Quality assurance policy - Each organization should have a written quality assurance policy. This policy should be distributed so that all organization personnel know the policy and scope of coverage.

Quality assurance objectives<sup>1,2,3</sup> - To administer a quality assurance program, the objectives of the program must be defined, documented, and issued to all involved in activities that affect the quality of the data. Such written objectives are needed because they:

1. Unify the thinking of those concerned with quality assurance.
2. Stimulate effective action.
3. Are a necessary prerequisite to an integrated, planned course of action.
4. Permit comparison of completed performances against stated objectives.

Data can be considered to be complete if a prescribed percentage of the total possible measurements is present. Precision and accuracy (bias) represent measures of the data quality. Data must be representative of the condition being measured.

Ambient air sampling at midnight is not representative of carbon monoxide levels during rush hour traffic. Stationary source emission measurements are not representative if measured at reduced load production conditions when usual operation is at full load. Data available from numerous agencies and private organizations should be in consistent units and should be corrected to the same standard conditions of temperature and pressure to allow comparability of data among groups.

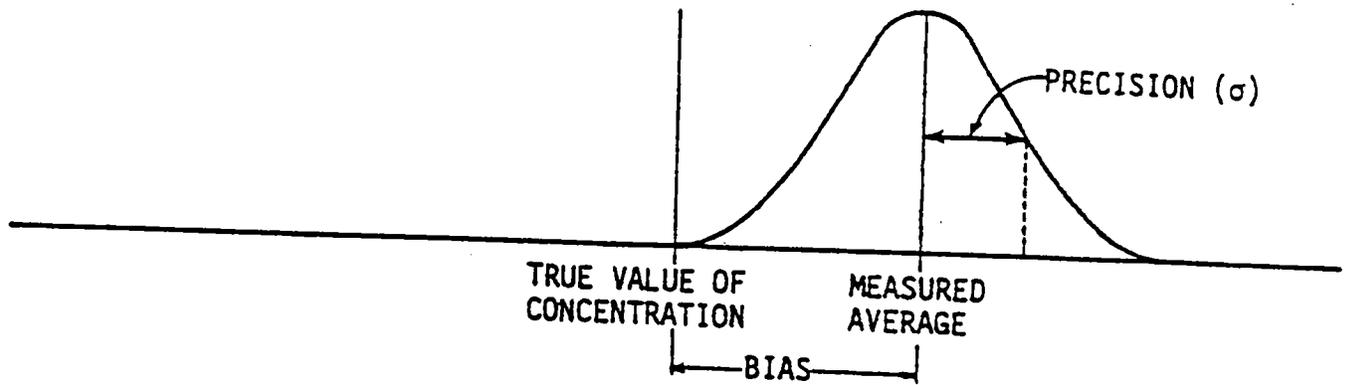
Figure 1.4.2.1 shows three examples of data quality with varying degrees of precision and bias. These examples hypothesize a true value that would result if a perfect measurement procedure were available and an infinitely large number of measurements could be made under specified conditions. If the average value coincides with the true value (reference standard), then the measurements are not biased. If the measurement values also are closely clustered about the true value, the measurements are both precise and unbiased. Figure 1.4.2.2 shows an example of completeness of data.

Each laboratory should have quantitative objectives set forth for each monitoring system in terms of completeness, precision, and bias of data. An example is included below for continuous measurement of carbon monoxide (nondispersive infrared spectrometry) to illustrate the point.

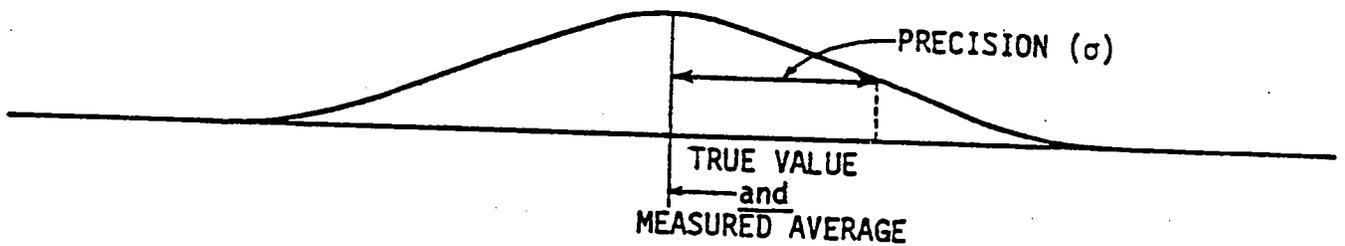
1. Completeness - For continuous measurements, 75 percent or more of the total possible number of observations must be present.<sup>4</sup>

2. Precision - Determined with calibration gases, precision is  $\pm 0.5$  percent full scale in the 0 through 58 mg/m<sup>3</sup> range.<sup>5,6</sup>

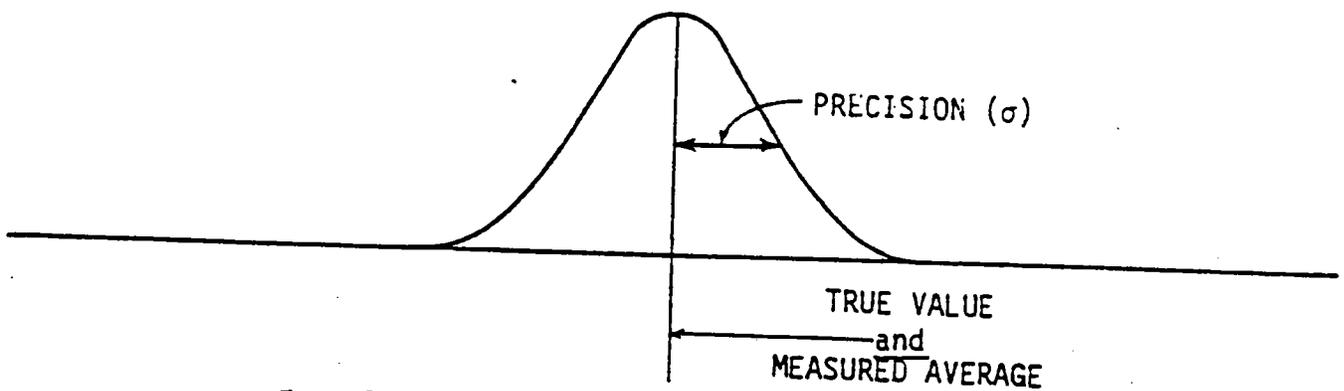
3. Accuracy - Depends on instrument linearity and the absolute concentrations of the calibration gases. An accuracy of  $\pm 1$  percent of full scale in the 0 through 58 mg/m<sup>3</sup> range can be obtained.<sup>5,6</sup>



Example of Positive Biased but Precise Measurements



Example of Unbiased but Imprecise Measurements



Example of Precise and Unbiased Measurements

Figure 1.4.2.1. Examples of data with varying degrees of precision and bias (normal distribution assumed).

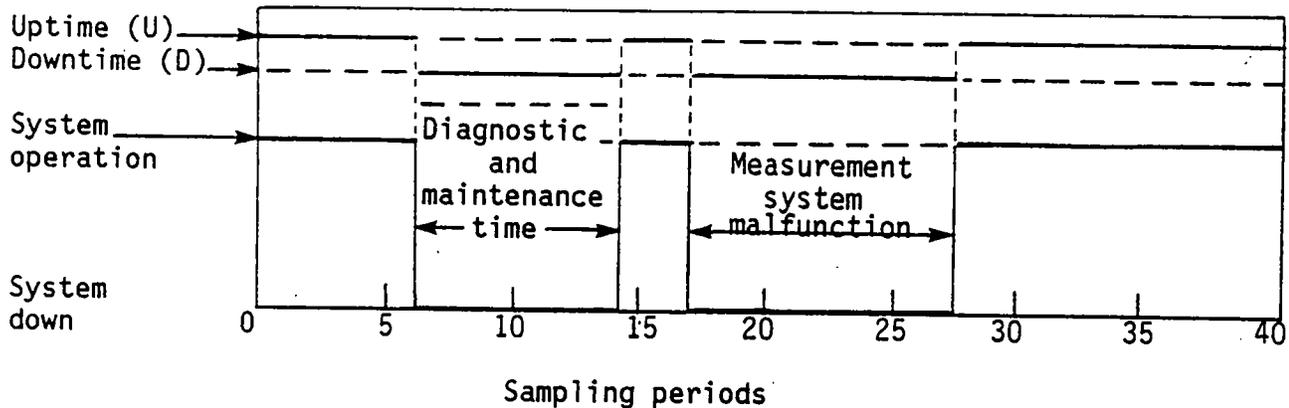


Figure 1.4.2.2. Example illustrating a measure of completeness of data,  $U/(D + U)$ .

For further discussion of completeness, precision, accuracy and comparability, see the following:

1. Completeness and comparability, Section 1.4.17 of this volume.
2. Precision and accuracy, Appendix G of this volume.

Employment of the elements of quality assurance discussed in Section 1.4 should lead to the production of data that are complete, accurate, precise, representative, and comparable.

#### 1.4.2.3 REFERENCES

1. Juran, J. M., (ed.). Quality Control Handbook. 3rd Ed. McGraw-Hill, New York, 1974. Sec. 2, pp. 4-8.
2. Feigenbaum, A. V. Total Quality Control. McGraw-Hill, New York, 1961. pp. 20-21.
3. Juran, J. M., and Gryna, F. M. Quality Planning and Analysis. McGraw-Hill, New York, 1970. pp. 375-377.
4. Nehls, G. J., and Akland, G. G. Procedures for Handling Aerometric Data. Journal of the Air Pollution Control Association, 23 (3):180-184, March 1973.
5. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS), Federal Register, Vol. 44, Number 92, May 10, 1979.
6. Appendix B - Quality Assurance Requirements for Prevention of Significant Deterioration (PSD) Air Monitoring, Federal Register, Vol. 44, Number 92, May 10, 1979.

### 1.4.3 ORGANIZATION

#### 1.4.3.1 ABSTRACT

1. Organizing a quality assurance function includes establishing objectives, determining the amount of emphasis to place on each quality assurance activity, identifying quality assurance problems to be resolved, preparing a quality assurance program and/or project plan, and implementing the plan.

2. The overall responsibility for quality assurance is normally assigned to a separate individual or group in the organization.

3. Quality assurance has input into many functions of an air pollution control agency. (See Figure 1.4.3.2 for details.)

4. The basic organizational tools for quality assurance implementation are:

- a. Organization chart and responsibilities.
- b. Job descriptions. (See Figure 1.4.3.3 for job description for the Quality Assurance Coordinator.)
- c. Quality assurance plan.

#### 1.4.3.2 DISCUSSION

Organizing the quality assurance function<sup>1</sup> - Because of the differences in size, workloads, expertise, and experience in quality assurance activities among agencies adopting the use of a quality assurance system, it is useful here to outline the steps for planning an efficient quality assurance system.

1. Establish quality assurance objectives (precision, accuracy, and completeness) for each measurement system (Section 1.4.2).

2. Determine the quality assurance elements appropriate for the agency (Figure 1.4.1).

3. Prepare quality assurance project plans for all measurement projects (Section 1.4.23).

4. Identify the quality assurance problems which must be resolved on the basis of the quality assurance project plan.

5. Implement the quality assurance project plan.

Location of the responsibility for quality assurance in the organization<sup>2</sup> - If practical, one individual within an organization should be designated the Quality Assurance (QA) Coordinator. The QA Coordinator should have the responsibility for coordinating all quality assurance activity so that complete integration of the quality assurance system is achieved. The QA Coordinator could also undertake specific activities such as quality planning and auditing. The QA Coordinator should, therefore, gain the cooperation of other responsible heads of the organization with regard to quality assurance matters.

As a general rule, it is not good practice for the quality assurance responsibility to be directly located in the organization responsible for conducting measurement programs. This arrangement could be workable, however, if the person in charge maintains an objective viewpoint.

Relationship of the quality assurance function to other functions - The functions performed by a comprehensive air pollution control program at the state or local level are shown in Figure 1.4.3.1.<sup>3</sup> The relationship of the quality assurance function to the other agency functions is shown in Figure 1.4.3.2. The role of quality assurance can be grouped into two categories:

1. Recommend quality assurance policy and assist its formulation with regard to agency policy, administrative support (contracts and procurements), and staff training.

2. Provide quality assurance guidance and assistance for monitoring networks, laboratory operations, data reduction and validation, instrument maintenance and calibration, litigation, source testing, and promulgation of control regulations.

Management Services

- Agency policy
- Administrative and clerical support
- Public information and community relations
- Intergovernmental relations
- Legal counsel
- Systems analysis, development of strategies, long-range planning
- Staff training and development

Technical Services

- Laboratory operations
- Operation of monitoring network
- Data reduction
- Special field studies
- Instrument maintenance and calibration

Field Enforcement Services

- Scheduled inspections
- Complaint handling
- Operation of field patrol
- Preparation for legal actions
- Enforcement of emergency episode procedures
- Source identification and registration

Engineering Services

- Calculation of emission estimates
- Operation of permit system
- Source emission testing
- Technical development of control regulations
- Preparation of technical reports, guides, and criteria on control
- Design and review of industrial emergency episode procedures

Figure 1.4.3.1. List of functions performed by comprehensive air pollution control programs.

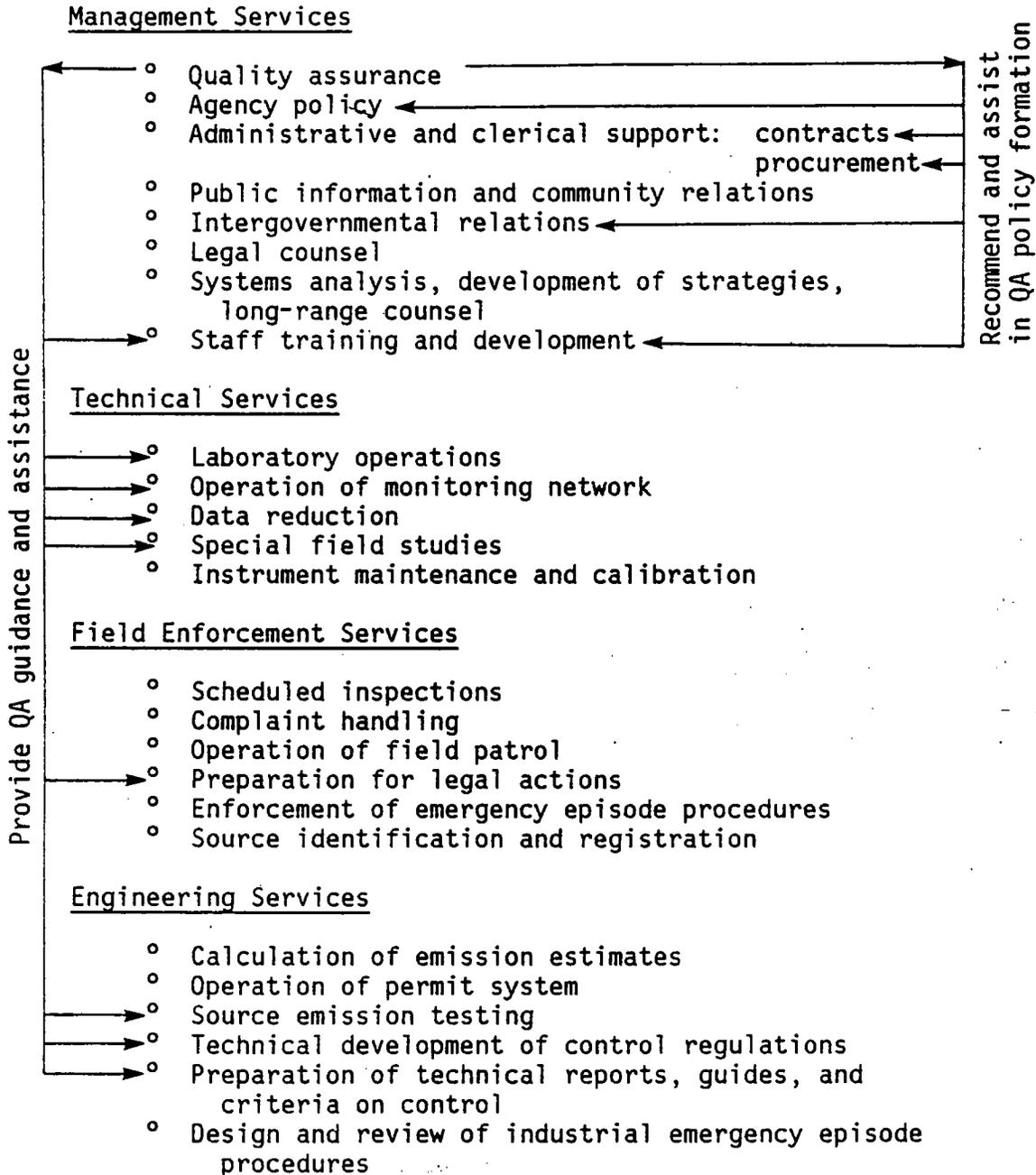


Figure 1.4.3.2. Relationship of the quality assurance function to other air pollution control program functions.

Basic organizational tools for quality assurance implementation are:

1. The organization chart<sup>4</sup> - The quality assurance organization chart should display line and staff relationships, and lines of authority and responsibility. The lines of authority and responsibility, flowing from the top to bottom, are usually solid, while staff advisory relationships are depicted by dashed lines.

2. The job description<sup>5</sup> - The job description lists the responsibilities, duties, and authorities of the job and relationships to other positions, individuals, or groups. A sample job description for a Quality Assurance Coordinator is shown in Figure 1.4.3.3.

3. The quality assurance plan - To implement quality assurance in a logical manner and identify problem areas, a quality assurance program plan and a quality assurance project plan are needed. For details on preparation of quality assurance program and project plans, see Sections 1.4.22 and 1.4.23, respectively.

TITLE: Quality Assurance Coordinator

Basic Function

The Quality Assurance Coordinator is responsible for the conduct of the quality assurance program and for taking or recommending measures.

Responsibilities and Authority

1. Develops and carries out quality control programs, including statistical procedures and techniques, which will help agencies meet authorized quality standards at minimum cost.
2. Monitors quality assurance activities of the agency to determine conformance with policy and procedures and with sound practice; and makes appropriate recommendations for correction and improvement as may be necessary.
3. Seeks out and evaluates new ideas and current developments in the field of quality assurance and recommends means for their application wherever advisable.
4. Advises management in reviewing technology, methods, and equipment, with respect to quality assurance aspects.
5. Coordinates schedules for measurement system functional check calibrations, and other checking procedures.
6. Coordinates schedules for performance and system audits and reviews results of audits.
7. Evaluates data quality and maintains records on related quality control charts, calibration records, and other pertinent information.
8. Coordinates and/or conducts quality-problem investigations.

Figure 1.4.3.3. Job description for the Quality Assurance Coordinator.

1.4.3.3 REFERENCES

1. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York. 1961. Chapter 4, pp. 43-82.
2. Covino, C.P., and Meghri, A.W. Quality Assurance Manual. Industrial Press, Inc., New York. 1967. Step 1, pp. 1-2.
3. Walsh, G.W., and von Lehmden, D.J. Estimating Manpower Needs of Air Pollution Control Agencies. Presented at the Annual Meeting of the Air Pollution Control Association, Paper 70-92, June 1970.
4. Juran, J.M., (ed.). Quality Control Handbook, 3rd Edition. McGraw-Hill, New York. 1974.
5. Industrial Hygiene Service Laboratory Quality Control Manual. Technical Report No. 78, National Institute for Occupational Safety and Health, Cincinnati, Ohio. 1974.

BIBLIOGRAPHY

1. Brown, F.R. Management: Concepts and Practice. Industrial College of the Armed Forces, Washington, D.C. 1967. Chapter II, pp. 13-34.



#### 1.4.4 QUALITY PLANNING

##### 1.4.4.1 ABSTRACT

1. Planning is thinking in advance the sequence of actions needed to accomplish a proposed objective and to communicate to the person or persons expected to execute these actions. Quality planning for air pollution measurements is designed to deliver acceptable quality data at a reasonable quality cost. Acceptable quality data is defined in terms of accuracy, precision, completeness, and representativeness.

2. The critical characteristics in the total measurement system must be identified and controlled. These critical characteristics may be located in any one or all of the following activities:

- a. Sample collection.
- b. Sample analysis.
- c. Data processing.
- d. Associated equipment and analyzers.
- e. Users, namely operators and analysts.

3. Interlaboratory collaborative test results and performance audits have been completed for many air pollution measurement methods. The studies/reports serve as a guide for the user of this Handbook for estimating the performance of the measurement methods.

Some of the results are used in the method descriptions in Volume II, Ambient-Air Specific Methods and in Volume III, Stationary-Source Specific Methods. Collaborative study reports available from EPA are listed in Figure 1.4.15.1. Interlaboratory performance audit reports are referenced herein.<sup>1,2</sup>

#### 1.4.4.2 DISCUSSION

Approach to planning - The act of planning is thinking in advance the sequence of actions necessary to accomplish certain objectives. In order that the planner may communicate his plan to the person or persons expected to execute it, the plan is written down with necessary criteria, diagrams, tables, etc.

Planning in the field of quality assurance for air pollution measurements must, of course, fundamentally be geared to deliver acceptable quality data at a reasonable quality cost. This objective is realized only by carefully planning many individual elements that relate properly to each other. The 23 elements which make up the Quality Assurance Wheel shown in Figure 1.4.1, are discussed in Sections 1.4.1 through 1.4.23 of this volume of the Handbook. These sections and in particular Section 1.4.23 can be used as a guide in developing the quality assurance project plan.

Specifications for data quality - The quality of data considered acceptable at each step of the measurement process must be defined as quantitatively as possible. The three basic measures of quality of air pollution data are accuracy, precision, and completeness. Acceptance limits should be established for these measures of data quality.

Acceptance limits for accuracy and precision of data are measurement method specific.<sup>3</sup> Recommended acceptance limits, when available, are given for each method in Volumes II and III of this Handbook. In addition, accuracy and precision data obtained by collaborative testing are available in EPA collaborative study publications (Figure 1.4.15.1 of Section 1.4.15).

Data quality assessments, in terms of overall precision, accuracy, and completeness, should be determined for the environmental data reported for each project or program.

Identification of critical characteristics - In the application of quality assurance measures, the total measurement system may be viewed as a complex system consisting of (1) the sample collection, (2) sample analyses, (3) data processing and

the associated test equipment or analyzers and (4) the operators and analysts. The critical characteristics of this complex system are identified by functional analysis of which ruggedness testing\* is an experimental form of analysis. The method activity matrices in each measurement method in Volumes II and III are tabulations of the most important operations/activities (not necessarily all critical) in the measurement method for which control may be required.

Development of a QA project plan<sup>4</sup> - The next step in the quality planning sequence is to determine which quality assurance elements shown on the Quality Assurance Wheel (Figure 1.4.1) should be included as part of a QA project plan. This is done by analyzing answers to questions posed to gain an understanding of the situation and needs. This analysis aids in the selection of the most productive steps leading toward accomplishment of the objectives concerning data quality. Such questions might be:

1. What activities should be considered?
2. Which of these activities are most critical?
3. What acceptance limits should be assigned to the activities, particularly the most critical ones?
4. How often should these activities be checked?
5. What methods of measurement should be used to check the activities?
6. What action should be taken if the acceptance limits for activities are not met?

The answers to many of these questions are the intended purpose of the activity matrices included in Volumes II and III.

---

\* A series of empirical tests performed to determine the sensitivity (hopefully to confirm the insensitivity) of a measurement system to specific operations/activities.

The finale of the quality planning process should be a written document which includes the most important information that should be communicated to the person or persons executing the plan. This is called the QA project plan. The recommended minimum content for a QA project plan is discussed in Section 1.4.23. The QA project plan serves three main functions:

1. The culmination of a planning cycle, the purpose of which is to design into a project or program necessary provisions to assure quality data.

2. A historical record that documents the project plan in terms of, for example, (1) measurement methods used, (2) calibration standards and frequencies planned, (3) auditing planned.

3. A document that can be used by the project officer, program manager, or quality assurance auditor to assess whether the QA activities planned are being implemented and their importance for accomplishing the goal of quality data.

#### 1.4.4.3 REFERENCES

1. Streib, E. W. and M. R. Midgett, A Summary of the 1982 EPA National Performance Audit Program on Source Measurements. EPA-600/4-83-049, December 1983.
2. Bennett, B. I., R. L. Lampe, L. F. Porter, A. P. Hines, and J. C. Puzak, Ambient Air Audits of Analytical Proficiency 1981, EPA-600/4-83-009, April 1983.
3. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS), Federal Register, Vol. 44, Number 92, May 10, 1979.
4. Feigenbaum, A. V. Total Quality Control. McGraw-Hill, New York. 1961. pp. 134-149.

BIBLIOGRAPHY

1. Sindelar, F. J. Management Planning and Controls for an Effective Quality Function. Industrial Quality Control XVIII, 3:28-29. September 1961.
2. Ewing, David W., (ed.). Long-Range Planning for Management. Harper and Row, New York. 1964.
3. LeBreton, P. P., and Henning, D. A. Planning Theory. Prentice-Hall, Englewood Cliffs, New Jersey. 1961.
4. Steiner, G. A., (ed.). Managerial Long-Range Planning. McGraw-Hill, New York. 1963.



## 1.4.5 TRAINING

### 1.4.5.1 ABSTRACT

All personnel involved in any function affecting data quality (sample collection, analysis, data reduction, and quality assurance) should have sufficient training in their appointed jobs to contribute to the reporting of complete and high quality data. The first responsibility for training rests with organizational management, program and project managers. In addition, the QA coordinator should recommend to management that appropriate training be available.

The training methods commonly used in the air pollution control field are the following:

1. On-the-job training (OJT).
2. Short-term course training (including self-instruction courses). A list of recommended short-term training courses is in Figure 1.4.5.1.
3. Long-term training (quarter or semester in length).

Training should be evaluated in terms of the trainee and the training per se. The following are techniques commonly used in the air pollution control field to evaluate training.

1. Testing (pretraining and posttraining tests).
2. Proficiency checks.
3. Interviews (written or oral with the trainee's supervisor and/or trainee).

### 1.4.5.2 DISCUSSION

All personnel involved in any function affecting data quality should have sufficient training in their appointed jobs to contribute to the reporting of complete and high quality data. The first responsibility for training rests with organizational management, program and project managers. In addition,

Course number and title	Days/h	Contact
<u>Quality Assurance /Quality Control Training</u>		
470 Quality Assurance for Air Pollution Measurement Systems	4	APTI <sup>a</sup>
556 Evaluation and Treatment of Outlier Data	3	NIOSH <sup>b</sup>
587 Industrial Hygiene Laboratory Quality Control	5	NIOSH
597 How to Write a Laboratory Quality Control Manual	3	NIOSH
---- Quality Management	5	UC <sup>c</sup>
9104 Quality Engineering	5	ETI <sup>d</sup>
9108 Quality Audit-Development and Administration	3	ETI
9101 Managing for Quality	5	ETI
9114 Probability and Statistics for Engineers and Scientists	5	ETI
9113 Managing Quality Costs	3	ETI
514Y Practical Application of Statistics to Quality Control	3	SAMI <sup>e</sup>
210Y Quality Management	5	SAMI
215Y Managing Quality Costs	3	SAMI
138Y Quality Program - Preparation and Audit	5	SAMI
919Y Software Quality Assurance	4	SAMI
284 Operating Techniques for Standards and Calibration	5	GWU <sup>f</sup>
641 Software Quality Assurance	3	GWU
---- Effective Quality Control Management	4	CPA <sup>g</sup>
---- Corporate Quality Assurance	3	MCQR <sup>h</sup>
<u>Air Pollution Measurement Method Training</u>		
413 Control of Particulate Emissions	4	APTI
415 Control of Gaseous Emissions	4	APTI
420 Air Pollution Microscopy	4.5	APTI
427 Combustion Evaluation	5	APTI
435 Atmospheric Sampling	4.5	APTI
444 Air Pollution Field Enforcement	3.5	APTI
450 Source Sampling for Particulate Pollutants	4.5	APTI
464 Analytical Methods for Air Quality Standards	5	APTI

Figure 1.4.5.1. Selected quality assurance and air pollution training available in 1984. (continued)

Course number and title	Days/h	Contact
<u>Air Pollution Measurement Method Training</u>		
468 Source Sampling and Analysis of Gaseous Pollutants	4	APTI
474 Continuous Emission Monitoring	5	APTI
<u>Air Pollution Measurement Systems Training</u>		
411 Air Pollution Dispersion Models: Fundamental Concepts	4.5	APTI
423 Air Pollution Dispersion Models: Application	4.5	APTI
426 Statistical Evaluation Methods for Air Pollution Data	4.5	APTI
452 Principles and Practice of Air Pollution Control	4.5	APTI
463 Ambient Air Quality Monitoring Systems: Planning and Administrative Concepts	5	APTI
482 Sources and Control of Volatile Organic Air Pollutants	4	APTI
<u>Self Instruction, Video-Instruction, and Other Training</u>		
406 Effective Stack Height/Plume Rise	10 h	APTI
422 Air Pollution Control Orientation Course (3rd Edition)	30 h	APTI
448 Diagnosing Vegetation Injury Caused by Air Pollution	30 h	APTI
473 Introduction to Environmental Statistics	70 h	APTI
472 Aerometric and Emissions Reporting System (AEROS)	-	APTI
475 Comprehensive Data Handling System (CDHS--AQDHS-II, EIS/P&R)	-	APTI
409 Basic Air Pollution Meteorology	25 h	APTI
410 Introduction to Dispersion Modeling	35 h	APTI
412A Baghouse Plan Review	20 h	APTI
414 Quality Assurance for Source Emission Measurements	35 h	APTI
416 Inspection Procedures for Organic Solvent Metal Cleaning (Degreasing) Operations	20 h	APTI
417 Controlling VOC Emissions from Leak Process Equipment	20 h	APTI
424 Receptor Model Training	30 h	APTI
431 Introduction to Source Emission Control	40 h	APTI
434 Introduction to Ambient Air Monitoring	50 h	APTI

Figure 1.4.5.1 (continued)

Course number and title	Days/h	Contact
436 Site Selection for Monitoring of SO <sub>2</sub> and TSP in Ambient Air	35 h	APTI
437 Site Selection for Monitoring of Photochemical Pollutants and CO in Ambient Air	35 h	APTI
412B Electrostatic Precipitator Plan Review	20 h	APTI
412C Wet Scrubbers Plan Review	-	APTI <sup>i</sup>
483A Monitoring the Emissions of Organic Compounds to the Atmosphere	-	APTI <sup>i</sup>
476A Transmissometer Operation and Maintenance	-	APTI <sup>i</sup>
438 Reference and Automated Equivalent Measurement Methods for Ambient Air Monitoring	30 h	APTI
443 Chain of Custody	2 h	APTI
453 Prevention of Significant Deterioration	15 h	APTI
449 Source Sampling Calculations	-	APTI
491A NSPS Metal-Coil Surface Coating	-	APTI <sup>i</sup>
491B NSPS Surface Coating of Metal Furniture	-	APTI <sup>i</sup>
491C NSPS Industrial Surface Coating	-	APTI <sup>i</sup>
491D NSPS Surface Coating Calculations	-	APTI <sup>i</sup>
428A NSPS Boilers	-	APTI <sup>i</sup>

Additional information may be obtained from:

<sup>a</sup>Air Pollution Training Institute, MD-20, Environmental Research Center, Research Triangle Park, North Carolina 27711, Attention: Registrar.

<sup>b</sup>R&R Associates, Post Office Box 46181, Cincinnati, Ohio 45246, Attention: Thomas Ratliff.

<sup>c</sup>The University of Connecticut, Storrs, Connecticut 06268.

<sup>d</sup>Education and Training Institute, American Society for Quality Control, 161 West Wisconsin Avenue, Milwaukee, Wisconsin 53203.

<sup>e</sup>Stat-A-Matrix Institute, New Brunswick, New Jersey.

<sup>f</sup>George Washington University, Continuing Engineering Education, Washington, D. C. 20052.

<sup>g</sup>The Center for Professional Advancement, Post Office Box H, East Brunswick, New Jersey 08816.

<sup>h</sup>Paul D. Krensky Associates, Inc., Adams Building, 9 Meriam Street, Lexington, MA 02173.

<sup>i</sup>Completion planned by October 1984.

the QA Coordinator should be concerned that the required training is available for these personnel and, when it is not, should recommend to management that appropriate training be made available.

Training objective<sup>1,2</sup> - The training objective should be to develop personnel to the necessary level of knowledge and skill required for the efficient selection, maintenance, and operation of air pollution measurement systems (ambient air and source emissions).

Training methods and availability - Several methods of training are available to promote achievement of the desired level of knowledge and skill required. The following are the training methods most commonly used in the air pollution control field; a listing of available training courses for 1984 is given in Figure 1.4.5.1.

1. On-the-job training (OJT) - An effective OJT program could consist of the following:

- a. Observe experienced operator perform the different tasks in the measurement process.
- b. Study the written operational procedures for the method as described in this Handbook (Volume II or III), and use it as a guide for performing the operations.
- c. Perform procedures under the direct supervision of an experienced operator.
- d. Perform procedures independently but with a high level of quality assurance checks, utilizing the evaluation technique described later in this section to encourage high quality work.

2. Short-term course training - A number of short-term courses (usually 2 weeks or less) are available that provide knowledge and skills for effective operation of an air pollution measurement system. Some of the courses are on the measurement methods per se and others provide training useful in the design

and operation of the total or selected portions of the measurement system. In addition, Figure 1.4.5.1 lists self-instruction courses and video-tapes available from:

Registrar  
Air Pollution Training Institute (MD-20)  
U.S. Environmental Protection Agency  
Research Triangle Park, North Carolina 27711  
(919) 541-2401

3. Long-term course training - Numerous universities, colleges, and technical schools provide long-term (quarter and semester length) academic courses in statistics, analytical chemistry, and other disciplines. The agency's training or personnel officer should be contacted for information on the availability of long-term course training.

Training evaluation - Training should be evaluated in terms of (1) level of knowledge and skill achieved by the trainee from the training; and (2) the overall effectiveness of the training, including determination of training areas that need improvement. If a quantitative performance rating can be made on the trainee during the training period (in terms of knowledge and skill achieved), this rating may also provide an assessment of the overall effectiveness of the training as well.

Several techniques are available for evaluating the trainee and the training per se. One or more of these techniques should be used during the evaluation. The most common types of evaluation techniques applicable to training in air pollution measurement systems are the following:

1. Testing - A written test before (pretest) and one after (post-test) training are commonly used in short-term course training. This allows the trainee to see areas of personal improvement and provides the instructor with information on training areas that need improvement.

2. Proficiency checks - A good means of measuring skill improvement in both OJT and short-term course training is to assign the trainee a work task. Accuracy and/or completeness

are commonly the indicators used to score the trainee's proficiency. The work tasks could be of the following form:

a. Sample collection - Trainee would be asked to list all steps involved in sample collection for a hypothetical case. In addition, the trainee could be asked to perform selected calculations. Proficiency could be judged in terms of completeness and accuracy.

b. Analysis - Trainee could be provided unknown samples for analysis. As defined here, an unknown is a sample whose concentration is known to the work supervisor (OJT) or training instructor (short-term course training) but unknown to the trainee. Proficiency could be judged in terms of accuracy of analysis.

c. Data reduction - Trainees responsible for data reduction could be provided data sets to validate. Proficiency could be judged in terms of completeness and accuracy.

If proficiency checks are planned on a recurring basis, a quality control or other type chart may be used to show progress during the training period as well as after the training has been completed. Recurring proficiency checks are a useful technique for determining if additional training may be required.

3. Interviews - In some cases, a written or oral interview with the trainee's supervisor and/or trainee is used to determine if the training was effective. This interview is normally not conducted until the trainee has returned to the job and has had an opportunity to use the training. This technique is most often used to appraise the effectiveness of a training program (OJT or short-term course) rather than the performance of the trainee.

#### 1.4.5.3 REFERENCES

1. Feigenbaum, A. V. Total Quality Control. McGraw-Hill, New York. 1961. pp. 605-615.

2. Feigenbaum, A. V. Company Education in the Quality Problem. Industrial Quality Control, X(6):24-29, May 1974.

#### BIBLIOGRAPHY

1. Juran, J. M., (ed.). Quality Control Handbook. 2nd edition. , McGraw-Hill, New York, 1962. Section 7, pp. 13-20.
2. Reynolds, E.A. Industrial Training of Quality Engineers and Supervisors. Industrial Quality Control, X(6):29-32, May 1954.
3. Industrial Quality Control, 23(12), June 1967. (All articles deal with education and training.)
4. Seder, L. A. QC Training for Non-Quality Personnel. Quality Progress, VII(7):9.
5. Reynolds, E. A. Training QC Engineers and Managers. Quality Progress, III(4):20-21, April 1970.

## 1.4.6 PRETEST PREPARATION

### 1.4.6.1 ABSTRACT

1. A common practice in both ambient air monitoring and stationary source emission monitoring includes pretest preparation for a new project. The proper selection of sampling sites and probe siting in ambient monitoring is fundamental in providing high quality and representative monitoring data. These activities are described in further detail in Volumes II and III and in References 1 and 2.

2. The pretest activities most important in ambient air monitoring system design are:

- a. Monitoring network size.
- b. Sampling station location.
- c. Probe siting.
- d. Method/equipment selection.

3. The pretest activities most important in stationary source emission monitoring system design are:

- a. Process design and operation familiarity.
- b. Measurements performed to gather data for design of the sample collection program.
- c. Monitoring of process to determine representative conditions of operation.

### 1.4.6.2 DISCUSSION

A common practice in both ambient air monitoring and stationary source emission monitoring includes pretest preparation for a new project. During the pretest preparation, an on-site visit may be conducted in order to complete administrative details for sample collection and to gather technical information for monitoring system design.

Ambient air monitoring system design - Factors that could be considered during ambient air monitoring system design are discussed in Volume II and in References 1 and 2. The items most important during the pretest preparation are summarized here.

1. Monitoring network size. The design of the monitoring network depends on the objective of the project. The objective is normally one of the following: compliance monitoring, emergency episode monitoring, trend monitoring, or research monitoring. In considering the location of the network, one or more of the following will be important considerations: monitoring must be pollution oriented; monitoring must be population oriented; monitoring must be source oriented; and/or monitoring must provide area-wide representation of air quality.

Criteria should be provided for new project monitoring network design. By way of example, criteria for design of the NAMS network for TSP and SO<sub>2</sub> are shown in Table 1.4.6.1. Table 1.4.6.2 shows the NAMS requirements for CO, O<sub>3</sub>, and NO<sub>2</sub>.

2. Sampling station location. The location of sampling stations within a monitoring network is influenced primarily by meteorological and topographic restraints. Meteorology (wind direction and speed) not only affect the geographical location of the sampling station but also the height of the sampling probe or sampler. Topographic features that have the greatest influence on final sampling station location are physical obstructions in the immediate area that may alter air flows, (e.g., trees, fences, and buildings). Criteria should be provided for sampling station location for new projects.

Providing project criteria for network and station design before the on-site inspection is an important factor in the success of the project and in the quality and representativeness of the data. Table 1.4.6.3 summarizes the probe siting criteria in Reference 2.

TABLE 1.4.6.1. SO<sub>2</sub> AND TSP NATIONAL AIR MONITORING STATION (NAMS) CRITERIA  
(Approximate number of stations/area)<sup>a</sup>

Population category	Concentration					
	High <sup>b</sup>		Medium <sup>c</sup>		Low <sup>d</sup>	
	SO <sub>2</sub>	TSP	SO <sub>2</sub>	TSP	SO <sub>2</sub>	TSP
High (>500,000)	6-8	6-8	4-6	4-6	0-2	0-2
Medium (100,000-500,000)	4-6	4-6	2-4	2-4	0-2	0-2
Low (50,000-100,000)	2-4	2-4	1-2	1-2	0	0

<sup>a</sup>This table is based on Reference 1. Urban areas and the number of stations/area will be jointly selected by EPA and the State agency.

<sup>b</sup>High concentration: SO<sub>2</sub> violating primary NAAQS; TSP violating primary NAAQS by  $\geq 20$  percent.

<sup>c</sup>Medium concentration: SO<sub>2</sub> violating 60 percent of primary NAAQS; TSP violating secondary NAAQS.

<sup>d</sup>Low concentration: SO<sub>2</sub> <60 percent of primary NAAQS; TSP less than secondary NAAQS.

TABLE 1.4.6.2. CO, O<sub>3</sub>, AND NO<sub>2</sub> NAMS CRITERIA<sup>a</sup>

Pollutant	Criteria
CO	<p>Two stations per major urban area:</p> <ul style="list-style-type: none"> <li>(1) one in a peak conc area (micro scale),</li> <li>(2) one in a neighborhood where conc exposures are significant (neighborhood scale).</li> </ul>
O <sub>3</sub>	<p>Two O<sub>3</sub> NAMS in each urban area having a population <u>&gt;</u>200,000.</p> <ul style="list-style-type: none"> <li>(1) one representative of maximum O<sub>3</sub> conc (urban scale),</li> <li>(2) one representative of high density population areas on the fringe of central business districts.</li> </ul>
NO <sub>2</sub>	<p>Two NO<sub>2</sub> NAMS in area having a population <u>&gt;</u>1,000,000.</p> <ul style="list-style-type: none"> <li>(1) one where emission density is highest (urban scale),</li> <li>(2) one downwind of the area of peak NO<sub>x</sub> emissions.</li> </ul>

<sup>a</sup>This table is based on Reference 1.

TABLE 1.4.6.3. SUMMARY OF PROBE SITING CRITERIA<sup>2</sup>

Pollutant	Scale	Height above ground, meters	Distance from supporting structure, meters		Other spacing criteria
			Vert	Horiz <sup>a</sup>	
TSP	All	2 - 15	--	>2	<ol style="list-style-type: none"> <li>Should be &gt;20 meters from trees</li> <li>Distance from sampler to obstacle, such as buildings, must be at least twice the height the obstacle protrudes above the sampler</li> <li>Must have unrestricted airflow 270° around the sampler</li> <li>No furnace or incineration flues should be nearby<sup>c</sup></li> <li>Must have minimum spacing from roads; this varies with height of monitor and spatial scale</li> </ol>
SO <sub>2</sub>	All	3 - 15	>1	>1	<ol style="list-style-type: none"> <li>Should be &gt;20 meters from trees</li> <li>Distance from inlet probe to obstacle, such as buildings, must be at least twice the height the obstacle protrudes above the inlet probe<sup>d</sup></li> <li>Must have unrestricted airflow 270° around the inlet probe, or 180° if probe is on the side of a building</li> <li>No furnace or incineration flues should be nearby<sup>c</sup></li> </ol>
CO	Micro	3 ± 1/2	>1	>1	<ol style="list-style-type: none"> <li>Must be &gt;10 meters from intersection and should be at a mid-block location</li> <li>Must be 2-10 meters from edge of nearest traffic lane</li> <li>Must have unrestricted airflow 180° around the inlet probe</li> </ol>
	Middle, neighborhood	3 - 15	>1	>1	<ol style="list-style-type: none"> <li>Must have unrestricted airflow 270° around the inlet probe, or 180° if probe is on the side of a building</li> <li>Spacing from roads varies with traffic</li> </ol>

(continued)

Table 1.4.6.3 (continued)

Pollutant	Scale	Height above ground, meters	Distance from supporting structure, meters		Other spacing criteria
			Vert	Horiz <sup>a</sup>	
O <sub>3</sub>	A11	3 - 15	>1	>1	<ol style="list-style-type: none"> <li>1. Should be &gt;20 meters from trees</li> <li>2. Distance from inlet probe to obstacle, such as buildings, must be at least twice the height the obstacle protrudes above the inlet probe</li> <li>3. Must have unrestricted airflow 270° around the inlet probe, or 180° if probe is on the side of a building</li> <li>4. Spacing from roads varies with traffic</li> </ol>
NO <sub>2</sub>	A11	3 - 15	>1	>1	<ol style="list-style-type: none"> <li>1. Should be &gt;20 meters from trees</li> <li>2. Distance from inlet probe to obstacle, such as buildings, must be at least twice the height the obstacle protrudes above the inlet probe<sup>b</sup></li> <li>3. Must have unrestricted airflow 270° around the inlet probe, or 180° if probe is on the side of a building</li> <li>4. Spacing from roads varies with traffic</li> </ol>

<sup>a</sup>When probe is located on rooftop, this separation distance is in reference to walls, parapets, or penthouses located on the roof.

<sup>b</sup>Sites not meeting this criterion would be classified as middle scale (see text).<sup>2</sup>

<sup>c</sup>Distance is dependent on height of furnace or incineration flue, type of fuel or waste burned, and quality of fuel (sulfur and ash content). This is to avoid undue influences from minor pollutant sources.

Stationary source emission system design - Factors that should be considered during stationary source emission monitoring system design are discussed in Volume III. The items most important during the pretest preparation are:

1. Familiarization with process design and operation. The success of source emission monitoring requires familiarity with the process to be monitored. The following are areas in which familiarity is particularly important:
  - a. Process operation principle.
  - b. Process flow chart.
  - c. Variability of process operation in terms of flow rate of effluent to be monitored and concentration of pollutant in the effluent.
  - d. Process data that must be collected during sample collection (e.g., fuel burning rate).
  - e. Identify the key parameters and their representative levels of operation.
  - f. Sample collection site considerations, including sample site location (no turbulence due to upstream obstructions or bends), sample port access and size, size of platform for sample collection work, and utilities availability for sample collection equipment.
2. Measurements needed for design of the sample collection program - During the on-site visit, certain measurements are normally made that are required for the design of the sample collection program. These measurements include:
  - a. Dimensions of the stack or duct cross-section so that a sampling plan by equal cross-sectional areas can be determined.
  - b. Gas velocity, gas temperature, and gas moisture content so that requirements for isokinetic sampling can be calculated.

The selection of proper sampling sites and preliminary measurements required for monitoring system design during the pretest preparation is fundamental in providing monitoring data that are both high quality and representative.

#### 1.4.6.3 REFERENCES

1. Appendix D - Network Design for State and Local Air Monitoring Stations (SLAMS) and National Air Monitoring Stations (NAMS), Federal Register 40 CFR 58, Number 92, May 10, 1979, p. 27586-27592.
2. Appendix E - Probe Siting Criteria for Ambient Air Quality Monitoring, Federal Register 40 CFR 58, Number 92, May 10, 1979, p. 27592-27597.

## 1.4.7 PREVENTIVE MAINTENANCE

### 1.4.7.1 ABSTRACT

1. The most important benefit of a good preventive maintenance program is to increase measurement system availability (proportion of up time) and thus increase data completeness. In addition, the quality of the data should improve as a result of good equipment operation.

2. Continuous pollutant analyzers commonly require daily service checks. An example of a daily checklist is shown in Figure 1.4.7.1. Daily service checks, preventive maintenance, and calibration should be integrated into a schedule. An example of a combined preventive maintenance-calibration schedule is shown in Figure 1.4.7.2.

### 1.4.7.2 DISCUSSION

Importance of preventive maintenance - As defined here, preventive maintenance is an orderly program of positive actions (equipment cleaning, lubricating, reconditioning, scheduled replacement, adjustment and/or testing) for minimizing the failure of monitoring systems or parts thereof during use. The most important benefit of a good preventive maintenance program is to increase measurement system availability and thus increase data completeness. Conversely, a poor preventive maintenance program will result in increased downtime (i.e., decrease in data completeness) and in increased unscheduled maintenance costs; and may yield invalid data.

Preventive maintenance schedule - Project officers should prepare and implement a preventive maintenance schedule for measurement systems. The planning required to prepare the preventive maintenance schedule will have the effect of:

City \_\_\_\_\_

Site location \_\_\_\_\_

Site number \_\_\_\_\_

Serial number \_\_\_\_\_

At last calibration	NO	NO <sub>2</sub>
Scale range		
Zero knob setting		
Span knob setting		

Date last calibration \_\_\_\_\_

Date								
Oxygen pressure - inst.								
Oxygen cylinder pressure								
Unadjusted NO zero reading								
Unadjusted NO <sub>2</sub> zero reading								
NO zero knob setting (new)								
NO span knob setting								
NO <sub>2</sub> zero knob setting (new)								
NO <sub>2</sub> span knob setting								
√ Valve in NO-NO <sub>2</sub> -NO <sub>x</sub> position								
√ NO-NO <sub>2</sub> range in 0.5 position								
√ Inspect oxygen line								
√ Inspect inlet line, probe, and filter holder								
Vacuum gauge reading								
Comments or problems:								
Operator initial								

Figure 1.4.7.1. Daily checklist for NO<sub>2</sub> analyzer.

	1	2	3	4	5	6	7	8	9	10
0		1,7	5,14, 27	12,13	2,38, 40	6,52	23,26, 28	3	6	31,32
1	33,34	4,49, 50,53	43,45, 47	9,10, 11,28	1	7	5,27		2	8
2	23,26, 28	3	6	21,22, 30,31	33,34	4,46 53	51,52	9,10, 11,28	1	7
3	5,27	55	2	3,49 50	23,26, 28	3	6	31,32	33,34, 35	4,53
4	45,47 48	9,10, 11,28	1	7	5,27		2	8,52	23,26, 28	3
.										
.										
.										
34	2	8	23,26	3,28	6	33,34	31,32	4	45,47, 48,53	9,10, 11,33, 34,52
35	1,28	7	5,54	27	2	8,49, 50	23,26	3,38	6	
36	31	4	53	25	9,10, 11,28					

Task number	Task	Operations manual reference
1	Calibrate MF-1D, MF-2D, MF-3D, MF-2G, MF-1H, and MF-2V	3.2.2
2	Calibrate MF-1G, MF-1V, and MF-1C	3.2.2
3	Calibrate MF-2H, MF-1S, and MF-4H	3.2.2
4	Calibrate MF-3H and MF-1P	3.2.2
5	Calibrate NO <sub>2</sub> analyzer (5-point)	3.2.1
6	Calibrate SO <sub>2</sub> analyzer (5-point)	3.2.1
7	Calibrate NO and NO <sub>x</sub> analyzers (5-point)	3.2.1
8	Calibrate O <sub>3</sub> analyzer (5-point)	3.2.1
9	Inspect compressor and vacuum pumps shock mounts	3.2.9
10	Clean compressor intake filters	3.2.10
11	Change pump box filters	3.2.11
12	Clean vacuum panel filter	3.2.20
13	Replace vacuum pump filters	3.2.12
14	Replace vacuum pump vanes	Refer to vendors literature

Figure 1.4.7.2. Combined preventive maintenance-calibration schedule (by Julian date) and tasks for a major EPA ambient-air monitoring project. (continued)

Task number	Task
15	Replace vacuum relief valve
16	Replace dryer ball check valve
17	Clean dryer solenoid valves
18	Clean dryer check valve
19	Replace water trap automatic drain and filter
20	Clean dryer relief valve
21	Replace compressor rings, valves, and springs
22	Lubricate pump box blower motors
23	Replace H <sub>2</sub> /O <sub>2</sub> generator water tank
24	Service and adjust H <sub>2</sub> /O <sub>2</sub> generator
25	Change molecular sieves
26	Leak check H <sub>2</sub> /O <sub>2</sub> generator
27	Inspect sample lines
28	Replace analyzers sample filters
29	Clean sample manifold
30	Lubricate sample manifold blower motor
31	Leak check calibration system
32	Leak check pressure panel
33	Change data tape
34	Clean tape deck transport mechanism
35	Check tape deck skew and tape tracking
36	Check tape deck head wear
37	Replace tape deck reel motors and capstan motor
38	Clean air conditioner filters
39	Lubricate air conditioner motors
40	Check air conditioner cabinet water drain
41	Clean air conditioner coil
42	Wax air conditioner cabinet
43	Calibrate wind speed instrument
44	Calibrate wind direction instrument
45	Fill water bath
46	Change SO <sub>2</sub> permeation tube
47	Check meteorological tower guy line tension
48	Inspect meteorological tower and instruments
49	Replace particulate manifold motor
50	Replace hi-vol motor
51	Replace NO-NO <sub>x</sub> analyzer exhaust filter (charcoal)
52	Calibrate ambient temperature sensor
53	Calibrate relative humidity sensor
54	Calibrate Xincom
55	Weigh fire extinguishers

(Note: MF means mass flow meter.)

Figure 1.4.7.2 (continued)

1. Highlighting that equipment or those parts thereof that are most likely to fail without proper preventive maintenance or scheduled replacement.

2. Defining a spare parts inventory that should be maintained to replace worn-out parts with a minimum of downtime.

A specific preventive maintenance schedule should relate to the purpose of testing, environmental influences, physical location of analyzers, and the level of operator skills. Checklists are commonly used to list specific maintenance tasks and frequency (time interval between maintenance).

Continuous pollutant analyzers commonly require daily service checks. By way of example, Figure 1.4.7.1 is a daily instrument service checklist and record used in a major EPA monitoring network for ambient air measurement of NO<sub>2</sub>. In this particular monitoring network, ambient air measurement data for NO<sub>2</sub> are recorded on strip charts and sent weekly to a central location for validation and data reduction. The instrument checklist shown in Figure 1.4.7.1 is sent with the strip charts. This checklist provides a record of daily service checks and, in addition, includes operations data used in the validation of the strip-chart data.

When a project includes several sensors (air pollution and/or meteorological), it becomes important to integrate checklists into a preventive maintenance schedule. Since calibration sensors are commonly the responsibility of the operator in addition to preventive maintenance, and since calibration tasks may be difficult to separate from preventive maintenance tasks, a combined preventive maintenance-calibration schedule is often advisable. An example of a combined preventive maintenance-calibration schedule for a major EPA ambient air monitoring project is shown in Figure 1.4.7.2.

This combined schedule is read as follows: Numbers along the vertical (1-36) and horizontal (1-10) part of the schedule refer to the Julian date. The numbers in the boxes indicate the

tasks to be completed according to Julian date. The following should be noted in this combined schedule:

1. The schedule is prepared by Julian date, not calendar date. For example, Julian date 10 is January 10 and Julian date 33 is February 2 (i.e., 33 days into the calendar year).

2. The operator is provided a reference to the project operations manual for the exact procedure to follow.

Preventive maintenance records - A record of all preventive maintenance and daily service checks should be maintained. This record should be coordinated with the record on equipment reliability (failures and unscheduled maintenance) for the purpose of coordinating and assessing the overall equipment maintenance program (Section 1.4.20).

Normally the daily service checklists shown in Figure 1.4.7.1 are stored with the measurement data. An acceptable practice for recording completion of tasks listed in Figure 1.4.7.2 is to maintain a preventive maintenance-calibration duplicate copy record book. After tasks have been completed and entered in the record book, a duplicate copy of each task is removed by the operator and sent to the supervisor for review and filing. The record book is stored at the sampling station for future reference.

#### 1.4.7.3 BIBLIOGRAPHY

1. EPA, Office of Air Programs. Field Operations Guide for Automatic Air Monitoring Equipment. Publication No. APTD-0736. Research Triangle Park, NC. 27711 1971.
2. Hubert, C. I. Preventive Maintenance of Electrical Equipment. McGraw-Hill, New York, N.Y. 1955. Part 1.
3. Spencer, J. Maintenance and Servicing of Electrical Instruments. Instruments Publishing Co., Inc., Pittsburgh, PA. 1951.

## 1.4.8 SAMPLE COLLECTION

### 1.4.8.1 ABSTRACT

1. Presample and postsample collection checks should be performed by the operator on the sample collection system. These checks are measurement-method specific. Recommended checks for each sample collection system are given in each method in Volumes II and III of this Handbook. These checks are summarized in the activity matrices at the ends of the appropriate sections.

2. Control checks should be conducted during the sample collection to determine the system performance.

3. Control charts should be used to record results from selected control checks to determine when the sample collection system is out-of-control and to record the corrective action taken.

### 1.4.8.2 DISCUSSION

Presample and postsample collection checks - Checks should be made by the operator prior to the actual sample collection. Presample collection checks normally include:

1. A leak check on the sample collection system. Since the entire sample collection system is normally under a vacuum, it is very important to be sure that all parts of the system are assembled so that air does not leak into the sample gas stream.

2. Specific checks on components of the sample collection system, such as the liquid level in bubblers.

The amount of postsample collection checking is dependent on the sample collection technique. When samples are collected for subsequent laboratory analyses, selected postsample checks and the inspection of the sample per se are a common practice.

When continuous air pollution sensors are used, sensor calibration checks replace the postsample collection checks. See Section 1.4.12 for more information on sensor calibration.

Control checks during sample collection - Control checks should be conducted during the actual sample collection to determine the system performance. For example, some operational checks for TSP hi-vol measurement are initial and final flow rate checks, timer checks, and filter checks for signs of air leakage.

Control charts - Results from selected control checks should be recorded on control charts. This will help the operator and supervisor to determine when the sample collection system is out-of-control and will provide a record of corrective action taken. Periodic zero and span checks conducted on continuous air pollution sensors are good examples of data which should be plotted on control charts.

## 1.4.9 SAMPLE ANALYSIS

### 1.4.9.1 ABSTRACT

1. Function checks should be conducted by the analyst to check the validity of the sample and performance of the equipment.

2. Control checks should be conducted during the analyses to:

- a. determine the performance of the analytical system.
- b. estimate the variability in results from the analytical system in terms of precision.

3. Control charts should be used to record results from selected quality control checks to determine when the analytical system is out-of-control and to record the corrective action taken.

### 1.4.9.2 DISCUSSION

Function checks - Function checks are performed to verify the stability and validity of the sample and the performance of the analytical equipment. The analyst should be provided with written performance specifications for each function check, accompanied by recommended action if the specifications are not met.

Control checks - Control checks should be performed during analysis. These checks are made on all analyses or intermittently after a specified number of analyses.

Some control checks are part of the routine analysis and are performed by the analyst to determine the performance of the analytical system. These control checks include the use of

sample blanks to observe zero concentration drift; spiked samples to determine percentage of sample recovery during intermediate analysis extraction steps; and sample aliquots to observe within- and between-run variability for the entire analysis.

Other control checks are performed by the analyst intermittently to estimate analysis variability in terms of precision and accuracy. Control samples normally used for this purpose are sample aliquots to determine precision and standard reference materials and standard reference samples to determine accuracy. A list of NBS environmental standard reference materials is shown in Figure 1.4.12.3 of Section 1.4.12.

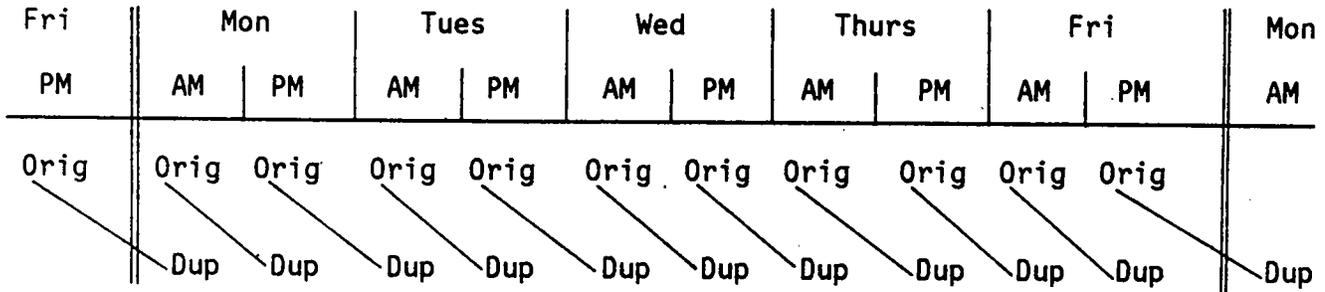
Analysis precision must be explained in terms of possible sources of variability to make this measurement most meaningful. The three common ways to report precision, are: replicability, repeatability, and reproducibility, Table 1.4.12.1. A more detailed discussion of these three forms of precision is included in Appendix A. The following summarizes those source variables that are the same or different for each form of precision.

Source of variability	Replicability	Repeatability	Reproducibility
Sample	Same	Same	Same
Analyst	Same	At least one of these three sources must be different	Different
Apparatus	Same		Different
Day	Same		Same or different
Laboratory	Same	Same	Different

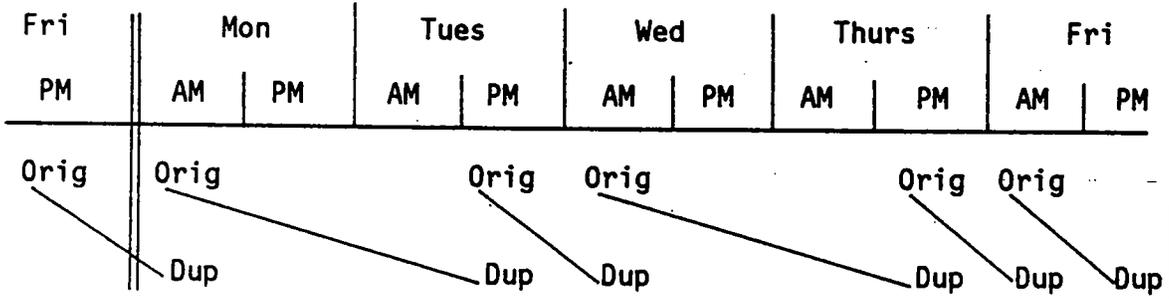
A good scheme using environmental samples to determine repeatability is shown in Figure 1.4.9.1. The procedure involves duplicate analyses performed on a staggered time schedule, which allows for a better appraisal of laboratory variability than if the duplicate is analyzed immediately after the original sample. The number of duplicates removed for analysis is dependent on the number of analyses performed during the a.m. or p.m. time cycle shown in Figure 1.4.9.1. The number of duplicates removed for analysis would normally range from 1 to 5, with a minimum of one duplicate being analyzed for every a.m. or p.m. time span in which a sample set is analyzed. If the variability of the analyst is being monitored during routine analysis, efforts should be made to submit control samples in a manner so that the analyst does not give them special attention.

Control charts - Results from selected function checks and control samples should be recorded on control charts. This will allow the analyst and supervisor to know exactly when the analytical system is out-of-control, which part of the system is the probable cause, and when a corrective action is taken.

A. Analysis performed every day



B. Analysis performed intermittently



Orig - analysis on original sample.  
Dup - analysis on duplicate of original sample.

Figure 1.4.9.1. Analysis of duplicate samples.

## 1.4.10 DATA REPORTING ERRORS

### 1.4.10.1 ABSTRACT

1. Human error is the most common source of error in data reporting. However, individuals using measurement systems with data logging devices to automate data handling from continuous sensors should be concerned that only the sensor analog signal and not electronic interferences are converted during the digital readout.

2. Data validation procedures (Section 1.4.17) should be used for reviewing data at the operational, as well as the managerial levels. Control charts (Appendix H) are a common tool used to review data from critical characteristics in a measurement system.

### 1.4.10.2 DISCUSSION

Source of data reporting errors - Measurements of the concentration of pollutants, either in the ambient atmosphere or in the emissions from stationary sources, are assumed to be representative of the conditions existing at the time of the sample collection. The extent to which this assumption is valid depends on the sources of error and bias inherent in the collection, handling, and analysis of the sample.

Besides the sampling and analytical error and bias, human error may be introduced any time between sample collection and sample reporting. Included among the human errors are such things as failure of the operator/analyst to record pertinent information, mistakes in reading an instrument, mistakes in calculating results, and mistakes in transposing data from one record form to another. Data handling systems involving the use of computers are susceptible to keypunching errors and errors

involving careless handling of magnetic tapes and other storage media. Although it cannot be completely avoided, human error can be minimized.

Data reporting techniques and error sources depend on the type of sensor measurement system - Measurement sensors for pollutant concentration and meteorological conditions may be classified by their sample collection principle into two categories: (1) integrated, and (2) continuous. Pollutant measurement systems may be either integrated or continuous, whereas meteorological measurement systems are normally always continuous.

In the integrated sample collection principle, a discrete sample is collected in some medium and is normally sent to a laboratory for analysis. Both the field operator and the laboratory analyst can make errors in data handling.

In the continuous sample collection principle, an analytical sensor produces a direct and continuous readout of the pollutant concentration or meteorological parameter. The readout may be a value punched on paper tape or recorded on magnetic tape. In addition, some continuous measurement systems may also use telemetry to transmit data to a data processing center. Both human and machine errors can occur in data handling in this type of system.

Data errors in integrated sampling - For ambient air monitoring, the operator records information before and after the sample collection period. For example, acceptance limits are recommended for flow rate data for hi-vol measurement of TSP and the operator should invalidate or "flag" data when values fall outside these limits. These limits are included as part of the measurement method descriptions and are in the activity matrices at the ends of pertinent sections of Volumes II and III of the Handbook. Questionable measurement results (outside of acceptance limits) may indicate the need for calibration or maintenance.

The analyst in the laboratory reads measurements from balances, colorimeters, spectrophotometers, and other instruments; and records the data on standard forms or in laboratory notebooks. Each time values are recorded, there is a potential for incorrectly entering results. Typical recording errors are transposition of digits (e.g., 216 might be incorrectly entered as 126) and incorrect decimal point location (e.g., 0.0635 might be entered as 0.635). These kinds of errors are difficult to detect. The supervisors must continually stress the importance of accuracy in recording results.

Acceptance limits contained in the measurement method write-up and those shown in the method activity matrices should be used by the analyst to invalidate or "flag" analysis data when values fall outside these limits. In addition, data validation procedures should be used to identify questionable data (Section 1.4.17).

Data errors in continuous sampling - Continuous air monitoring systems may involve either manual or automated data recording. Automated data recording may involve the use of a data logging device to record data on paper tape or magnetic tape at the remote sampling station, or the use of telemetry to transmit data on-line to a computer at a central facility.

Manual reduction of pollutant concentration data from strip charts can be a significant source of data errors. In addition to making those errors associated with recording data values on record forms, the individual who reads the chart can also err in determining the time-average value. When the temporal variability in concentration is large, it is difficult to estimate an average concentration. Two people reading the same chart may yield results that vary considerably.

Technicians responsible for reducing data from strip charts should be given training. After a technician is shown how to read a chart, his results should be compared with those of an experienced technician. Only after he has demonstrated the

capability to obtain satisfactory results should a technician be assigned to a data reduction activity.

Periodically a senior technician should check strip charts read by each technician. Up to 10 percent of the data reported by each technician should be checked, depending on time availability and past history of error occurrence. If an individual is making gross errors, additional training may be necessary.

Because manual chart reading is a tedious operation, a drop in productivity and an increase in errors might be expected after a few hours. Ideally, an individual should be required to spend only a portion of a day at this task.

The use of a data logging device to automate data handling from a continuous sensor is not a strict guarantee against data recording errors. Internal validity checks are necessary to avoid serious data recording errors. There are two sources of error between the sensor and the recording medium: (1) the output signal from the sensor and (2) the errors in recording by the data logger.

The primary concern about the sensor output is to ensure that only the sensor analog signal and not electronic interferences be converted to a digital readout. Internal validity checks should be planned to "flag" spurious data resulting from electronic interferences.

For a system involving the use of telemetry, it is also necessary to include a validity check for data transmission.

Errors in computations - To minimize computational errors, operators and analysts should follow closely the formulae, calculation steps, and examples given for each method, using the calculation instructions and forms provided. As an example, Volume II provides instructions for calculations of air volume and air concentration for the Hi-Vol Method. Recommended audits on computations are given for each method in Volumes II and III.

Control charts - Procedures for reviewing data at the operational as well as the managerial levels should be provided.

Review of measurement results from control samples used during analysis, for example, can indicate out-of-control conditions that would yield invalid data from subsequent analyses, if the conditions are not corrected immediately. At the managerial level, periodic review of data can indicate trends or problems that need to be addressed to maintain the desired level of precision and accuracy. One common tool for statistical analysis of data at both the operational and the managerial levels is the control chart. The major steps in constructing the control chart are outlined in Appendix H. A typical control chart for sample average and indicated actions is in Figure 1.4.10.1.

#### 1.4.10.3 REFERENCES

1. Control Chart Method for Controlling Quality During Production. ANSI/ASQC Standard B3-1958 (reaffirmed in 1975).

#### BIBLIOGRAPHY

1. Kelley, W. D. Control Chart Techniques. Statistical Method-Evaluation and Quality Control for the Laboratory. August 1968. U.S. DHEW, Public Health Service, Cincinnati, Ohio. p. 3.

LABORATORY QUALITY CONTROL CHART

Operation \_\_\_\_\_ Dates \_\_\_\_\_

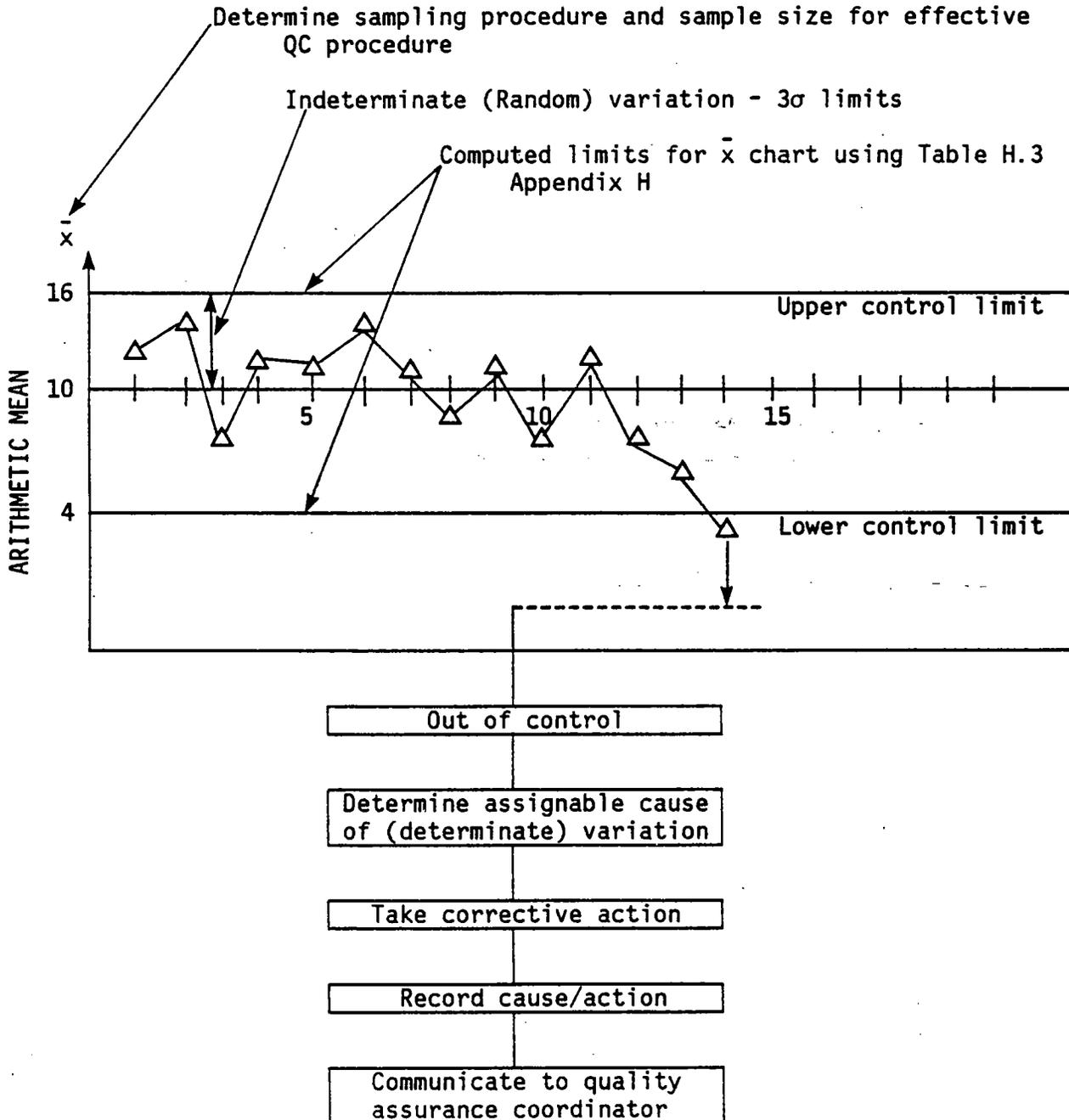


Figure 1.4.10.1. A typical control chart for sample averages and indicated actions.

#### 1.4.11 PROCUREMENT QUALITY CONTROL

##### 1.4.11.1 ABSTRACT

The quality of equipment and supplies used in a measurement process significantly affects the quality and the amount of data generated from the process. Quality control procedures for procurement should be used to ensure that the equipment (and supplies) will yield data consistent with the objectives of the measurement process.

The quality control procedures for the procurement of ambient air quality analyzer include:

1. Make prepurchase evaluation and selection
2. Contact users of the analyzers being evaluated
3. Prepare contract specifications
4. Conduct acceptance test
5. Compare to old analyzer if appropriate
6. Maintain records of equipment--performance specifications, acceptance test data, maintenance data, and vendor performance.

Similar quality control procedures for procurement of calibration standards, chemicals, and materials should be followed. These will be described in the following section.

##### 1.4.11.2 DISCUSSION

In this section, the quality control procedures are given for the procurement of research and monitoring services (and interagency agreements) or of equipment and supplies.

1.4.11.2.1 Procurement of Research and Monitoring Services - EPA's policy requires all Regional Offices, Program Offices, and the States to participate in a centrally managed Agencywide QA Program. The goal of the Program is to ensure that all

environmentally-related measurements which are funded by EPA or which generate data mandated by EPA are scientifically valid, defensible, and of known precision and accuracy. In a memorandum dated June 14, 1979, the administrator specifically addressed the QA requirements for all EPA extramural projects, including contracts, interagency agreements, grants, and cooperative agreements, that involve environmental measurements.

This subsection is to provide assistance to EPA personnel involved in the administration of contracts and interagency agreements to uniformly implement the intent of the memo. Guidance and review forms are provided in References 1 and 2 to assist in the systematic review of QA requirements for projects covered by contracts and interagency agreements. For contracts, the steps of the review include:

1. Requirements of the RFP (Request for Proposal),
2. Evaluation of the offerors' proposals, and
3. Requirements of the contract.

The RFP shall state that a QA Program Plan and/or Project Plan must be submitted as a separate identifiable part of the technical proposal. In addition, the RFP should include requirements that offerors in the competitive range participate in appropriate EPA, QA performance audits and that they permit QA system audits by EPA as part of the evaluation process to determine the awardee.

Contracts should require submittal of a QA Program and/or QA Project Plan(s) to EPA for approval prior to the initiation of environmental measurements. The term "environmental measurements" applies to essentially all field and laboratory investigations that generate data such as measuring chemical, physical, or biological parameters in the environment; determining presence or absence of pollutants in waste streams; health and ecological effects; clinical and epidemiological investigations; engineering and process evaluations; studies involving laboratory simulation of environmental events; and studies on pollutant transport and dispersion modeling.

The Office of Research and Development (ORD) is delegated the responsibility to develop, implement, and manage the QA Program for the Agency. For Level-of-Effort type contracts, separate QA Project Plans should be required for each work assignment involving environmental measurements. The contract should also include requirements for the submission of separate periodic QA reports, the right of EPA to conduct QA system audits, and, as appropriate, requirements for participation in EPA performance audits. The requirements shall extend through the awardee to all subcontractors.

The QA review for interagency agreement shall be similar to those for awarded contracts.

1.4.11.2.2 Procurement of Analyzer<sup>3</sup> - The following QC procedures for the procurement of an analyzer should be considered. Similar procedures should be followed for other equipment (recorders, data loggers, flow meters, etc.).

1. Make prepurchase evaluation and selection - This includes defining the performance specifications and indicating the relative importance of these requirements. The advantages and disadvantages of each type of analyzer should be determined from information provided by the manufacturers' operating manuals.

2. Contact users of the analyzers being evaluated - A user's list should be requested from the manufacturers/supplier. These users should be contacted with regard to the performance, dependability, ease of operation, and other pertinent factors relative to the analyzer. Analyzers can then be selected for in-house testing and comparison. The analyzers which are found to yield satisfactory performance should then be field tested and a specific analyzer is then selected based on a review of all of the evaluation data.

3. Prepare purchase specifications - The performance specifications should be written into the purchase contract. It should:

- a. require manufacturer test data documenting that the specific analyzer purchased meet the specifications,
- b. specify that payment is not due until the analyzer has passed an acceptance test,
- c. include a warranty covering a free repair period of at least one year,
- d. specify that operating manuals be supplied and that they be consistent with the analyzer purchased,
- e. include operator training, and
- f. require that a two-year supply of consumable and spare parts be furnished.

4. Conduct acceptance test - Upon receipt of the analyzer, be sure that it meets the performance specifications. This includes an inspection that all components and optional equipment are present and the tests to evaluate critical performance parameters. In addition the analyzer should be tested simultaneously with an analyzer that is onsite (if applicable) to determine if the new analyzer yields comparable or improved quality data.

5. Maintain records - A record should be maintained of the analyzer and should include:

- a. Performance specs
- b. Acceptance test data
- c. Maintenance operations
- d. Vendor performance.

1.4.11.2.2 Calibration standards - Purchase contracts should require that (1) the standards be traceable to an NBS-SRM or commercial Certified Reference Materials (CRM) (Section 1.4.12), (2) vendor supply traceability protocol test data, (3) calibration curves (or formulas) for determining permeation rates be supplied with permeation tubes, and (4) detailed instructions for use and care of calibration standards be supplied.

1.4.11.2.3 Chemicals - Purchase contracts should require certified analyses of critical chemicals. Upon receipt, check the

chemicals against those on hand to ensure equivalency, overlap the "new" and "old" chemicals to ensure equivalency, maintain a record to aid in tracing problems to supplier/container.

1.4.11.2.4 Materials - Critical performance parameters should be specified in the contract. Upon receipt be sure the materials meet the requirements and overlap the use of the "new" and "old" materials to ensure equivalency.

#### 1.4.11.3 REFERENCES

1. Guidelines and Specifications for Implementing Quality Assurance Requirements for EPA Contracts and Interagency Agreements Involving Environmental Measurements, QAMS-002/80, ORD, USEPA, May 19, 1980.
2. Quality Assurance (QA) Requirements for Contracts in Excess of \$10,000, Procurement Information Notice 82-26, USEPA, February 23, 1982.
3. Kopecky, Mary Jo and B. Rodger, "Quality Assurance for Procurement of Air Analyzers," ASQC Technical Conference Transactions, 1979.



#### 1.4.12 CALIBRATION

This section contains a brief discussion of the major elements of a calibration program. Appendix J contains a discussion of the statistical aspects of calibration.

##### 1.4.12.1 ABSTRACT

Calibration is the single most important operation in the measurement process. Calibration is the process of establishing the relationship between the output of a measurement process and a known input. A calibration plan should be developed and implemented for all data measurement equipment, test equipment, and calibration standards to include:

1. A statement of the maximum allowable time between multipoint calibrations and calibration checks.
2. A statement of the minimum quality of calibration standards (e.g., standards should have four to ten times the accuracy of the instruments that they are being used to calibrate). A list of calibration standards should be provided.
3. Provisions for standards traceability (e.g., standards should be traced to NBS-SRM's or commercial Certified Reference Materials (CRM) if available).
4. Provisions for written procedures to aid in ensuring that calibrations are always performed in the same manner. The procedures should include the intended range of validity.
5. Statement of proper environmental conditions to ensure that the equipment is not significantly affected by its surroundings.
6. Provisions for proper record keeping and record forms to ensure that adequate documentation of calibrations is available for use in internal data validation and in case the data are used in enforcement actions.

7. Documentation on qualifications and training of personnel performing calibrations.

#### 1.4.12.2 DISCUSSION

A calibration plan should be developed and implemented for measuring and test equipment and calibration standards to include the items listed under the previous subsection. These items are described below:

1. Maximum allowable time between multipoint calibrations and zero/span checks - Calibration Intervals<sup>1,2</sup> - All calibration standards and measuring and test equipment should be assigned a maximum allowable time interval between multipoint calibrations and zero/span calibration checks (Figure 1.4.12.1). In the absence of an established calibration interval (based on equipment manufacturer's recommendation, Government specifications, etc.) for a particular item, an initial calibration interval should be assigned by the laboratory/quality assurance coordinators. The calibrations should be specified in terms of time or, in the case of certain types of test and measuring equipment, a usage period.

The establishment of the calibration intervals should be based on inherent stability, purpose or use, precision, bias, and degree of usage of the equipment. The time intervals may be shortened or lengthened (but not to exceed specifications/regulations) by evaluating the results of the previous and present calibrations and adjusting the schedule to reflect the findings. These evaluations must provide positive assurance that the adjustments will not adversely affect the accuracy of the system. The laboratory should maintain proper usage data and historical records for all equipment, to determine whether an adjustment of the calibration interval is warranted.

Adherence to the calibration frequency schedule is mandatory. One means of maintaining the schedule is to prepare a Calibration Control Card (e.g., Figure 1.4.12.2) as a reminder



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec																			
<b>INSTRUMENT RECORD CARD</b>																														
Instr type _____						Ident number _____						Instr code number _____																		
Model number _____						Ser number _____																								
Date rec'd _____						Date insp _____						O.K.'d by _____																		
Checking interval _____						App'd by _____																								
Location _____						Rec'd by _____																								
Calibration responsibility _____																														
Calibration instruction number _____																														

FRONT

Date checked	Checked by	Results of check

REVERSE

Figure 1.4.12.2. Calibration control card.

and a means for documenting pertinent information. It may be necessary to calibrate between normal calibration dates if there is evidence of inaccuracy or damage.

2. Quality of calibration standards - Transfer standards should have four to ten times the accuracy of field and laboratory instruments and gauges. For example, if a thermometer used in the field to determine air temperature has a specified accuracy (precision and bias) of  $\pm 2^{\circ}\text{F}$ , it should be calibrated against a laboratory thermometer with an accuracy of at least  $\pm 0.5^{\circ}\text{F}$ . The calibration standards used in the measurement system should be calibrated against higher-level, primary standards having unquestionable and higher accuracy. The highest-level standards available are National Bureau of Standards (NBS) Standard Reference Materials (SRM). A listing of environmental SRM available from the NBS is shown in Figure 1.4.12.3. These environmental SRM's may be purchased from the National Bureau of Standards, Office of Standard Reference Materials, Washington, D. C. 20234.

Calibration gases purchased from commercial vendors normally contain a certificate of analysis. Whenever an SRM gas is available from the NBS, commercial gas vendors should be requested to establish traceability of the certificate of analysis to this SRM gas (Section 2.0.7 of Volume II of this Handbook). Another standard of high accuracy has been recognized by EPA as equivalent to an NBS-SRM. These standards are commercial Certified Reference Materials (CRM). In the current EPA regulations where traceability of gas working standards (used for calibration and auditing) are required, this traceability may be established to either an NBS-SRM or a commercial CRM. A CRM is prepared by a commercial vendor according to a CRM procedure<sup>3</sup> developed by NBS and EPA. In brief, the CRM procedure requires the gas vendor to prepare a batch of 10 or more standards with the batch average concentration within 1.0 percent of the SRM it is duplicating. The gas vendor must conduct analyses to demonstrate the batch is both homogenous and stable. After the gas

Analyzed Gases

SRM	Type	Certified component	Nominal concentration	
1658a	Methane in air	CH <sub>4</sub>	1	ppm
1659a	Methane in air	CH <sub>4</sub>	10	ppm
1660a	Methane-propane in air	CH <sub>4</sub> /C <sub>3</sub> H <sub>8</sub>	4/1	ppm
1661	Sulfur dioxide in nitrogen	SO <sub>2</sub>	500	ppm
1662a	Sulfur dioxide in nitrogen	SO <sub>2</sub>	1000	ppm
1663a	Sulfur dioxide in nitrogen	SO <sub>2</sub>	1500	ppm
1664	Sulfur dioxide in nitrogen	SO <sub>2</sub>	2500	ppm
1665b	Propane in air	C <sub>3</sub> H <sub>8</sub>	3	ppm
1666b	Propane in air	C <sub>3</sub> H <sub>8</sub>	10	ppm
1667b	Propane in air	C <sub>3</sub> H <sub>8</sub>	50	ppm
1668b	Propane in air	C <sub>3</sub> H <sub>8</sub>	100	ppm
1669b	Propane in air	C <sub>3</sub> H <sub>8</sub>	500	ppm
1670	Carbon dioxide in air	CO <sub>2</sub>	0.033	percent
1671	Carbon dioxide in air	CO <sub>2</sub>	0.034	percent
1672	Carbon dioxide in air	CO <sub>2</sub>	0.035	percent
1674b	Carbon dioxide in nitrogen	CO <sub>2</sub>	7.0	percent
1675b	Carbon dioxide in nitrogen	CO <sub>2</sub>	14.0	percent
1677c	Carbon monoxide in nitrogen	CO	10	ppm
1678c	Carbon monoxide in nitrogen	CO	50	ppm
1679c	Carbon monoxide in nitrogen	CO	100	ppm
1680b	Carbon monoxide in nitrogen	CO	500	ppm
1681b	Carbon monoxide in nitrogen	CO	1000	ppm
1683b	Nitric oxide in nitrogen	NO	50	ppm
1684b	Nitric oxide in nitrogen	NO	100	ppm
1685b	Nitric oxide in nitrogen	NO	250	ppm
1686b	Nitric oxide in nitrogen	NO	500	ppm
1687b	Nitric oxide in nitrogen	NO	1000	ppm
1693	Sulfur dioxide in nitrogen	SO <sub>2</sub>	50	ppm
1694	Sulfur dioxide in nitrogen	SO <sub>2</sub>	90	ppm
1969	Sulfur dioxide in nitrogen	SO <sub>2</sub>	3500	ppm
1805	Benzene in nitrogen	C <sub>6</sub> H <sub>6</sub>	0.25	ppm
1806	Benzene in nitrogen	C <sub>6</sub> H <sub>6</sub>	9.5	ppm
1808	Perchloroethylene in nitrogen	C <sub>2</sub> Cl <sub>4</sub>	0.25	ppm
1809	Perchloroethylene in nitrogen	C <sub>2</sub> Cl <sub>4</sub>	10	ppm

Figure 1.4.12.3. Environmental standard reference materials available from the National Bureau of Standards in 1983. (Continued)

SRM	Type	Certified component	Nominal concentration	
1811	Benzene, toluene, bromobenzene, chlorobenzene in nitrogen	$C_6H_6/C_6H_5CH_3$ $C_6H_5Br/C_6H_5Cl$	0.25 ppm each component	
1812	Benzene, toluene, bromobenzene, chlorobenzene in nitrogen	$C_6H_6/C_6H_5CH_3$ $C_6H_5Br/C_6H_5Cl$	9.5 ppm each component	
2605	$N_2O$ & $CO_2$ in air (size - 150 ft <sup>3</sup> )	$N_2O$	270	ppb
2606	$N_2O$ & $CO_2$ in air (size - 30 ft <sup>3</sup> )	$CO_2$	305	ppm
2607	$N_2O$ & $CO_2$ in air (size - 150 ft <sup>3</sup> )	$N_2O$	270	ppb
2608	$N_2O$ & $CO_2$ in air (size - 30 ft <sup>3</sup> )	$CO_2$	305	ppm
2609	$N_2O$ & $CO_2$ in air (size - 150 ft <sup>3</sup> )	$N_2O$	300	ppb
2610	$N_2O$ & $CO_2$ in air (size - 30 ft <sup>3</sup> )	$CO_2$	340	ppm
2612a	Carbon monoxide in air	$N_2O$	300	ppb
2613a	Carbon monoxide in air	$CO_2$	340	ppm
2614a	Carbon monoxide in air	$N_2O$	330	ppb
2619a	Carbon dioxide in nitrogen	$CO_2$	375	ppm
2620a	Carbon dioxide in nitrogen	$N_2O$	330	ppb
2621a	Carbon dioxide in nitrogen	$CO_2$	375	ppm
2622a	Carbon dioxide in nitrogen	$CO$	9.5	ppm
2623	Carbon dioxide in nitrogen	$CO$	18	ppm
2624a	Carbon dioxide in nitrogen	$CO$	43	ppm
2625	Carbon dioxide in nitrogen	$CO_2$	0.5	percent
2626a	Carbon dioxide in nitrogen	$CO_2$	1.0	percent
2627	Nitric oxide in nitrogen	$CO_2$	1.5	percent
2628	Nitric oxide in nitrogen	$CO_2$	2.0	percent
2629	Nitric oxide in nitrogen	$CO_2$	2.5	percent
2630	Nitric oxide in nitrogen	$CO_2$	3.0	percent
2631	Nitric oxide in nitrogen	$CO_2$	3.5	percent
2632	Carbon dioxide in nitrogen	$CO_2$	4.0	percent
2633	Carbon dioxide in nitrogen	$NO$	5	ppm
2634	Carbon dioxide in nitrogen	$NO$	10	ppm
2635	Carbon monoxide in nitrogen	$NO$	20	ppm
2636	Carbon monoxide in nitrogen	$NO$	1500	ppm
2637	Carbon monoxide in nitrogen	$NO$	3000	ppm
2638	Carbon monoxide in nitrogen	$CO_2$	300	ppm
2639	Carbon monoxide in nitrogen	$CO_2$	400	ppm
2639	Carbon monoxide in nitrogen	$CO_2$	800	ppm
2635	Carbon monoxide in nitrogen	$CO$	25	ppm
2636	Carbon monoxide in nitrogen	$CO$	250	ppm
2637	Carbon monoxide in nitrogen	$CO$	2500	ppm
2638	Carbon monoxide in nitrogen	$CO$	5000	ppm
2639	Carbon monoxide in nitrogen	$CO$	1	percent

Figure 1.4.12.3 (continued)

SRM	Type	Certified component	Nominal concentration	
2640	Carbon monoxide in nitrogen	CO	2	percent
2641	Carbon monoxide in nitrogen	CO	4	percent
2642	Carbon monoxide in nitrogen	CO	8	percent
2643	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	100	ppm
2644	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	250	ppm
2645	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	500	ppm
2646	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	1000	ppm
2647	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	2500	ppm
2648	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	5000	ppm
2649	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	1	percent
2650	Propane in nitrogen	C <sub>3</sub> H <sub>8</sub>	2	percent
2651	Propane & oxygen in nitrogen	C <sub>3</sub> H <sub>8</sub> /O <sub>2</sub>	0.01/ 5.0 percent	
2652	Propane & oxygen in nitrogen	C <sub>3</sub> H <sub>8</sub> /O <sub>2</sub>	0.01/10.0 percent	
2653	Nitrogen dioxide in air	NO <sub>2</sub>	250	ppm
2654	Nitrogen dioxide in air	NO <sub>2</sub>	500	ppm
2655	Nitrogen dioxide in air	NO <sub>2</sub>	1000	ppm
2656	Nitrogen dioxide in air	NO <sub>2</sub>	2500	ppm
2657	Oxygen in nitrogen	O <sub>2</sub>	2	percent
2658	Oxygen in nitrogen	O <sub>2</sub>	10	percent
2659	Oxygen in nitrogen	O <sub>2</sub>	21	percent
<b>Analyzed Liquids and Solids</b>				
1619	Residual fuel oil	S	0.7	wt%
1620a	Residual fuel oil	S	5	wt%
1621b	Residual fuel oil	S	1	wt%
1622a	Residual fuel oil	S	2	wt%
1623a	Residual fuel oil	S	0.2	wt%
1624a	Distillate fuel oil	S	0.2	wt%
1630	Trace mercury in coal	Hg	0.13	µg/g
1632a	Trace elements in coal, bituminous	-	30 elements	
1633a	Trace elements in coal fly ash	-	34 elements	
1634a	Trace elements in fuel oil	-	trace elements	
1635	Trace elements in coal, subbituminous	-	24 elements	
1636a	Lead in reference fuel	Pb	0.03,0.05,0.07,2.0 g/gal	
1637a	Lead in reference fuel	Pb	0.03,0.05,0.07 g/gal	
1638a	Lead in reference fuel	Pb	2.0 g/gal	
1641a	Mercury in water (set 6)	Hg	1.1 µg/g	
1642b	Mercury in water (950 ml)	Hg	1.1 µg/g	
1643a	Trace elements in water	-	18 elements	
1644	Polynuclear aromatic hydrocarbon generator columns	-	7 PAH	

Figure 1.4.12.3 (continued)

SRM	Type	Certified component	Nominal concentration
1647	Priority pollutant polynuclear aromatic hydrocarbons (in acetonitrile)	-	PAH
1648	Urban particulate (2 g)	-	33 elements
1649	Urban dust/organic	-	organics
1581	PCB's in oils	-	PCB
<u>Permeation Tubes</u>			
1625	SO <sub>2</sub> tube (10 cm)	SO <sub>2</sub>	2.8 µg/min
1626	SO <sub>2</sub> tube (5 cm)	SO <sub>2</sub>	1.4 µg/min
1627	SO <sub>2</sub> tube (2 cm)	SO <sub>2</sub>	0.56 µg/min
1629	NO <sub>2</sub> device	NO <sub>2</sub>	(0.5 to 1.5 µg/min)
1911	Benzene	C <sub>6</sub> H <sub>6</sub>	0.35 µg/min

Note: For SRM 1629, the individual rates are between the limits shown.

Figure 1.4.12.3 (continued)

vendor's analyses are complete, all results are sent to NBS. At the same time, the gas vendor provides EPA a list of individual standard numbers but no analyses results. At random, EPA selects two standards from each batch and conducts an analysis audit on these standards. The EPA sends the audit results to NBS who then decides whether to approve the candidate CRM batch for sale based on the gas vendor's analyses (for batch homogeneous and stability) and the EPA audit results. A description of the EPA audit program plus data demonstrating long-term stability of CRM is available.<sup>4</sup> A list of currently available CRM's may be obtained by writing to the following address:

U.S. Environmental Protection Agency  
Environmental Monitoring Systems Laboratory  
Quality Assurance Division  
Research Triangle Park, North Carolina 27711  
Attention: List of Current CRM

Inaccurate concentrations of working standards will result in serious errors in reported measured pollutant concentrations. By way of example, EPA audited 13 gas vendors that sold SO<sub>2</sub>, NO,

and CO cylinder gas standards. During the audit, the gas standards were purchased by an anonymous third party for EPA. In the procurement package to each gas vendor, a certificate of analysis was required for each cylinder gas. Traceability to an NBS-SRM was not required in the procurement package. From the EPA audit report,<sup>5,6</sup> four of the gas vendors provided certified SO<sub>2</sub> cylinder gases (at 90 ppm) that were in error by more than 30 percent. One of the four was in error by more than 100 percent. These audit results illustrate the need to require traceability to SRM or CRM during the procurement of gas standards.

3. Standards traceability - Calibration Source - All calibrations performed by or for the laboratory should be traced through an unbroken chain (supported by reports or data forms) to some ultimate or national reference standards maintained by a national organization such as the NBS. The ultimate reference standard can also be an independent reproducible standard (i.e., a standard that depends on accepted values of natural physical constants). Traceability is needed because calibration gas users often receive inaccurate and/or unstable calibration gases.

An up-to-date report for each calibration standard used in the calibration system should be provided. If calibration services are performed by a commercial laboratory on a contract basis, copies of reports issued by them should be maintained on file.

All reports should be kept in a suitable file and should contain the following information:

- a. Report number.
- b. Identification or serial number of the calibration standard to which the report pertains.
- c. Conditions under which the calibration was performed (temperature, relative humidity, etc.).
- d. Accuracy of calibration standard (expressed in percentage or other suitable terms).

- e. Deviation or corrections.
- f. Corrections that must be applied if standard conditions of temperature, etc., are not met or differ from those at place of calibration.

Contracts for calibration services should require the commercial laboratory to supply records on traceability of their calibration standards.

4. Written calibration procedures - Written step-by-step procedures for calibration of measuring and test equipment and use of calibration standards should be provided by the laboratory in order to eliminate possible measurement inaccuracies due to, for example, differences in techniques, environmental conditions, choice of higher-level standards. These calibration procedures may be prepared by the laboratory, or the laboratory may use published standard practices or written instructions that accompany purchased equipment. These procedures should include the following information:

- a. The specific equipment or group of equipment to which the procedure is applicable. ("Like" equipment or equipment of the same type, having compatible calibration points, environmental conditions, and accuracy requirements, may be serviced by the same calibration procedure.)
- b. A brief description of the scope, principle, and/or theory of the calibration method.
- c. Fundamental calibration specifications, such as calibration points, environmental requirements, and accuracy requirements.
- d. A list of calibration standards and accessory equipment required to perform an effective calibration. Manufacturer's name, model number, and accuracy should be included as applicable.
- e. A complete procedure for calibration arranged in a step-by-step manner, clearly and concisely written.

f. Calibration procedures should provide specific instructions for obtaining and recording the test data, and should include data forms.

When available, published procedures may be used. NBS Handbook 77, Precision Measurement and Calibration,<sup>7</sup> published by the National Bureau of Standards, provides calibration procedures for many types of electrical, hydraulic, electronic, and mechanical measuring instruments.

Many calibration procedures require statistical analysis of results. A detailed example of computations for calibration of an NO<sub>2</sub> monitor is provided in Appendix J.

5. Environmental conditions for equipment - Measuring and test equipment and calibration standards should be calibrated in an area that provides control on environmental conditions to the degree necessary to assure required precision and bias. The calibration area should be reasonably free of dust, vapor, vibration, and radio frequency interferences; and it should not be close to equipment that produces noise, vibration, or chemical emissions.

The laboratory calibration area should have adequate temperature and humidity controls. A temperature of 68 to 73°F and a relative humidity of 35 to 50 percent usually provide a suitable environment.

A filtered air supply is desirable in the calibration area. Dust particles are more than just a nuisance; they can be abrasive, conductive, and damaging to instruments.

Other environmental conditions for consideration are:

a. Electric power. Recommended requirements for electrical power within the laboratory should include voltage regulation of at least 10 percent (preferably 5 percent); low values of harmonic distortion; minimum voltage fluctuations caused by interaction of other users on main line to laboratory (separate input power if possible); and a suitable grounding system established to assure equal potentials to ground throughout the laboratory.

b. Lighting. Adequate lighting (suggested values-- 80- to 100-foot candles) should be provided for workbench areas. The lighting may be provided by overhead incandescent or fluorescent lights. Fluorescent lights should be shielded properly to reduce electrical noise.

6. Record keeping - Proper and complete documentation of calibrations performed may be needed if monitoring data are used in an enforcement action and for internal data validation. Bound calibration logbooks should be used. Traceability should be supported by reports or data forms. Items that should be recorded for each instrument calibration include:

- a. Description of calibration material/device, including serial number(s),
- b. Description of instrument calibrated, including serial number(s),
- c. Instrument location,
- d. Date of calibration,
- e. Signature of person performing calibration, and
- f. Calibration data, including environmental conditions during calibration.

7. Qualifications and training of personnel - The personnel performing the calibrations must be adequately trained for the particular calibrations, in the record keeping, and in adherence to the calibration plan. On-the-job training must be monitored until the operator can perform accurate calibrations.

#### 1.4.12.3 REFERENCES

1. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS), Federal Register, Vol. 44, Number 92, May 10, 1979.
2. Appendix B - Quality Assurance Requirements for Prevention of Significant Deterioration (PSD) Air Monitoring, Federal Register, Vol. 44, Number 92, May 10, 1979.
3. A Procedure for Establishing Traceability of Gas Mixtures to Certain National Bureau of Standards Standard Reference

Materials. EPA-600/7-81-010, Joint publication by NBS and EPA. Available from the U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Quality Assurance Division, Research Triangle Park, North Carolina 27711, May 1981.

4. Wright, R. S., Eaton, W. C., Decker, C. E., and von Lehmden, D. J., The Performance Audit Program for Gaseous Certified Reference Materials, for presentation at APCA Annual Meeting, June 1983. Available from the U.S. Environmental Protection Agency at address shown in Reference 3.
5. Decker, C. E., Eaton, W. C., Shores, R. C., and Wall, C. V., Analysis of Commercial Cylinder Gases of Nitric Oxide, Sulfur Dioxide, and Carbon Monoxide at Source Concentrations--Results of Audit 5, EPA-600/S4-81-080, December 1981.
6. Decker, C. E., Saeger, M. L., Eaton, W. C. and von Lehmden, D. J., Analysis of Commercial Gases of Nitric Oxide, Sulfur Dioxide, and Carbon Monoxide at Source Concentrations, Proceedings of APCA Specialty Conference on Continuous Emission Monitoring: Design, Operation and Experience, pp. 197-209, 1981.
7. Precision Measurement and Calibration. National Bureau of Standards, Washington, D.C. NBS Handbook 77.

#### BIBLIOGRAPHY

1. Evaluation of Contractor's Calibration System. "MIL-HDBK-52, Department of Defense, Washington, D. C. July 1964.
2. Beckwith, T. G., and Buck, N. L. Mechanical Measurements. Addison-Wesly Publishing Company, Reading, Massachusetts. 1969. Chapter 10.
3. Covino, C. P., and Meghri, A. W. Quality Assurance Manual. Industrial Press, Inc., New York. 1967.
4. Feigenbaum, A. V. Total Quality Control. McGraw-Hill Book Company, New York. 1961. pp. 512-514.
5. Kennedy, C. W., and Andrews, D. E. Inspection and Gauging. 4th ed. Industrial Press, Inc., New York. 1967. Chapter 14.
6. Calibration System Requirements. MIL-C-45662A, Department of Defense, Washington, D. C. February 1962.

#### 1.4.13 CORRECTIVE ACTION

##### 1.4.13.1 ABSTRACT

1. Corrective actions are of two kinds:
  - a. Corrective action - on-the-spot or immediate, to correct nonconforming data or repair equipment.
  - b. Corrective action - long-term, to eliminate causes of nonconformance.
2. Steps comprising a closed-loop corrective action system are:
  - a. Define the problem.
  - b. Assign responsibility for investigating the problem.
  - c. Investigate and determine the cause of the problem.
  - d. Determine a corrective action to eliminate the problem.
  - e. Assign and accept responsibility for implementing the corrective action.
  - f. Establish effectiveness of the corrective action and implement the correction.
  - g. Verify that the corrective action has eliminated the problem.

Corrective action procedures recognize the need for an assigned individual to test the effectiveness of the system and the corrective actions.

##### 1.4.13.2 DISCUSSION

On-the-spot or immediate corrective action - This is the process of correcting malfunctioning equipment.

In a quality assurance program, one of the most effective means of preventing trouble is to respond immediately to reports from the operator of suspicious data or equipment malfunctions. Application of proper corrective actions at this point can reduce or prevent the collection of poor quality data. Established procedures for corrective actions are available in the methods if the performance limits are exceeded (either through direct observation of the parameter or through review of control charts). Specific control procedures, calibration, presampling or preanalysis operational checks, are designed to detect instances in which corrective action is necessary. A checklist for logical alternatives for tracing the source of a sampling or analytical error is provided to the operator. Trouble-shooting guides for operators (field technicians or lab analysts) are generally found in instrument manufacturer's manuals. On-the-spot corrective actions routinely made by field technicians or lab analysts should be documented as normal operating procedures, and no specific documentation other than notations in operations logbooks need be made.

Long-term corrective action - The purpose of long-term corrective action is to identify and eliminate causes of nonconformance; hopefully they will be eliminated permanently. To improve data quality to an acceptable level and to maintain data quality at an acceptable level, it is necessary that the quality assurance system be sensitive and timely in detecting out-of-control or unsatisfactory conditions. It is equally important that, once the conditions of unacceptable quality data are indicated, a systematic and timely mechanism is established to assure that the condition is reported to those who can correct it and that a positive loop mechanism is established to assure that appropriate corrective action has been taken. For major problems it is desirable that a formal system of reporting and recording of corrective actions be established.

Closed-loop corrective action system for major problems - Experience has shown that most problems will not disappear until positive action has been taken by management. The significant characteristic of any good management system is the step that closes the loop--the determination to make a change if the system demands it.

The following discussion outlines the considerations and procedures necessary to understand and implement an effective closed-loop corrective action system for major problems. Effective corrective action occurs when many individuals and departments cooperate in a well planned program. There are several essential steps that must be taken to plan and implement a corrective action program that achieves significant results.

Corrective actions should be a continual part of the laboratory system for quality, and they should be formally documented. Corrective action is not complete until it is demonstrated that the action has effectively and permanently corrected the problem. Diligent follow-up is probably the most important requirement of a successful corrective action system.

Initiation, use, and completion of the corrective action request - A corrective action request may be initiated by any individual who observes a major problem. The corrective action request should be documented and limited to a single problem. If more than one problem is involved, each should be documented on a separate form.

Use of a Master Log - Corrective action can be casual when the organization is small or the problems few. When this is not the case and the problems are severe and numerous, action documentation and status records are required. All requests for corrective action, and action taken should be entered into a master log for control purposes and for visibility to management.



## 1.4.14 QUALITY COSTS<sup>1,2</sup>

### 1.4.14.1 ABSTRACT

Cost categories can be identified for a quality assurance system. By assigning costs according to quality assurance activities and grouping these by cost categories, cost effectiveness can be appraised.

1. Identification of costs is a prerequisite to cost reduction.

2. The American Society for Quality Control categorizes costs as: (a) prevention costs, (b) appraisal costs, (c) internal-failure costs, and (d) external-failure costs. For air pollution measurement systems, a more practical cost categorization is: (a) prevention cost, (b) appraisal costs, and (c) correction-failure costs. The quality assurance activities listed in this Handbook have been placed in these three cost categories. Since accounting systems are not set up to accommodate cost breakdown by quality assurance activities, judgment is required to apportion the costs into the correct cost category.

3. Quality control (QC) cost figures should be reported periodically (e.g., quarterly) to management.

4. Allocation of cost figures from the accounting system into the applicable cost categories helps to identify quality assurance activities whose costs may be disproportionate relative to the total cost. Furthermore, quality cost figures provide input for budget forecasting.

### 1.4.14.2 DISCUSSION

Program managers with Governmental agencies and industrial organizations involved in environmental measurement programs are concerned with overall program cost-effectiveness including total

cost, data quality and timeliness. There are several costing techniques designed to aid the manager in monitoring and controlling program costs. One particular technique specifically applicable to the operational phase of a program is quality cost system.

#### 1.4.14.2.1 Objective of a Quality Cost System

The objective of a quality cost system for an environmental monitoring program is to minimize the cost of those operational activities directed toward controlling or improving data quality while maintaining an acceptable level of data quality. The basic concept of the quality cost system is to minimize total quality costs through proper allocation of planned expenditures for the prevention and appraisal efforts in order to control the unplanned correction costs. That is, the system is predicated on the idea that prevention is cheaper than correction.

#### 1.4.14.2.2 Structuring of Quality Costs

The first step in developing a quality cost system is identifying the cost of quality-related activities, including all operational activities that affect data quality, and dividing them into the major cost categories.

Costs are divided into category, group, and activity. Category, the most general classification, refers to the standard cost subdivisions of prevention, appraisal, and failure (or correction). The category subdivision of cost provides the basic format of the quality cost system. Activity is the most specific classification and refers to the discrete operations for which costs should be determined. Similar types of activities are summarized in groups for purposes of discussion and ease in reporting.

1.4.14.2.2.1 Cost categories--The quality cost system structure provides a means for identification of quality-related activities and for organization of these activities into prevention, appraisal, and failure cost categories. These categories are defined as follows:

1. Prevention costs--Costs associated with planned activities whose purpose is to ensure the collection of data of acceptable quality and to prevent the generation of data of unacceptable quality.

2. Appraisal costs--Costs associated with measurement and evaluation of data quality. This includes the measurement and evaluation of materials, equipment, and processes used to obtain quality data.

3. Failure costs--Costs incurred directly by the monitoring agency or organization producing the failure (unacceptable data).

1.4.14.2.2.2 Cost Groups--Quality cost groups provide a means for subdividing the costs within each category into a small number of subcategories which eliminates the need for reporting quality costs on a specific activity basis. Although the groups listed below are common to all environmental measurement methods, the specific activities included in each group may differ between methods.

Groups within prevention costs. Prevention costs are subdivided into five groups:

1. Planning and documentation--Planning and documentation of procedures for all phases of the measurement process that may have an effect on data quality.

2. Procurement specification and acceptance--Testing of equipment parts, materials, and services necessary for system operation. This includes the initial on-site review and performance test, if any.

3. Training--Preparing or attending formal training programs, evaluation of training status of personnel, and informed on-the-job training.

4. Preventive maintenance--Equipment cleaning, lubrication, and parts replacement performed to prevent (rather than correct) failures.

5. System calibration--Calibration of the monitoring system, the frequency of which could be adjusted to improve the accuracy of the data being generated. This includes initial calibration and routine calibration checks and a protocol for tracing the calibration standards to primary standards.

Groups within appraisal costs. Appraisal costs are subdivided into four groups:

1. Quality control measures--QC-related checks to evaluate measurement equipment performance and procedures.

2. Audit measures--Audit of measurement system performance by persons outside the normal operating personnel.

3. Data validation--Tests performed on processed data to assess its correctness.

4. Quality assurance assessment and reporting--Review, assessment, and reporting of QA activities.

Groups within failure costs. Under most quality cost systems, the failure category is subdivided into internal and external failure costs. Internal failure costs are those costs incurred directly by the agency or organization producing the failure.

Internal failure costs are subdivided into three groups:

1. Problem investigation--Efforts to determine the cause of poor data quality.

2. Corrective action--Cost of efforts to correct the cause of poor data quality, implementing solutions, and measures to prevent problem reoccurrence.

3. Lost data--The cost of efforts expended for data which was either invalidated or not captured (unacquired and/or unacceptable data). This cost is usually prorated from the total operational budget of the monitoring organization for the percentage of data lost.

External failure costs are associated with the use of poor quality data external to the monitoring organization or agency.

collecting the data. In air monitoring work these costs are significant but are difficult to systematically quantize. Only failure costs internal to the monitoring agency are considered herein. However, external-failure costs are important and should be considered when making decisions on additional efforts necessary for increasing data quality or for the allocation of funds for resampling and/or reanalysis.

Examples of external-failure cost groups are:

1. Enforcement actions--Cost of attempted enforcement actions lost due to questionable monitoring data.
2. Industry--Expenditures by industry as a result of inappropriate or inadequate standards established with questionable data.
3. Historical data--Loss of data base used to determine trends and effectiveness of control measures.

1.4.14.2.2.3 Cost Activities--Examples of specific quality-related activities which affect data quality are presented in Table 1.4.14.1. These activities are provided as a guide for implementation of a quality cost system for an air quality program utilizing continuous monitors. Uniformity across agencies and organizations in the selection of activities is desirable and encouraged, however, there are variations which may exist, particularly between monitoring agencies and industrial/research projects.

Agencies should make an effort to maintain uniformity regarding the placement of activities in the appropriate cost group and cost category. This will provide a basis for future "between agency" comparison and evaluation of quality cost systems.

#### 1.4.14.2.3 Development and Implementation of the Quality Cost System

Guidelines are presented in this section for the development and implementation of a quality cost system. These cover planning the system, selecting applicable activities, identifying

TABLE 1.4.14.1. EXAMPLE OF COST ACTIVITIES FOR A STATE AGENCY

Cost category	Cost group	Activity
Prevention	Planning and documentation	QA program plan for air monitoring system
		Interlaboratory comparisons
	Procurement specifications	Inspection and acceptance testing of equipment and reference materials
	Training	On-the-job and formal training
	Preventive maintenance	Preventive maintenance program for analyzers and equipment
	System calibration	Zero and span precision checks
Appraisal	QC measures	Analysis of control samples
		Duplicate samples operation
	Audit measures	Participation in EPA audit performance survey
		System audits
Data validation	Strip chart checks	
	Statistical checks	
	Graphical review of data	
QA assessment and reporting	Assessment of audit and precision data	
	Report preparation	
Failure (correction)	Problem investigation	Special testing for investigation of problem areas
	Corrective action	Reanalysis of samples
	Lost data	Missing or unacceptable data

sources of quality cost data, tabulating, and reporting the cost data.

1.4.14.2.3.1 Implementation of a quality cost system--Implementation of a quality cost system need not be expensive and time consuming. It can be kept simple if existing data sources are used wherever possible. The importance of planning cannot be overemphasized. Supervisors should be thoroughly briefed on quality cost system concepts, benefits, and goals.

System planning should include the following items:

1. Determining scope of the initial quality cost program.
2. Setting objectives for the quality cost program.
3. Evaluating existing cost data.
4. Determining sources to be utilized for the cost data.
5. Deciding on the report formats, distribution, and schedule.

To gain experience with quality cost system techniques, an initial pilot program could be developed for a single measurement or project within the agency. The unit selected should be representative, (i.e., exhibit expenditure for each cost category: prevention, appraisal, and failure). Once a working system for the initial effort has been established, a full-scale quality cost system can then be implemented.

1.4.14.2.3.2 Activity selection--The first step for a given agency to implement a quality cost system is to prepare a detailed list of the quality-related activities most representative of the agencies monitoring operation and to assign these activities to the appropriate cost groups and cost categories.

The general definitions of the cost groups and cost categories, presented in the previous section, are applicable to any measurement system. Specific activities contributing to these cost groups and categories, however, may vary significantly between agencies, depending on the scope of the cost system, magnitude of the monitoring network, parameters measured, and duration

of the monitoring operation. The activities listed in Table 1.4.14.1 are provided as a guide only, and they are not considered to be inclusive of all quality-related activities. An agency may elect to add or delete certain activities from this list. It is important, however, for an agency to maintain uniformity regarding the cost groups and categories for the activities.

1.4.14.2.3.3 Quality cost data sources-- Most accounting records do not contain cost data detailed enough to be directly useful to the operating quality cost system. Some further calculation is usually necessary to determine actual costs which may be entered on the worksheets. The cost of a given activity is usually estimated by prorating the person's charge rate by the percentage of time spent on that activity. A slightly rougher estimate can be made by using average charge rates for each position instead of the actual rates.

Failure costs are more difficult to quantize than either prevention or appraisal costs. The internal failure cost of lost data (unaquired and/or unacceptable data), for example, must be estimated from the total budget.

1.4.14.2.3.4 Quality cost analysis techniques--Techniques for analyzing and evaluating cost data range from simple charts comparing the major cost categories to sophisticated mathematical models of the total program. Common techniques include trend analysis and Pareto analysis.

Trend analysis. Trend analysis compares present to past quality expenditures by category. A history of quality cost data, typically a minimum of 1-year, is required for trend evaluation. (An example is given in Table 1.4.14.2 and Figure 1.4.14.1.)

Cost categories are plotted within the time frame of the reporting period (usually quarterly). Costs are plotted either as total dollars (if the scope of the monitoring program is relatively constant) or as "normalized" dollars/data unit (if the

TABLE 1.4.14.2. TOTAL QUALITY COST SUMMARY  
(Combined network costs, 1978-79)

Cost group	2nd quarter	3rd quarter	4th quarter	1st quarter
<b>Prevention</b>				
Planning and documentation	---	---	---	179
Procurement	---	---	---	179
Training	---	---	---	459
Preventive maintenance	588	559	587	1,046
System calibration and operation	<u>1,254</u>	<u>1,317</u>	<u>1,386</u>	<u>1,713</u>
<b>Total prevention costs</b>	<b>1,842</b>	<b>1,876</b>	<b>1,973</b>	<b>3,576</b>
<b>Appraisal</b>				
QC measures	768	806	742	1,631
Audits	1,308	1,508	1,470	1,913
Data validation	1,468	1,668	1,868	1,887
QA assessment and reporting	<u>1,748</u>	<u>1,839</u>	<u>1,686</u>	<u>2,179</u>
<b>Total appraisal costs</b>	<b>5,292</b>	<b>5,821</b>	<b>5,766</b>	<b>7,610</b>
<b>Failure</b>				
Problem investigation	1,579	1,886	1,760	704
Corrective action	1,361	1,334	1,365	546
Lost data (unacquired data)	<u>12,430</u>	<u>13,893</u>	<u>13,162</u>	<u>9,506</u>
<b>Total failure costs</b>	<b>15,370</b>	<b>17,113</b>	<b>16,287</b>	<b>10,256</b>
<b>Total quality costs</b>	<b>22,504</b>	<b>24,810</b>	<b>24,026</b>	<b>21,442</b>
<b>Measurement bases</b>				
Total program cost per quarter	48,304			
Total data units per quarter	33,792			

80-43.3

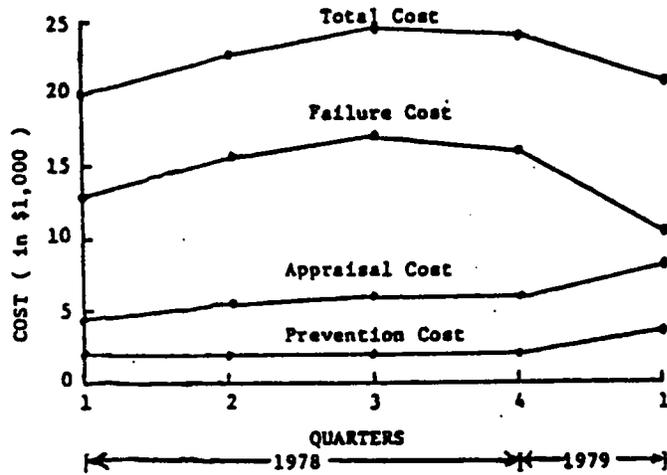


Figure 1.4.14.1. Quality cost trends.

scope may change). Groups and activities within the cost categories contributing the highest cost proportions are plotted separately (e.g., Figure 1.4.14.2).

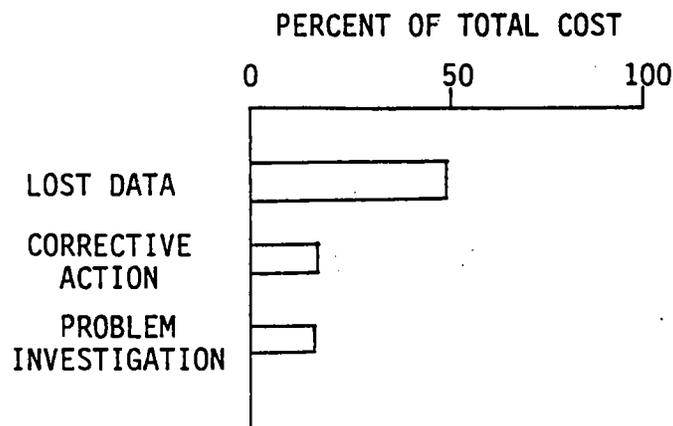


Figure 1.4.14.2. Failure cost distribution.

Pareto analysis. Pareto analysis identifies the areas with greatest potential for quality improvement by:

1. Listing factors and/or cost segments contributing to a problem area.
2. Ranking factors according to magnitude of their contribution.
3. Directing corrective action toward the largest contributor.

Pareto techniques may be used to analyze prevention, appraisal, or failure costs. They are not logically applied to the failure cost category, since the relative costs associated with activities in the failure category indicate the major source of data quality problems. Typically, relatively few contributors will account for most of the failure costs.<sup>3,4</sup> An example is given in Figure 1.4.14.2.

1.4.14.2.3.5 Quality cost reports--Quality cost reports prepared and distributed at regular intervals should be brief and factual, consisting primarily of a summary discussion, a tabulated data summary, and a graphic representation of cost category relationships, trends, and data analysis. The summary discussion should emphasize new or continuing problem areas and progress achieved during the reporting period.

Written reports should be directed toward specific levels of management. Managers and supervisors receiving reports should be thoroughly briefed on the concepts, purpose, and potential benefits of a quality cost system, that is, identification of quality-related problems, potential input into problem solution, and quality cost budgeting.

#### 1.4.14.3 REFERENCES

1. Rhodes, Raymond C. and Seymour Hochheiser. "Quality Costs for Environmental Monitoring Systems," American Society for Quality Control, Technical Conference, Vol. 77, 1971, p. 151.

2. Strong, R.B., J.H. White and F. Smith, "Guidelines for the Development and Implementation of a Quality Cost System for Air Pollution Measurement Programs," Research Triangle Institute, Research Triangle Park, North Carolina, 1980, EPA Contract No. 68-02-2722.
3. American Society for Quality Control, Quality Costs Technical Committee. "Guide for Reducing Quality Costs," Milwaukee, Wisconsin, 1977.
4. American Society of Quality Control, Quality Cost-Effectiveness Committee. "Quality Costs--What and How," Milwaukee, Wisconsin, 1977.

#### BIBLIOGRAPHY

1. Juran, J. M., and Gryna, F. M. Quality Planning and Analysis. McGraw-Hill, New York. 1970. Chapter 5, pp. 54-67.
2. "Quality Costs - What and How." American Society for Quality Control, Quality Cost Technical Committee. Milwaukee, Wisconsin. May 31, 1967.
3. Rhodes, R. C. "Implementing a Quality Cost System," Quality Progress. 5(2):16-19, February 1972.
4. Feigenbaum, A. V. Total Quality Control. McGraw-Hill, New York. 1961. Chapter 5, pp. 83-106.
5. Masser, W. J. "The Quality Manager and Quality Costs," Industrial Quality Control. XIV(4):5-8, October 1957.

## 1.4.15 INTERLABORATORY AND INTRALABORATORY TESTING

### 1.4.15.1 ABSTRACT

There are two major types of interlaboratory tests: (1) collaborative tests and (2) performance tests such as the EPA national performance audit program.<sup>1,2</sup> The collaborative test is a special form of an interlaboratory test and involves several laboratories for the purpose of defining the limits of a method.<sup>3</sup>

The interlaboratory performance tests such as the current EPA national performance audit program is used not only by EPA but other agencies (e.g., NIOSH). This test may involve over 100 participating laboratories and provides a means for participants to compare their results with those of other labs. This test allows the participants to take corrective action when their results are outside of specified limits stated for the audit materials.

Intralaboratory tests have as their purpose the identification of sources of measurement error and the estimation of bias and variability (repeatability and replicability) in the measurements resulting from these sources. The intralaboratory test of primary interest here is the ruggedness test. A ruggedness test is used for studying the effects on the measurement of several factors in the test procedure. The important factors or steps can be identified and limits determined for the test conditions in order that more precise and accurate data can be derived from the routine use of the measurement method.

## 1.4.15.2 DISCUSSION

1.4.15.2.1 Interlaboratory performance testing - The ultimate goal of interlaboratory testing is to improve the quality of data (both bias and precision) generated by all laboratories measuring the particular pollutant. The method of measurement is commonly not specified. However, the participant must report the method used. Because of its particular interest, the EPA national performance audit program is described briefly herein.

1. Audit materials are sent to participating laboratories.

2. These laboratories analyze the audit materials and send their results to EPA.

3. EPA compiles and analyzes the test results and reports their findings to the participants.

4. EPA prepares a summary report for all audits conducted during each year. This report summarizes all of the data but does not reveal individual lab results. For the annual audit report:

a. Results are analyzed at each concentration/flow level, usually 3 or 5 levels.

b. Results are examined and outliers are eliminated.

c. Averages and standard deviations are computed along with other pertinent statistics (e.g., relative standard deviation or coefficient of variation, mean value, accuracy and precision estimates). See Appendix K for an example of the reported results.

5. Performance audit schedules are announced in:

a. Journal of the Air Pollution Control Association

b. Stack Sampling News

c. Quality Assurance Newsletter

d. Additional information can be obtained from the Regional QA Coordinator or Environmental Monitoring Systems Laboratory, Quality Assurance Division, USEPA, Research Triangle Park, N. C. 27711.

6. An example of EPA national performance audit annual reports for source measurements and ambient measurements are shown in References 1 and 2, respectively.

1.4.15.2.2 Collaborative tests - A special form of interlaboratory tests is a collaborative test. In this type of test, several organizations participate simultaneously in the sampling and analysis of a test method in order to define the performance characteristics of the method, including precision and accuracy. Because of the high cost involved in collaborative testing, these tests are normally conducted only on methods that are or will be promulgated into EPA regulations as EPA test methods. A short discussion of this type of test is given in Appendix K. It is sufficient to indicate here that these tests use selected laboratories, and the test is usually performed over several days with all participants at the same location(s). The data analysis presents results on variation among and within labs, with the latter being subdivided into that among days and within days (or between replicates). Reports on collaborative tests of ambient air and source emission test methods are listed in Table 1.4.15.1.

1.4.15.2.3 Intralaboratory tests - One of the most frequently used intralaboratory test is the ruggedness test. In this test a single laboratory (and usually a single analyst) conducts the entire test. The purpose of the test is to check on the effects of perturbation of the test conditions on the results of the measurement method. Reports on ruggedness tests of ambient air and source emission test methods are listed in Table 1.4.15.1. The major steps in performing a ruggedness test are:

1. Select those conditions in the test method which may affect the variability of the test results.

TABLE 1.4.15.1. EPA METHOD EVALUATION AND COLLABORATIVE  
TEST REPORTS FOR AMBIENT AIR AND STATIONARY-  
SOURCE SPECIFIC METHODS

Ambient Air	
Report title	Reference number
<u>1983</u>	
Performance Test Results and Comparative Data for Designated Reference and Equivalent Methods for O <sub>3</sub>	EPA-600/S4-83-003 PB-83-166686
Performance Test Results and Comparative Data for Designated Reference Methods for CO	EPA-600/S4-83-013 PB-83-196808
Performance Test Results and Comparative Data for Designated Reference Methods fo NO <sub>2</sub>	EPA-600/S4-83-019 PB-83-200238
Technical Assistance Document for Sampling and Analysis of Toxic Organic Compounds in Ambient Air	EPA-600/4-83-027
<u>1982</u>	
A Comparative Evaluation of Seven Automatic Ambient Nonmethane Organic Compound Analyzers	EPA-600/S4-82-046
Laboratory Evaluation of Nonmethane Organic Carbon Determination in Ambient Air by Cryogenic Preconcentration and Flame Ionization Detection	EPA-600/S4-82-019
<u>1981</u>	
Technical Assistance Document for the Calibration and Operation of Automated Ambient Nonmethane Organic Compound Analyzers	EPA-600/4-81-015
<u>1980</u>	
Evaluation of Ozone Calibration Procedures	EPA-600/4-80-050
<u>1979</u>	
Improvement and Evaluation of Methods for Sulfate Analysis	EPA-600/4-79-028
(continued)	

Table 1.4.15.1 (continued)

Ambient Air	
Report title	Reference number
Transfer Standards for the Calibration of Ambient Air Monitoring Analyzers for Ozone - Technical Assistance Document	EPA-600/4-79-056
Technical Assistance Document for the Calibration of Ambient Ozone Monitors	EPA-600/4/79-057
<u>1978</u>	
Investigation of Flow Rate Calibration Procedure Associated with the High Volume Method for Determination of Suspended Particulates	EPA-600/4-78-047 PB-291386
Use of the Flame Photometric Detector Method for Measurement of Sulfur Dioxide in Ambient Air	EPA-600/4-78-024 PB-285171
<u>1977</u>	
Comparison of Wet Chemical and Instrumental Methods for Measuring Airborne Sulfate	EPA-600/7-77-128
Evaluation of 1 Percent Neutral Buffered Potassium Iodide Procedure for Calibration of Ozone Monitors, Environmental Monitoring Series	EPA-600/4-77-005
<u>1976</u>	
Measurement of Atmospheric Sulfates: Evaluation of the Methylthymol Blue Method	EPA-600/4-76-015 PB-253349/AS
Measurement of Atmospheric Sulfates: Literature Search and Methods Selection	EPA-600/4-76-008 PB-254387/AS
Effect of Temperature on Stability of Sulfur Dioxide Samples Collected by the Federal Reference Method	EPA-600/4-76-024
<u>1975</u>	
Technical Assistance Document for the Chemiluminescence Measurement of Nitrogen Dioxide	EPA-600/4-75-003
Evaluation of Effects of NO, CO, and Sampling Flow Rate on Arsenite Procedure for Measurement of NO <sub>2</sub> in Ambient Air	EPA-650/4-75-019 PB-242285/AS

(continued)

Table 1.4.15.1 (continued)

Ambient Air	
Report title	Reference number
Evaluation of Continuous Colorimetric Method for Measurement of Nitrogen Dioxide in Ambient Air	EPA-650/4-75-022 PB-243462/AS
Evaluation of Gas Phase Titration Technique as Used for Calibration of Nitrogen Dioxide Chemiluminescence Analyzers	EPA-650/4-75-021 PB-242294/AS
Summary Report: Workshop on Ozone Measurement by the Potassium Iodide Method	EPA-650/4-75-007 PB-240939/AS
Collaborative Study of Reference Method for Measurement of Photochemical Oxidants in the Atmosphere (Ozone-Ethylene Chemiluminescent Method)	EPA-650/4-75-016 PB-244105/AS
Collaborative Test of the Chemiluminescent Method for Measurement of NO <sub>2</sub> in Ambient Air	EPA-650/4-75-013 PB-246843/AS
Collaborative Test of the Continuous Colorimetric Method for Measurement of Nitrogen Dioxide in Ambient Air	EPA-650/4-75-011
<u>1974</u>	
An Evaluation of Arsenite Procedure for Determination of Nitrogen Dioxide in Ambient Air	EPA-650/4-74-048 PB-239727/AS
Collaborative Test of the TGS-ANSA Method for Measurement of Nitrogen Dioxide in Ambient Air	EPA-650/4-74-046 PB-257976/AS
An Evaluation of TGS-ANSA Procedure for Determination of Nitrogen Dioxide in Ambient Air	EPA-650/4-74-047 PB-238097
Evaluation of Triethanolamine Procedure for Determination of Nitrogen Dioxide in Ambient Air	EPA-650/4-74-031 PB-237348/AS
Collaborative Testing of Methods for Measurements of NO in Ambient Air. Volume I - Report of Testing (Sodium Arsenite Procedure)	EPA-650/4-019a PB-244902/AS
<u>1973</u>	
Collaborative Study of Reference Method for Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method)(24-Hour Sampling)	EPA-650/4-74-027 PB-239731/AS

(continued)

Table 1.4.15.1 (continued)

Ambient Air	
Report title	Reference number
<u>1972</u>	
Collaborative Study of Reference Method for the Continuous Measurement of Carbon Monoxide in the Atmosphere (Non-Dispersive Infrared Spectrometry)	EPA-72-009 PB-211265
<u>1971</u>	
Collaborative Study of Reference Method for Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method)	EPA-APTD-0903 PB-205891
Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High-Volume Method)	EPA-APTD-0904 PB-205892
<u>Publications</u>	
Performance Testing of Ambient Air Analyzers for SO <sub>2</sub>	American Laboratory, <u>12:19</u> , December 1980
Collaborative Testing of a Manual Sodium Arsenite Method for Measurement of Nitrogen Dioxide in Ambient Air	Environmental Science & Technology, 12:294, March 1978
Evaluation of the Sodium Arsenite Method for Measurement of NO <sub>2</sub> in Ambient Air	APCA Journal 27(6):553-556, June 1977
Performance of an NO <sub>2</sub> Permeation Device	Analytical Chemistry, <u>49:1823-1829</u> (1977)
Qualification of Ambient Methods as Reference Methods	American Society for Testing and Materials, Special Tech. Publication 598 (1976)

(continued)

Table 1.4.15.1 (continued)

<u>Stationary Sources</u>	
<u>Report title</u>	<u>Reference number</u>
<u>1983</u>	
Field Evaluation of EPA Reference Method 23	EPA-600/4-83-024 PB-83-214551
Technical Assistance Document: Quality Assurance Guideline for Visible Emission Training Programs	EPA-600/4-83-011 PB-83-193656
Assessment of the Adequacy of the Appendix F Quality Assurance Procedure for Maintaining CEMS Data Accuracy	EPA-600/4-83-047 PB-83-26440
Laboratory Evaluation of an Impinger Collection/Ion Chromatographic Source Test Method for Formaldehyde	EPA-600/4-83-031 PB-83-225326
Validation and Improvement of EPA Reference Method 25 - Determination of Gaseous Nonmethane Organic Emissions as Carbon	EPA-600/4-83-008 PB-83-191007
<u>1982</u>	
Evaluation of Method 16A - Determination of Total Reduced Sulfur Emissions from Stationary Sources	EPA-450/3-82-028
Reliability of CO and H <sub>2</sub> S Continuous Emission Moni- tors at a Petroleum Refinery	EPA-600/4-82-064
A Study to Evaluate and Improve EPA Reference Method 16	EPA-600/4-82-043 PB-83-165571
Techniques to Measure Volumetric Flow and Particulate Concentrations in Stacks with Cyclonic Flow	EPA-600/4-82-062
<u>1981</u>	
Method to Measure Polychlorinated Biphenyls in Natural Gas Pipelines	EPA-600/4-81-048
<u>1980</u>	
Evaluation of Emission Test Methods for Halogenated Hydrocarbons (Volume II)	EPA-600/4-80-003
An Evaluation Study of EPA Method 8	EPA-650/4-80-018

(continued)

Stationary Sources

Report title	Reference number
A Study to Improve EPA Methods 15 and 16 for Reduced Sulfur Compounds	EPA-600/4-80-023
Comparative Testing of EPA Methods 5 and 17 at Non-metallic Mineral Plants	EPA-600/4-80-022
<u>1979</u>	
Angular Flow Insensitive Pitot Tube Suitable for Use with Standard Stack Testing Equipment	EPA-600/4-79-042
Test Methods to Determine the Mercury Emissions from Sludge Incineration Plants	EPA-600/4-79-058
<u>1978</u>	
Collaborative Testing of EPA Method 106 (Vinyl Chloride) that will Provide for a Standardized Stationary Source Emission Measurement Method	EPA-600/4-78-058
<u>1977</u>	
Collaborative Study of EPA Method 13A and Method 13B	EPA-600/4-77-050 PB-278849/5BE
Survey of Continuous Source Emission Monitors: Survey No. 1 - NO <sub>x</sub> and SO <sub>2</sub>	EPA-600/4-77-022
Standardization of Method 11 at a Petroleum Refinery: Volume I	EPA-600/4-77-008a
Standardization of Method 11 at a Petroleum Refinery: Volume II	EPA-600/4-77-008b
Standardization of Stationary Source Method for Vinyl Chloride	EPA-600/4-77-026
<u>1976</u>	
Stationary Source Emission Test Methodology - A Review	EPA-600/4-76-044
The Application of EPA Method 6 to High Sulfur Dioxide Concentrations	EPA-600/4-76-038
(continued)	

Table 1.4.15.1 (continued)

Stationary Sources	
Report title	Reference number
Collaborative Study of Particulate Emissions Measurements by EPA Methods 2, 3, and 5 Using Paired Particulate Sampling Trains (Municipal Incinerators)	EPA-600/4-76-014 PB-252028/6
<u>1975</u>	
A Method to Obtain Replicate Particulate Samples from Stationary Sources	EPA-650/4-75-025 PB-245045/AS
Collaborative Study of Method 10 - Reference Method for Determination of Carbon Monoxide Emissions from Stationary Sources - Report of Testing	EPA-650/4-75-001 PB-241-284/AS
Evaluation and Collaborative Study of Method for Visual Determination of Opacity of Emissions from Stationary Sources	EPA-650/4-75-009
<u>1974</u>	
Collaborative Study of Method for the Determination of Sulfuric Acid Mist and Sulfur Dioxide Emissions from Stationary Sources	EPA-650/4-74-003 PB-240752/AS
Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators)	EPA-650/4-74-025 PB-238555/AS
Collaborative Study of Method for Stack Gas Analysis and Determination of Moisture Fraction with Use of Method 5	EPA-650/4-74-026
Collaborative Study of Method of Determination of Stack Gas Velocity and Volumetric Flow Rate in Conjunction with EPA Method 5	EPA-650/4-74-033 PB-241284/AS
Collaborative Study of Method 104 - Reference Method for Determination of Beryllium Emission from Stationary Sources	EPA-650/4-74-023 PB-245011/AS
Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Nitric Acid Plants)	EPA-650/4-74-028 PB-236930/AS
(continued)	

Table 1.4.15.1 (continued)

<u>Stationary Sources</u>	
Report title	Reference number
<u>1973</u>	
Collaborative Study of Method for the Determination of Sulfur Dioxide Emissions from Stationary Sources (Fossil-Fuel Fired Steam Generators)	EPA-650/4-74-024 PB-238293/AS
Laboratory and Field Evaluations of EPA Methods 2, 6, and 7	EPA-650/4-74-039 PB-238267/AS
Survey of Manual Methods of Measurements of Asbestos, Beryllium, Lead, Cadmium, Selenium, and Mercury in Stationary Source Emissions	EPA-650/4-74-015 PB-234326/AS
<u>Publications</u>	
Evaluation of Selected Gaseous Halocarbons for Use in Source Test Performance Audits	Journal of the Air Pollution Control Assoc. 33(9):823-826, September 1983
Analysis of Commercial Gases of Nitric Oxide, Sulfur Dioxide, and Carbon Monoxide at Source Concentrations	Proceedings of Journal of the Air Pollution Control Assoc. Specialty Conf. on Continuous Emission Monitoring: Design, Operation and Experience, pp. 197-209, 1981
The Collaborative Test of Method 6B: Twenty-Four-Hour Analysis of SO <sub>2</sub> and CO <sub>2</sub>	Journal of the Air Pollution Control Assoc. 33(10):968-973, October 1983
The Area Overlap Method for Determining Adequate Chromatographic Resolution	Journal of Chromatographic Science 20:221-114, May 1982

(continued)

Table 1.4.15.1 (continued)

Stationary Sources	
Report title	Reference number
A Device to Check Pitot Tube Accuracy	Journal of the Air Pollution Control Assoc. <u>31</u> (10):1092-1093, October 1981
Role of Quality Assurance in Collaborative Testing	Journal of the Air Pollution Control Assoc. <u>29</u> (7):708-709, July 1979
Measuring Inorganic and Alkyl Lead Emissions from Stationary Sources	Journal of the Air Pollution Control Assoc. <u>29</u> (9):959-962, September 1979
Precision Estimates for EPA Test Method 8 - SO <sub>2</sub> and H <sub>2</sub> SO <sub>4</sub> Emissions from Sulfuric Acid Plants	Atmospheric Environment <u>13</u> : 179-182 (1979)
Adequacy of Sampling Trains and Analytical Procedures Used for Fluoride	Atmospheric Environment <u>10</u> : 865-872, March 1976
Improved Procedure for Determining Mercury Emissions from Mercury Cell Chlor-Alkali Plants	Journal of the Air Pollution Control Assoc. <u>26</u> (7):674-677, July 1976
Means to Evaluate Performance of Stationary Source Test Methods	Environmental Science & Technology, <u>10</u> (6):85, January 1976
Field Reliability of the Orsat Analyzer	Journal of the Air Pollution Control Assoc. <u>26</u> (5):492-495, May 1976

PB reports are available from the National Technical Information Service, Department of Commerce, 5885 Port Royal Road, Springfield, Virginia 22161.

EPA reports are available from the U.S. Environmental Protection Agency, ORD Publications, 26 West St. Clair Street, Cincinnati, Ohio 45268

Internal reports are available from the Quality Assurance Division (MD-77), Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711.

2. Design an experiment to test for these conditions using methods of statistical design of experiments. For example, if there are seven factors or conditions to be varied, the experiment can be performed in eight complete analyses, provided the pattern of variation of the seven conditions follows the specified statistical plan.

3. Analyze the data to determine if any one or more of the seven factors has a significant effect on the results.

Other intralaboratory tests may be performed for the purpose of studying the effect of specific test conditions or operators. In fact the results of the ruggedness test may suggest further testing of one or two specific conditions. Another type of test may compare results from different analysts/instruments or from different measurement methods.

The major problems with designing a program to audit the analyst's proficiency are concerned with the following:

- a. What kinds of samples to use.
- b. How to prepare and introduce samples into the run without the analyst's knowledge.
- c. How often to check the analyst's proficiency.

The problems and suggested solutions or criteria for decision are given in Table 1.4.15.2.

TABLE 1.4.15.2. PROBLEMS IN ASSESSING ANALYST PROFICIENCY

Problem	Solutions and decision criteria
Kinds of samples	<ol style="list-style-type: none"> <li>1. Use replicate samples of unknowns or reference standards.</li> <li>2. Consider cost of samples.</li> <li>3. Samples must be exposed by the analyst to same preparatory steps as are normal unknown samples.</li> </ol>
Introducing the sample	<ol style="list-style-type: none"> <li>1. Samples should have same labels and appearance as unknowns.</li> <li>2. Because checking periods should not be obvious, supervisor and analyst should overlap the process of logging in samples.</li> <li>3. Supervisor can place knowns or replicates into the system occasionally.</li> <li>4. Save an aliquot from one day for analysis by another analyst. This technique can be used to detect bias.</li> </ol>
Frequency of checking performance	<ol style="list-style-type: none"> <li>1. Consider degree of automation.</li> <li>2. Consider total method precision.</li> <li>3. Consider analyst's training, attitude, and performance record.</li> </ol>

1.4.15.3 REFERENCES

1. Streib, E. W. and M. R. Midgett, A Summary of the 1982 EPA National Performance Audit Program on Source Measurements. EPA-600/4-83-049, December 1983.
2. Bennett, B. I., R. L. Lampe, L. F. Porter, A. P. Hines, and J. C. Puzak, Ambient Air Audits of Analytical Proficiency 1981, EPA-600/4-83-009, April 1983.
3. Youden, W. J. and E. H. Steiner. Statistical Manual of the Association of Analytical Chemists. Published by the Association of Official Analytical Chemists, P. O. Box 340, Benjamin Franklin Station, Washington, D. C. 20044. 1975.

BIBLIOGRAPHY

1. Pooler, F. The St. Louis Regional Air Pollution Study: A Coherent Effort Toward Improved Air Quality Simulation Models. Presented at the Summer Computer Simulation Conference, Houston, Texas. July 1974. Copies available from RAPS, EPA, Research Triangle Park, North Carolina 27711.
2. Bromberg, S. M., Akland, G. G., and Puzak, J. C. Survey of Laboratory Performance Analysis of Simulated Ambient SO<sub>2</sub> Bubbler Samples. Journal of the Air Pollution Control Association 24, 11. November 1974.
3. WHO International Air Pollution Monitoring Network--Data User's Guide. EP/72.6. June 1972. Available from Division of Environmental Health, WHO, 1211 Geneva 27, Switzerland.



## 1.4.16 AUDIT PROCEDURES

### 1.4.16.1 ABSTRACT

1. Performance audits are made to quantitatively evaluate the quality of data produced by the total measurement system (sample collection, sample analysis and data processing). The individuals performing the audit, their standards and equipment are different from the regular team (operating the measurement system) and their standards and equipment in order to obtain an independent assessment. The performance audit is commonly limited to a portion of the total measurement system (e.g., flow rate measurement, sample analysis) but may include the entire measurement system (e.g., continuous ambient air analyzer).

2. A system audit is a qualitative on-site inspection and review of the the total measurement system. The auditor should have extensive background experience with the measurement system being audited.

### 1.4.16.2 DISCUSSION

1.4.16.2.1 Performance Audits - The purposes of performance audits include:

1. Objective assessment of the accuracy of the data collected by a given measurement system,
2. Identification of sensors out-of-control,
3. Identification of systematic bias of a sensor or of the monitoring network,
4. Measurement of improvement in data quality based on data from previous and current audits.

The role of audits in the overall management program is verification. While audits do not improve data quality if all work is correctly performed, they do provide assurance that the work prescribed for the measurement program has been conducted

properly. Audits conducted by individuals not responsible for the day-to-day operations provide a control and assessment mechanism to program managers. A performance audit procedure for continuous ambient air analyzers is given herein to illustrate items that must be considered in conducting a performance audit.

1. Select audit materials

a. Use high concentration (10 to 100 ppm) audit cylinder gas in conjunction with a dilution system. Advantage--better gas stability at high concentration; disadvantage--dilution system calibration errors are possible.

b. Use low concentration (<1 ppm except for CO) audit cylinder gas. Advantage--no dilution system needed; disadvantages--probability of gas instability and thus inaccurate concentration, and number of cylinders.

c. Use permeation tubes. Advantage--better stability than low concentration cylinder gas; disadvantages--permeation rate, which is temperature dependent, must stabilize before performing audit and possibility of dilution system calibration error.

d. Use materials traceable to NBS-SRM or commercial CRM if possible.

e. Table 1.4.16.1 lists the primary standards applicable to ambient audit equipment calibration. The list is not all inclusive but includes the standards of high accuracy that will fulfill the traceability requirements.

2. Select audit concentration levels - As a minimum, use a low scale and a high scale point in order to check the analyzer's linearity, and use a third point near the sites' expected concentration level. Audit concentration levels are specified in 40 CFR Part 58, Appendices A and B for a minimum QA program.<sup>1,2</sup>

3. Determine auditor's proficiency - Auditor must analyze audit materials (including the verification of their stability) and his results compared with the known values prior to his performing an audit.

TABLE 1.4.16.1. PRIMARY STANDARDS

Parameter	Range	Usable standard	Primary standard
Flow rate	0-3 l/min	Soap bubble flow kit	NBS-traceable flow kit or gravimetrically calibrated flow tubes
Flow rate	0.5-3 l/min	1 l/revolution wet test meter 3 l/revolution wet test meter	Primary standard spirometer
Flow rate	0.1-2.5 m <sup>3</sup> /min	Positive displacement Roots meter	Roots meter
Time	0-5 minutes	Stopwatch	NBS-time
SO <sub>2</sub>	0-0.5 ppm	Permeation tube	NBS-SRM 1626
	50-90 ppm	Cylinder gas (SO <sub>2</sub> /N <sub>2</sub> )	NBS-SRM 1693, 1694 or commercial CRM
NO-NO <sub>2</sub> -NO <sub>x</sub>	0-0.5 ppm	NO <sub>2</sub> permeation tube	NBS-SRM 1629
	50 ppm	NO cylinder gas (NO/N <sub>2</sub> /GPT)	NBS-SRM 1683 or commercial CRM
O <sub>3</sub>	0-1.0 ppm	O <sub>3</sub> generator/UV photometer	Standard laboratory photometer
CO	10-100 ppm	Cylinder gas CO/N <sub>2</sub> or CO/air	NBS-SRM 1677, 1678, 1679, 2635, 2612, 2613, 2614 or commercial CRM

Note: Descriptions of NBS-SRM are shown in Figure 1.4.12.3. A list of currently available CRM may be obtained from EPA at address shown in Section 1.4.12.

4. Select analyzers out-of-control limits - Select the maximum allowable difference between the analyzer and auditor results. For gaseous analyzers, limits of 10 to 20% are commonly used.

5. Conduct the audit in the field

a. Record site data (address, operating organization, type of analyzer being audited, zero and span post settings, type of in-station calibration used, and general operating procedures.

b. Mark the data recording, indentifying the time interval in which the audit was performed. A data stamp may be used to document the station data system. This will ensure that recorder traces cannot be switched in future reference.

c. Have the station operator make necessary notations on the data acquisition system prior to disconnecting a monitor or sampler from the normal sampling mode. Initiate the audit. Audit techniques are listed in Table 1.4.16.2.

d. Have the station operator convert all station data to engineering units (ppm, m<sup>3</sup>/min, etc.) in the same manner that actual data are handled.

e. All pertinent data should be recorded in an orderly fashion on field data forms.

f. Return all equipment to normal sampling mode upon completion of the audit, so that no data are lost.

g. Make data computations and comparisons prior to vacating the test site. This is to ensure that no extraneous or inconsistent differences exist that are found after vacating the test site. It is often impossible to rectify a difference after leaving the test site. Hence calculations and comparisons made in the field are cost effective. Verbally relate as much information as possible to the analyzer operator immediately after the audit.

6. Verify the audit material stability after the audit (e.g., reanalysis of audit material).

TABLE 1.4.16.2. AUDIT TECHNIQUES

Pollutant/ parameter	Audit technique	Audit standard	Traceability to primary standard
SO <sub>2</sub>	Dynamic dilution of a stock cylinder	50 ppm SO <sub>2</sub> in air or N <sub>2</sub>	NBS-SRM 50 ppm SO <sub>2</sub> /N <sub>2</sub> standard  <u>or</u>
SO <sub>2</sub>	Dynamic dilution of a permeation tube	Permeation tube	NBS-SRM permea- tion tube
CO	Dynamic dilution of a stock cylinder	900 ppm CO in air or N <sub>2</sub>	NBS-SRM 1000 ppm CO/N <sub>2</sub> standard
CO	Separate cylinders	5, 20, 45 ppm CO in air or N <sub>2</sub> cylinders	NBS-SRM 50 ppm CO/N <sub>2</sub> standard
NO-NO <sub>x</sub> -NO <sub>2</sub>	Dynamic dilution/gas phase titration	50 ppm NO/N <sub>2</sub> with 0.5 ppm NO <sub>2</sub> impurity	NBS-SRM 50 ppm NO/N <sub>2</sub>
NO-NO <sub>x</sub> -NO <sub>2</sub>	Dynamic dilution of stock cylin- der/dynamic permeation dilution	50 ppm NO/N <sub>2</sub> cylinder; NO <sub>2</sub> permeation tube	NBS-SRM 50 ppm NO/N <sub>2</sub> cylinder; NBS NO <sub>2</sub> permea- tion tube
O <sub>3</sub>	O <sub>3</sub> generation with verifica- tion by UV photometry	Standard photometer	Standard labora- tory maintained UV photometer
TSP flow rate	Simultaneous flow rate comparison	ReF device	Primary standard Roots Meter system

7. Prepare Audit Report - Prepare a written report and mail to the pertinent personnel, it should include:

- a. Assessment of the accuracy of the data collected by the audited measurement system
- b. Identification of sensors out-of-control
- c. Identification of monitoring network bias
- d. Measurement of improvement in data quality since the previous audit(s).

8. Corrective Action - Determine if corrective actions are implemented.

Detailed guidance to State and local agencies on how to conduct performance audits of ambient air measurement systems are described in Section 2.0.12 of Volume II of this Handbook.

System Audit - Detailed guidance to State and local agencies for conducting a system audit of an ambient air monitoring program are in Section 2.0.11 of Volume II of this Handbook. Data forms are provided as an aid to the auditor. These forms should be submitted to the agency being evaluated 4 to 6 weeks prior to the on-site system audit. This allows the agency to locate and enter detailed information (often not immediately accessible) required by the forms. When the completed forms are returned, they should be reviewed and the auditor should prepare a list of specific questions he would like to discuss with the agency. An entrance interview date should be arranged to discuss these questions.

The next step is the systems audit. A convenient method is to trace the ambient data from the field measurement through the submittal to EPA, noting each step in the process, documenting the written procedures that are available and followed, and noting the calibration and quality control standards that are used.

After the auditor collects the information, an exit interview is conducted to explain the findings of the evaluation to

the agency representatives. A written report is then prepared as soon as possible to summarize the results of the audit.

Guidance on how to evaluate the capabilities of a source emission test team are described in Reference 3. Data forms are included as an aid to the auditor.

#### 1.4.16.3 REFERENCES

1. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS), Federal Register, Vol. 44, Number 92, May 19, p. 27574-27582.
2. Appendix B - Quality Assurance Requirements for Prevention of Significant Deterioration (PSD) Air Monitoring, Federal Register, Vol. 44, Number 92, May 1979, p. 27582-27584.
3. Estes, E. D. and Mitchell, W. J., Technical Assistance Document: Techniques to Determine A Company's Ability to Conduct A Quality Stack Test, EPA-600/4-82-018, March 1982.



## 1.4.17 DATA VALIDATION

### 1.4.17.1 ABSTRACT

Data validation can be accomplished by several methods. Validation can be manual or computerized.

1. Data validation is the process whereby data are filtered and either accepted or flagged for further investigation based on a set of criteria. Validation is performed to isolate spurious values since values are not automatically rejected. Records of invalid data should be maintained.

2. Validation methods can include review by supervisory personnel as well as application of validation criteria by computer. Criteria depend on the types of data and on the purpose of the measurement.

3. A number of statistical techniques are useful.<sup>1,2</sup> Periodic checking of manually reduced data values is important. Important statistical techniques are:

- a. Tests for outliers
- b. Gross limit tests<sup>2</sup>
- c. Parameter relationship tests<sup>2</sup>
- d. Inter- and intra-site correlations<sup>1</sup>
- e. Gap test.<sup>2</sup>

4. Data validation procedures and specific criteria for the EPA National Aerometric Data Bank (NADB) are given in order to illustrate important areas of concern which should be considered. These areas include:

- a. Screening data for representativeness; instrument averaging time; sampling program duration; and comparability with other reported data.

b. Providing criteria for: completeness of data; use of accuracy and precision data; handling of data reported as below the minimum detectable limits; and handling of data reported with negative values.

#### 1.4.17.2 DISCUSSION

Several data validation procedures are described briefly. They are presented in increasing order of analytical complexity and in four categories of use: tests for routine validation, for internal consistency, for historical or temporal consistency, and for parallel consistency. Criteria for selecting the most beneficial data validation procedures are discussed. Examples for most of the procedures are in a report, the basis for this discussion.<sup>1</sup>

##### 1.4.17.2.1 INTRODUCTION

The primary purpose of this section is to describe several data validation procedures which can be used by either local, State, or Federal agencies for ambient air monitoring data. A secondary purpose is to suggest criteria for selecting the procedures which would be most suitable to the particular application.

Data validation will refer to those activities performed after the fact, that is, after the data have been collected. The difference between data validation and quality control techniques is that the quality control techniques attempt to minimize the amount of bad data being collected, while data validation seeks to prevent any bad data from getting through the data collection and storage systems. Thus data validation serves as a final screen before the data are used in decision making.

The validation may be performed by a data validator, a researcher using an existing data bank, or by a member of a

field team or local agency. It is preferable that data validation be performed as soon as possible after the data collection, so that the questionable data can be checked by recalling information on unusual events and on meteorological conditions which can aid in the validation. Also, timely corrective actions may be taken when indicated to minimize further generation of questionable data.

The following sections describe the data validation procedures and the selection criteria, abstracted from a data validation report.<sup>1</sup> Because of the limitation in space, the interested reader should refer to the report for detailed examples. In addition the reader would benefit by referring to several other pertinent references.<sup>2-7</sup>

#### 1.4.17.2.2 DATA VALIDATION PROCEDURES

Descriptions of the several data validation procedures are subdivided for convenience of use into four categories:

1. Routine check and review procedures which should be used to some extent in every validation process,
2. Tests for internal consistency of the data,
3. Tests for consistency of data sets with previous data (historical or temporal consistency), and
4. Tests for consistency with other data sets, collected at the same time or under similar conditions (consistency of parallel data sets).

The four categories are described in the following four subsections in order of increasing statistical sophistication in each category.

1.4.17.2.2.1 Routine validation procedures - Routine checks should include the following:

1. Data identification checks,
2. Unusual event review,
3. Deterministic relationship checks, and

4. Data processing procedures.

Data Identification Checks - Data with improper identification codes are useless. Three equally important identification fields which must be correct are time, location, and parameter. Examples of data identification errors noted by the EPA regional offices include: (1) improper State identification codes; (2) data identified for a nonexistent day (e.g., October 35); and (3) duplicate data from one monitoring site, but no data from another. Since most of these are human error, an individual other than the original person preparing the forms should scan the data coding forms prior to using the data as computer input or in a manual summary. If practical, the data listings should also be checked after entry into a computer system or data bank.

Unusual Event Review - A log should be maintained by each agency to record extrinsic events (e.g., construction activity, duststorms, unusual traffic volume, and traffic jams) that could explain unusual data. Depending on the purpose of data collection, this information could also be used to explain why no data are reported for a specified time interval, or it could be the basis for deleting data from a file for specific analytical purposes.

Deterministic Relationship Checks - Data sets which contain two or more physically or chemically related parameters should be routinely checked to ensure that the measured values on an individual parameter do not exceed the corresponding measured values of an aggregate parameter which includes the individual parameter. For example,  $\text{NO}_2$  values should not exceed  $\text{NO}_x$  values recorded at the same time and location. The following table lists some, but not all, of the possible deterministic relationship checks involving air quality and meteorological parameters.

<u>Individual parameter</u>		<u>Aggregate parameter</u>
NO <sub>2</sub> (nitrogen dioxide)	must be less than	NO <sub>x</sub> (nitrogen oxides)
CH <sub>4</sub> (methane)	must be less than	THC (total hydrocarbon)
SO <sub>2</sub> (sulfur dioxide)	must be less than	total sulfur
Pb (lead)	must be less than	TSP (total suspended particulates)

Data sets in which individual parameter values exceed the corresponding aggregate values should be flagged for further investigation. Minor exceptions to allow for measurement system noise may be permitted in cases where the individual value is a large percentage of the aggregate value.

Data Processing Practices - Reference 5 identifies 67 procedures currently in use for detecting and, when possible, correcting errors as they occur in computer systems. A review of this reference reveals that several of these procedures are within the categories of internal, historical, and parallel data consistency checks.

1.4.17.2.2 Tests for Internal Consistency - These tests check values in a data set which appear atypical when compared to the whole data set. Common anomalies of this type include unusually high or low values (outliers) and large differences in adjacent values. These tests will not detect errors which alter all values of the data set by either an additive or multiplicative factor (e.g., an error in the use of the scale of a meter or recorder). The following tests for internal consistency are listed in order of increasing statistical sophistication.

1. Data plots,
2. Dixon ratio test,
3. Grubbs test,
4. Gap test,

5. "Johnson" p test, and
6. Multivariate test.

Data Plots - Data plotting is one of the most effective means of identifying possible data anomalies. However, plotting all data points may require considerable manual effort or computer time. The number of data plots required can be reduced by plotting only those data which have been identified by a statistical test (or tests) (e.g., a Dixon ratio test) to be questionable. Nevertheless, data plots will often identify unusual data that would not ordinarily be identified by other internal consistency tests.

Dixon Ratio Test - The Dixon ratio test is the simplest of the statistical tests recommended for evaluating the internal consistency of data. The test for the largest value requires only the identification of the lowest ( $x_1$ ) and two highest values ( $x_{n-1}$  and  $x_n$ ) in the data set. The ratio

$$R = \frac{x_n - x_{n-1}}{x_n - x_1} \quad (1)$$

is calculated and then compared to a tabulated value in the appropriate table.<sup>1</sup> Consistency is indicated by a ratio near zero; a possible data anomaly is indicated by a ratio near unity. This test is ideally suited for moderate-sized data sets (e.g., a month of daily average values). The critical values of the ratio are derived on the assumption of a normal distribution; hence, a logarithmic transformation is usually required for TSP or other pollutant data.

Grubbs Test - This test, like the Dixon ratio assumes the normal distribution; however, it requires computation of the mean ( $\bar{x}$ ) and the standard deviation ( $s$ ) of the data. The test statistic is

$$T = \frac{x_n - \bar{x}}{s} \quad (2)$$

where  $x_n$  is the largest value in the data set. The calculated T is compared to a tabulated value at an appropriate level of risk.

Gap Test - This test identifies possible data anomalies by examining the length of the gap (or distance) between the two largest values ( $x_n$  and  $x_{n-1}$ ), the second and third largest values ( $x_{n-1}$  and  $x_{n-2}$ ), and similarly for other gaps. The two-parameter exponential distribution is fitted to the upper tail of the distribution of the sample data, and the probabilities of the observed gap sizes determined. If the probability is very small, the larger value is considered as a possible data anomaly.

"Johnson" p Test - This test fits a distribution function to the upper tail of the sample data distribution,<sup>1</sup> and then compares the consistency of the largest value with that predicted by the fitted distribution (e.g., lognormal or Weibull distribution).

Multivariate Test Procedures - The procedures given previously in this subsection can be used for testing data sets involving more than one variable by applying them independently to each variable; however, this approach may be inefficient, particularly when the variables are statistically correlated. In some cases a multivariate test will show that a value of one variable that appears to be an outlier using a single variable test procedure is consistent with the data set when one or more other variables are considered. Conversely, there may be a value of one variable which is consistent with the other data in the set when considering only one variable, but which is definitely a possible outlier when considering two variables.

Multivariate tests which have been successfully used for data validation checks include cluster analysis techniques,<sup>8</sup> principal component analysis,<sup>9</sup> and correlation methods. Applications of these methods usually require computerized procedures. For example, the cluster analysis technique can be applied using a program called NORMIX.<sup>10</sup>

1.4.17.2.2.3 Tests for Historical Consistency - These tests check the consistency of the data set with respect to similar data recorded in the past. In particular these procedures will detect changes where each item is increased (decreased) by a constant or by a multiplicative factor. This is not the case for the procedures in the previous section. These tests for historical consistency include:

1. Gross limit checks,
2. Pattern and successive difference tests,
3. Parameter relationship tests, and
4. Shewhart control chart.

Gross Limit Checks - Gross limit checks are useful in detecting data values that are either highly unlikely or generally considered impossible. Upper and lower limits are developed by examining historical data for a site (or for other sites in the area). Whenever possible, the limits should be specific for each monitoring site and should consider both the parameter and instrument/method characteristics. Table 1.4.17.1 shows examples of gross limit checks that have been used for ambient air monitoring data.<sup>11,12</sup> Although these checks can easily be adapted to computer applications, they are particularly appropriate for technicians who reduce data manually or who scan the strip charts to detect unusual events.

TABLE 1.4.17.1. EXAMPLES OF HOURLY GROSS LIMIT CHECKS FOR AMBIENT AIR MONITORING<sup>11,12</sup>

Parameter	Limits	
	Lower	Upper
Ozone	0 ppm	1 ppm
NO <sub>2</sub>	0 ppm	2 ppm
CO (carbon monoxide)	0 ppm	100 ppm
Total hydrocarbons	0 ppm	25 ppm
Total sulfur	0 ppm	1 ppm
Windspeed	0 m/s	22.2 m/s
Barometric pressure	950 mb	1050 mb

Pattern Tests - These tests check the data for pollutant behavior which has never or very rarely occurred in the past. Like the gross limit checks, they require that a set of limits be determined empirically from prescreened historical data. Values representing pollutant behavior outside of these predetermined limits are then flagged for further investigation. EPA has recommended the use of the pattern tests which place upper limits on:

1. The individual concentration value (maximum-hour test),
2. The difference in adjacent concentration values (adjacent hour test),
3. The difference or percentage difference between a value and both of its adjacent values (spike test), and
4. The average of four or more consecutive values (consecutive value test).<sup>2</sup>

The maximum-hour test (a gross limit check) can be used with both continuous and intermittent data; the other three tests should be used only with continuous data.

Table 1.4.17.2 is a summary of limit values developed by EPA for hourly average data. These values were selected on the basis of empirical tests on actual data sets. Note that the limit values vary with data stratifications (e.g., day/night).

TABLE 1.4.17.2. PARTIAL LISTING OF LIMITS USED IN EPA REGION V FOR PATTERNS TESTS

Pollutant (units)	Data stratification	Maximum hour	Adjacent hour	Spike	Consecutive 4-hour
Ozone-total oxidant ( $\mu\text{g}/\text{m}^3$ )	summer day	1000	300	200(300%)	500
	summer night	750	200	100(300%)	500
	winter day	500	250	200(300%)	500
	winter night	300	200	100(300%)	300
Carbon monoxide ( $\text{mg}/\text{m}^3$ )	rush traffic hours	75	25	20(500%)	40
	nonrush traffic hours	50	25	20(500%)	40

These limit values are usually inappropriate for other pollutants, data stratifications, averaging times, or EPA regions; thus, the data analyst should develop the required limit values by examining historical data similar to the data being tested. These limit values can be later modified if they flag too many values that prove correct or if they flag too few errors. Pattern tests should continue to evolve to meet the needs of the analyst and the characteristics of the data.

Parameter Relationship Tests - Parameter relationship tests can be divided into deterministic tests involving the theoretical relationships between parameters (e.g.,  $\text{NO} < \text{NO}_x$ ) or empirical tests which determine whether or not a parameter is behaving normally in relation to the observed behavior of one or more other parameters (e.g.,  $\text{NO}$  and  $\text{O}_3$ ). Determining the "normal" behavior of related parameters requires the detailed review of historical data and usually the application of the least squares method.

The following area-specific example illustrates the testing of meteorological data using a combination of successive value tests, gross limit tests, and parameter relationship tests. The validation protocol specifies that the following procedures be applied to ambient temperature data based on the availability of hourly averages reported in monthly formats:

1. Check the hourly average temperature. The minimum should occur between 04-09 hours, and the maximum should occur between 12-17 hours.

2. Inspect the hourly data for each day. Hourly changes should not exceed  $10^\circ\text{F}$ . If a decrease of  $10^\circ\text{F}$  or more occurs, check the wind direction and the precipitation summaries. The wind direction should have changed to a more northerly direction and/or rainfall of 0.15 in. or more per hour should have fallen.

3. Hourly values should not exceed predetermined maximum or minimum values based on month of the year. For example, in November the maximum allowable temperature is  $85^\circ\text{F}$  and the minimum is  $10^\circ\text{F}$ .

If any of the above criteria are not met, the data for the appropriate time period should be flagged for anomaly investigation.

In this example, relationship checks have been developed for temperature and wind direction as well as temperature and precipitation. Other pairs of parameters for which these checks could be developed include solar insolation and cloud cover; windspeed aloft and ground windspeed; O<sub>3</sub> and NO; and temperature and humidity.

Shewhart Control Chart - The Shewhart control chart is a valuable supplement to the gross limit and pattern tests because the chart identifies data sets which have mean or range values that are inconsistent with past data sets. The normal procedure for using the control chart is to determine control limits from past "in control" data and to compare future data points to these limits. However, after-the-fact control chart analyses are also of considerable value. The steps involved in constructing a control chart are described in Appendix H. Also described in Appendix H are criteria commonly used to determine when the measured values have exceeded control limits. An example of the use of control charts to ambient air pollutant data is described in Reference 1.

1.4.17.2.2.4 Tests for Consistency of Parallel Data Base - The tests for internal consistency (previously described) implicitly assume that most of the values in a data set are correct. Consequently, if all of the values in a data set incorporate a small positive bias, tests such as the Dixon ratio test would not indicate that the data set is inconsistent. One method of identifying a systematic bias is to compare the data set with other data sets which presumably have been sampled from the same population (i.e., same air mass and time period) and to check for differences in the average value or overall distribution of values. Four tests are presented here in order of increasing computational complexity. The first three are nonparametric

(i.e., they do not assume that the data have a particular distribution) and can be used for the nonnormal data sets which frequently occur in air quality analysis. The four tests are:

1. Sign test,
2. Wilcoxon signed-rank test,
3. Rank sum test, and
4. Intersite correlation test.

Sign Test - The sign test is a relatively simple way of testing the assumption that two paired samples (e.g., data sets from adjacent monitoring instruments on the same days) have the same median. The data analyst determines the sign (+ or -) of the algebraic difference in the measurement pairs and then counts the total number of positive signs ( $n_+$ ) and negative signs ( $n_-$ ); zero differences are ignored. For  $N = n_+ + n_- > 25$ , the normal approximation is adequate, that is, the variable ( $z$ ) which is approximately normally distributed is computed,

$$z = \frac{2n - N}{\sqrt{N}},$$

where  $n$  is the lesser of  $n_+$  and  $n_-$ . If for example,  $z$  is  $< -2$ , the two data sets would be inferred to have different medians, at about 0.05 significance level.

Wilcoxon Signed Rank Test - This test is similar to the sign test, but the signed ranks are used instead of only the sign. This test is generally more powerful because it considers both the sign and the magnitude of the difference in terms of a rank. See the report for an example.<sup>1</sup>

Rank Sum Test - This test differs from the previous two tests in that the two data sets are not paired and hence unrelated. A detailed example is in the report.<sup>1</sup>

Intersite Correlation Test - This test is generally applicable to two correlated data sets (e.g., TSP measurements on the same days at two neighboring sites). An example is in the report to illustrate how the data from the two sites aid in the

correct identification of a possible data anomaly.<sup>1</sup> The plot is in Figure 1.4.17.1 for the example in the report.<sup>1</sup> An ellipse is drawn to include approximately 95% of the data points. Points outside the ellipse may be data anomalies, and each point should be investigated. Close examination reveals that reviewing one variable at a time may lead to an inconsistent decision relative to these data. For example, the value at (175,129) would appear to be a possible anomaly when studying the data from one site, but it would appear to be consistent when considering the data from both sites.

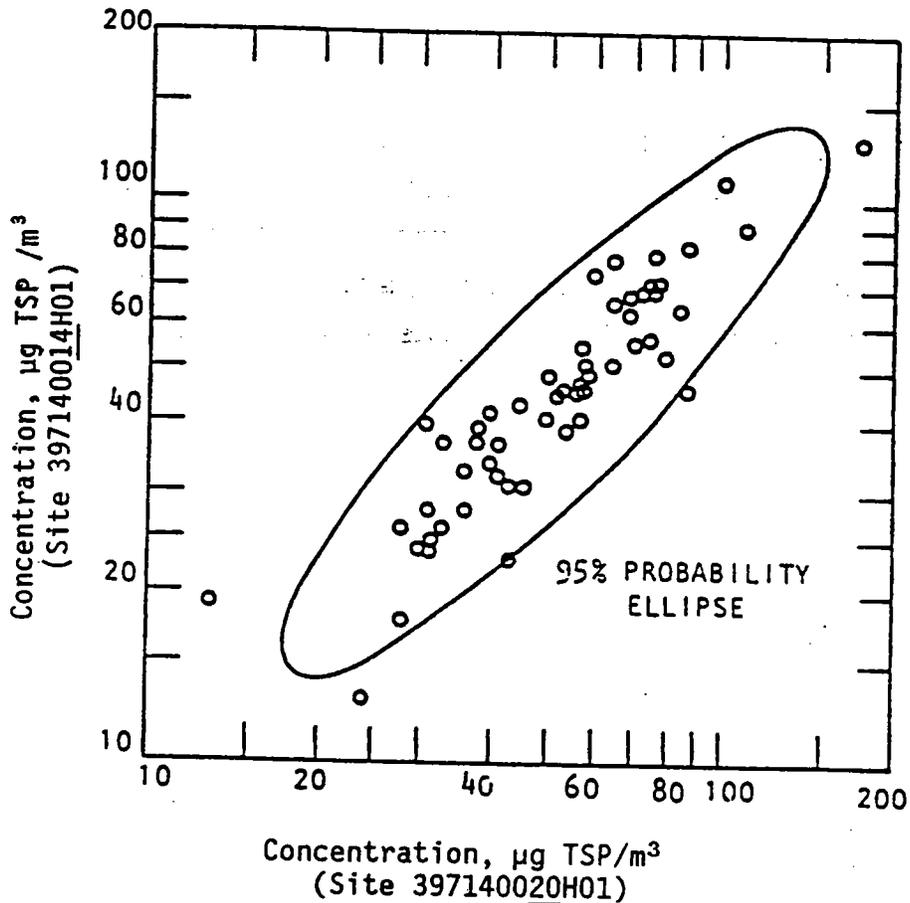


Figure 1.4.17.1. Intersite correlation test data.

#### 1.4.17.2.3 SELECTION OF THE DATA VALIDATION PROCEDURE

Selection of the most beneficial data validation procedures depends on several factors. For example, a local agency with no computer facility and with limited staff and minimum statistical support should consider the following procedures first: data ID checks, unusual event review, deterministic relationship checks, Dixon ratio test for a single questionable value, data plots, gross limit and pattern checks, and possibly the control chart.

On the other hand, a large agency with extensive computer capabilities and statistical support can use any of the validation procedures, especially those with heavy emphasis on computerized graphics, Shewhart control charts, distributions fitted to the data, and parameter relationships. After experience is gained with the types of data anomalies which occur, selection of the specific procedure can be more efficient, and a given procedure can be improved by altering the limits to change its sensitivity (e.g., redefining the gross limit or pattern checks or improving the pattern relationships). Thus it is necessary to maintain good documentation on the data identified as questionable; the source of error; if any, associated with these data; the number of questionable data values ultimately inferred to be correct; the techniques used in flagging the data; and other information pertaining to the cost of performing the data validation.

#### 1.4.17.3 REFERENCES

1. Nelson, A. C., D. W. Armentrout, and T. R. Johnson. Validation of Air Monitoring Data. EPA-600/4-80-030, June 1980.
2. U.S. Environmental Protection Agency. Screening Procedures for Ambient Air Quality Data. EPA-450/2-78-037. July 1978.
3. Rhodes, R. C., and S. Hochheiser. Data Validation Conference Proceedings. Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, EPA-600/9-79-042. September 1979.

4. Barnett, V., and T. Lewis. Outliers in Statistical Data. John Wiley and Sons, New York, 1978.
5. U.S. Department of Commerce. Computer Science and Technology: Performance Assurance and Data Integrity Practices. National Bureau of Standards, Washington, D.C., January 1978.
6. Naus, J. I. Data Quality Control and Editing. Margel Dekker, Inc., New York, New York.
7. 1978 Annual Book of ASTM Standards, Part 41. Standard Recommended Practice for Dealing with Outlying Observations, ASTM Designation: E 178-75. pp. 212-240.
8. Marriott, F. H. C. The Interpretation of Multiple Observations. Academic Press, New York, 1974.
9. Hawkins, D. M. "The Detection of Errors in Multivariate Data Using Principal Components." Journal of the American Statistical Association, Vol. 69, No. 346, 1974.
10. Wolfe, J. H. "NORMIX: Computation Methods for Estimating the Parameters of Multivariate Normal Mixtures of Distributions," Research Memo, SRM 68-2. U.S. Naval Personnel Research Activity, San Diego, California, 1967.
11. U.S. Environmental Protection Agency. Guidelines for Air Quality Maintenance Planning and Analysis. Vol. 11. Air Quality Monitoring and Data Analysis. EPA-450/4-74-012. 1974.
12. U.S. Environmental Protection Agency. Quality Assurance and Data Validation for the Regional Air Monitoring System of the St. Louis Regional Air Pollution Study. EPA-600/4-76-016. 1976.



## 1.4.18 STATISTICAL ANALYSIS OF DATA

### 1.4.18.1 ABSTRACT

A number of statistical tools and techniques are described in the appendices. The appendices are organized in part by functional or application area rather than by statistical nomenclature. For example, Appendix J concerns the subject of calibration; however, least squares or regression analysis, a useful tool for determining calibration curves, can also be used for estimating the resulting precision of the reported pollutant concentration from a specific analyzer reading. The statistical tools should be used with discretion.

A glossary of major statistical terms is included as Appendix A. Appendix B includes symbol definitions used throughout the remaining appendices.

### 1.4.18.2 DISCUSSION

Summary statistics - Summary statistics such as the mean and the standard deviation are used to simplify the presentation of data and at the same time to summarize essential characteristics. Appendix C includes a discussion of summary statistics.

Frequency distributions - Frequency distributions such as normal, log-normal, and Weibull distributions are used to summarize and present relatively large data sets, such as the daily concentrations of suspended particulates in ambient air. Appendix D discusses frequency distributions.

Estimation procedures - Statistical estimation procedures are used to make inferences concerning the conceptual population of measurements made under the same conditions based on a small sample of data. An example would be the estimation of the average pH of a large number (population) of filters based on a

sample of pH readings for seven filters. Appendix E discusses estimation procedures.

Outliers - Outliers, that is, unusually large or small values, are identified by appropriate statistical tests for outliers. These statistical tests are useful in data validation, for example, in identifying gross errors in data handling procedures. Appendix F is a treatment of outliers and data validation. For additional information on data validation refer to Section 1.4.17.

Audit data - Methods for treating performance audit data and for presenting the results in terms of bias and precision are included in Appendix G.

Control charts - Techniques for selecting the type of control chart, for determining the limits, and for interpreting plotted results are presented in Appendix H.

Sampling - Sampling techniques apply to many phases of a quality assurance program. Methods for selecting a random sample, as well as procedures for acceptance sampling, are briefly discussed in Appendix I.

Calibration - Calibration procedures represent one of the critical sources of measurement error. Appendix J is a discussion of calibration procedures. Control charts should be used to indicate when a new multipoint calibration is to be conducted.

Replication, repeatability, and reproducibility tests - The identification of sources of measurement error within and among laboratories is one of the important functions of the Quality Assurance Coordinator. Programs for doing this are discussed in Appendix K.

Reliability and maintainability - As measurement systems become more complex, system reliability becomes an increasingly important parameter in determining the completeness and accuracy of the results. Reliability is discussed in Appendix L.

1.4.18.3 BIBLIOGRAPHY

1. Burr, I. W. Engineering Statistics and Quality Control. McGraw-Hill, New York. 1953.
2. Duncan, A. J. Quality Control and Industrial Statistics. 3rd Ed. Richard D. Irwin, Inc., Homewood, Illinois. 1965.
3. Juran, J. M., (ed.). Quality Control Handbook. 2nd Ed. McGraw-Hill, New York. 1962.
4. Snedecor, G. W., and Cochran, W. G. Statistical Methods. 6th Ed. Iowa State College Press, Ames, Iowa. 1967.
5. Youden, W. J., Statistical Techniques for Collaborative Tests, The Association of Official Analytical Chemists, Box 540, Benjamin Franklin Station, Washington, D. C. 20044.
6. Hald, A., Statistical Theory with Engineering Applications, John Wiley and Sons, Inc., New York, 1952.
7. Bennett, C. A. and N. L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry, John Wiley & Sons, Inc., New York, 1954.
8. Grant, E. I., and Leavenworth, R. S., Statistical Quality Control, Fourth Edition, McGraw-Hill Book Co., New York, 1972.



## 1.4.19 CONFIGURATION CONTROL<sup>1,2</sup>

### 1.4.19.1 ABSTRACT

1. Configuration control is used to record changes in air pollution measurement method equipment and the physical arrangement of this equipment in the monitoring system.

2. Configuration control may be grouped into two types depending on the purpose:

a. Provides history (record) of changes during the life of the monitoring project.

b. Provides design and operation data on the first monitoring instrument or system when multiple instruments or systems are planned. This information is commonly obtained by a First Article Configuration Inspection (FACI). An example of a FACI is shown for a major EPA monitoring network in the discussion portion.

3. Configuration control record procedures are the same as those used for document control (Section 1.4.1).

### 1.4.19.2 DISCUSSION

Difference between Configuration and Document Control - Document control, described in Section 1.4.1, is used to make sure all personnel on a monitoring project are using the same and most current written procedures for sampling, analysis, calibration, data collection and reporting, auditing, etc. When revisions are made in these procedures, they should be documented as described in Section 1.4.1. Similarly, a system is needed to record changes made in the equipment and/or physical arrangement of this equipment in the monitoring system that are not included as part of document control. This system is called configuration control.

Types of Configuration Control - Configuration control may be grouped into two types, depending on the intended purpose of the information.

In the first type, a history of changes is maintained throughout the life of the monitoring project. This history is valuable during problem-solving investigations that may occur either during the project life or long after the project has been completed. Subtle changes in the equipment used in the monitoring system may have significant effects on the measured pollutant concentrations. Such equipment changes would normally not appear under document control on the procedure used for sampling and analysis. By way of example, these changes might include:

1. Replacement of monitoring instrument or component part with a different model type (equipment change).
2. Replacement of filter used to remove particulates prior to instrumental gaseous-pollutant analysis with a different filter type (equipment change).
3. Relocation of an air pollution sampler to a different spot at the sampling site (rearrangement of same equipment).

Each project officer must decide the scope of configuration control that should be applied to his project.

The second type of configuration control is used to provide information on engineering design and operation on the first monitoring instrument or station when multiples are planned. This information is commonly obtained and documented by a First Article Configuration Inspection (FACI). The FACI is most important for large complex monitoring projects, particularly when pollutant sensor outputs are stored on-site or transmitted to a central facility for computer storage. Purchase contracts that involve multiple instrument systems of identical design and/or monitoring stations of identical design should require a FACI as part of the contract.

By way of example, the FACI required as part of the contract for the EPA Regional Air Monitoring System (RAMS) will be briefly described. The RAMS was a network of 25 monitoring sites in and around the St. Louis area, designed to collect ambient air and meteorological measurements for diffusion modeling and other purposes. When the first monitoring station was installed, a FACI was completed as required by the contract. The FACI covered the following:

1. Shelter system.
2. Gas analyzing system (sensor for ozone, nitrogen oxides, total hydrocarbons, carbon monoxide, and total sulfur).
3. Particulate sampling system (including sensor for light scattering).
4. Meteorological system (sensor for wind speed, wind direction, temperature and dew point).
5. Data acquisition system.

For each system, the FACI consisted of a physical inspection, a functional demonstration, and an operational test consistent with requirements in the contract. To facilitate and semi-formalize the exchange of information between EPA and the contractor during the FACI, "squawk sheets" were used. These sheets allowed discrepancies to be noted by EPA and were responded to by the contractor. An example of the RAMS squawk sheet is shown in Figure 1.4.19.1. The contractor prepared a formal response to all squawk sheets.

The procedures described in Section 1.4.1 for document control are also applicable for configuration control of hardware over the project life.

Squawk title \_\_\_\_\_

Number \_\_\_\_\_

Date \_\_\_\_\_

Author \_\_\_\_\_

Squawk description:

EPA  
Coordinator \_\_\_\_\_

[EPA comment]

Contractor action/response

Contractor Program  
Engineer \_\_\_\_\_

[Contractor response to EPA comment]

Final disposition

Contractor Program Engineer \_\_\_\_\_

EPA Project Officer \_\_\_\_\_

1.4.19.3 REFERENCES

1. Lowers, H. R. Quality Assurance in Configuration Management. Quality Progress. V(6):17-19, June 1972.
2. Covino, C. P., and Meghri, A. W. Quality Assurance Manual. Industrial Press, Inc., New York. 1967. pp. 31a, 31b, and 32a.

BIBLIOGRAPHY

1. Configuration Management. Air Force Systems Command Manual 375-1. 1960.



## 1.4.20 RELIABILITY

### 1.4.20.1 ABSTRACT

Reliability of an air pollution measurement system (or any system) is defined as the probability that the system will perform its intended function for a prescribed period of time under the operating conditions specified, or, conversely, unreliability is the probability that a device will fail to perform as specified. Reliability is becoming increasingly important in air pollution measurement because of the increase in complexity and sophistication of sampling, analysis, automatic recording, and telemetering systems. Furthermore, data interpretation for trend analyses depends on a high percentage of data completeness (e.g., less than 10 to 20% missing data. Generally, as the measurement system becomes more complicated, its probability of failure increases. In order to ensure high equipment reliability the following should be considered:

1. Specify equipment reliability in contracts--select high reliability components.
2. Inspect and test incoming equipment for adherence to contract specifications (e.g., conduct performance acceptance tests) or have equipment supplier conduct these tests.
3. Control the operating environment that influences the reliability of the equipment and hence the measurements.
4. Provide for adequate training of personnel.
5. Provide preventive maintenance to reduce or minimize wear out failures.
6. Provide records of failures, analyze and use these data to initiate corrective actions, and predict failure rates.

#### 1.4.20.2 DISCUSSION

In order to ensure high reliability of equipment (and hence the completeness of data), the following should be considered:

Specify equipment reliability requirements in contracts<sup>1</sup> -

These requirements constitute a specification to be met by the manufactured product. This specification should consist of:

1. The product reliability definition, which includes:
  - a. All functional requirements of the equipment.
  - b. Safety requirements.
  - c. Environmental conditions for the reliability demonstration tests.

2. Where applicable, give required reliability expressed as a minimum mean time between failures (MTBF).<sup>2</sup> The MTBF is the average time that the system performs its required function without failure. This may be expressed as hours, days, or number of monitoring periods. It is estimated by averaging the recorded times of successful system performance.

3. Required performance demonstration tests.

Inspect and test incoming equipment for adherence to contract specifications<sup>3</sup>

1. Quality control tests should be conducted to determine whether the product in question meets performance and design specifications at the time of testing.

2. Burn-in tests should be conducted for specified times where there is an indication of early failures.

3. If appropriate, reliability demonstration and/or performance tests should be conducted on a sample of equipments, testing until failure or for a specified time, to:

- a. Verify adherence to specified reliability standards.
- b. Generate data for product improvement.

c. Provide an estimate of product service life and reliability.

Control the operating conditions<sup>4</sup> - Environmental factors affecting performance or reliability may be natural, induced, or a combination of both.

1. Natural environmental factors are:
  - a. Barometric pressure changes.
  - b. Temperature.
  - c. Particulate matter, such as sand, dust, insects, and fungus.
  - d. Moisture, such as icing and salt spray.
2. Induced factors are:
  - a. Temperature, self-generated or generated by adjacent or ancillary equipment.
  - b. Dynamic stresses, such as shock vibration.
  - c. Gaseous and particulate contamination, such as exhaust or combustion emissions.

3. Combined natural and induced conditions. Frequently, the stresses affecting an item result from a combination of one or more factors from both classes. Such combinations may intensify the stress, or the combined factors may tend to cancel out each other.

Provide for adequate training of personnel<sup>5,6</sup> - The implementation of a reliability assurance program requires a training program at both the operational and supervisory levels. At the operator level, instruction should be given in the collection of failure and maintenance data, in the maintenance function (both preventive and unscheduled maintenance or repair of the equipment), and in the control of operating conditions. This training can be accomplished by use of lectures, films, posters, and reliability information bulletins.

At the supervisory level, in addition to the above, training should be given in the analysis of reported data, program planning, and testing procedures.

The reliability of the measurement system depends to a large extent on the training of the operator. The completeness of the data, as measured by the proportion of valid data reported, is a function of both the reliability and maintainability of the equipment/measurement system.

Consider maintainability at time of purchase - Maintainability is the probability that the system will be returned to its operational state within a specified time after failure. For continuous air pollution monitoring instruments, maintainability is an important consideration during procurement, and in some cases should be included in the purchase contract. Maintainability items to consider at the time of procurement include:

1. Design factors.
  - a. Number of moving parts.
  - b. Number of highly stressed parts.
  - c. Number of heat producing parts.
2. Ease of repair after failure has occurred.
3. Maintainability cost.
  - a. Inventory of spare parts required.
  - b. Amount of technician training required for repair.
  - c. Factory service required.
  - d. Service repair contract required.
  - e. Estimated preventive maintenance required.

Provide preventive maintenance - In order to prevent or minimize the occurrence of wear out failure, the components of the system subject to wear out must be identified and a preventive maintenance schedule implemented. This aids in improving the completeness of the data. Maintenance can be performed

during nonoperational times for noncontinuous monitoring equipments, resulting in no downtime. Replacement units must be employed in continuous monitoring systems in order to perform the maintenance while the system is performing its function. Downtime may also be scheduled.

Provide records of failure and maintenance; analyze and use to initiate corrective actions - Field reliability data should be collected in order to:

1. Provide information upon which to base failure rate predictions.
2. Provide specific failure data for equipment improvement efforts.
3. Provide part of the information needed for corrective action recommendations.

A more complete discussion of reliability and maintainability is contained in Appendix L.

#### 1.4.20.3 REFERENCES

1. Muench, J. O. A Complete Reliability Program. Proceedings Annual Reliability and Maintainability Symposium. Institute of Electrical and Electronic Engineers, Inc. 1972. pp. 20-23.
2. Enrick, N. L. Quality Control and Reliability. 6th ed. Industrial Press, Inc., New York, New York. 1972. Chapter 16, pp. 219-238.
3. MIL-STD-718B, Reliability Tests, Exponential Distribution. 1967. Department of Defense, Washington, D.C.
4. Haviland, R. P. Engineering Reliability and Long Life Design. D. Van Nostrand Co., Inc., Princeton, New Jersey. 1964. Chapter 10, pp. 149-168.
5. Bazovsky, I. Reliability Theory and Practice. Prentice-Hall, Englewood Cliffs, New Jersey. 1961. Chapter 9, pp. 76-84.
6. Juran, J. M. Quality Control Handbook. 2nd edition. McGraw-Hill, New York. 1962. Sec. 20, pp. 20-2 - 20-38 and Sec. 13, pp. 13-17 - 13-19.



## 1.4.21 QUALITY REPORTS TO MANAGEMENT

### 1.4.21.1 ABSTRACT

Several reports are recommended in the performance of the quality assurance tasks. Concise and accurate presentation of the data and derived results is necessary. Some of the quality assurance reports for management are:

1. Data quality assessment reports (e.g., those specified in 40 CFR, Part 58, Appendices A and B),
2. Performance and system audit reports,
3. Interlaboratory comparison summaries,
4. Data validation reports,
5. Quality cost reports,
6. Instrument or equipment downtime,
7. Quality assurance program and project plans, and
8. Control charts.

Reports should be prepared with the following guidelines as appropriate.

1. All raw data should be included in the report when practical.
2. Objective of the measurement program, in terms of the data required and an uncertainty statement concerning the results.
3. Methods of data analysis should be described unless they are well-documented in the open literature.
4. A statement on any limitation and on applicability of the results should be included.
5. Precision and accuracy of the measurement methods should be stated.
6. Quality control information should be provided as appropriate.

7. Reports should be placed into a storage system in order that they may be retrieved as needed for future reference.

#### 1.4.21.2 DISCUSSION

There are several quality assurance reports that should be prepared periodically (quarterly or annually) summarizing the items of concern. These reports will be briefly discussed below.

##### 1. Data Quality Assessment Reports

40 CFR Part 58, Appendices A and B require that reports of the precision and accuracy calculations be submitted each quarter along with the air monitoring data. See References 1 and 2 for details of the calculations and for specific data/ results to be reported.

##### 2. Performance and System Audit Reports

Upon completion of a performance and/or system audit, the auditing organization should submit a report summarizing the audit and present the results to the auditee to allow initiation of any necessary corrective action.

##### 3. Interlaboratory Comparison Summaries

EPA prepares annual reports summarizing the interlaboratory comparisons for the National Performance Audit Program. In addition, the results from this audit are submitted to the participating labs as soon as possible after the audit. These data can then be used by the participants to take any necessary corrective action with regard to their measurement procedures. See Appendix K for a further discussion of the contents of the annual report.<sup>3,4</sup>

##### 4. Data Validation Report

It is recommended in Section 1.4.17 that a data validation process be implemented in order to minimize the reporting of data of poor quality. A periodic report of the results of the

data validation procedure should be made summarizing, for example, the number of items (values) flagged as questionable, the result of followup investigations of these anomalies, the final number of data values rejected or corrected as a result of the procedure, corrective action recommended, and effectiveness of the data validation procedures.<sup>5,6</sup>

5. Quality Cost Report

A quality cost system is recommended in Section 1.4.14. After the system has been implemented, a quality cost report should be made periodically to include the prevention, appraisal, and correction costs.<sup>7</sup>

6. Instrument or Equipment Downtime

In Section 1.4.7 it is recommended that records be maintained of the equipment in terms of failures, cause of failures, repair time, and total downtime. These data should be summarized periodically and submitted to management as an aid in future procurement.

7. Quality Assurance Program (or Project) Plans

Although these are not reports on results, they are plans for the QA activities for a QA program or project. They are the reports which indicate which QA reports should be prepared.

8. Control Charts

The control charts are a visual report of the analytical work and hence they are a significant part of the reporting system. A summary of the results of the control chart applications should appear in the summary report to management.

Some guidelines in the preparation of these reports are given in the Abstract portion of this section.

1.4.21.3 REFERENCES

1. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS), Federal Register, Vol. 44, Number 92, May 1979.
2. Appendix B - Quality Assurance Requirements for Prevention of Significant Deterioration (PSD) Air Monitoring, Federal Register, Vol. 44, Number 92, May 1979.
3. Streib, E. W. and M. R. Midgett, A Summary of the 1982 EPA National Performance Audit Program on Source Measurements. EPA-600/4-88-049, December 1983.
4. Bennett, B. I., R. L. Lampe, L. F. Porter, A. P. Hines, and J. C. Puzak, Ambient Air Audits of Analytical Proficiency 1981, EPA-600/4-83-009, April 1983.
5. Nelson, Jr., A. C., D. W. Armentrout, and T. R. Johnson. Validation of Air Monitoring Data, North Carolina, EPA-600/4-80-030, June 1980.
6. U.S. Environmental Protection Agency. Screening Procedures for Ambient Air Quality Data. EPA-450/2-78-037, July 1978.
7. Strong, R.B., J.H. White and F. Smith, "Guidelines for the Development and Implementation of a Quality Cost System for Air Pollution Measurement Programs," Research Triangle Institute, Research Triangle Park, North Carolina, 1980, EPA Contract No. 68-02-2722.

## 1.4.22 QUALITY ASSURANCE PROGRAM PLAN<sup>1</sup>

### 1.4.22.1 ABSTRACT

1. The QA Program Plan is a document which stipulates the policies, objectives, management structure, responsibilities, and procedures for the total QA programs for each major organization.<sup>1</sup> The EPA policy requires participation by all EPA Regional Offices, EPA Program Offices, EPA Laboratories, and States in a centrally managed QA program, and includes all monitoring and measurement efforts mandated or supported by EPA through regulations, grants, contracts, or other formalized means not currently covered by regulation.

2. Each EPA Program Office, EPA Regional Office, EPA Laboratory, and State and other organizations, is responsible for the preparation and implementation of the QA Program Plan to cover all environmentally-related measurement activities supported or required by EPA. A basic requirement of each plan is that it can be implemented and that its implementation can be measured.

3. Each QA Program Plan should include the following elements:

- a. Identification of office/laboratory submitting the plan,
- b. Introduction - brief background, purpose, and scope,
- c. QA policy statement,
- d. QA management structure,
- e. Personnel qualification and training needs,
- f. Facilities, equipment, and services - approach to selection, evaluation, calibration, operation, and maintenance,
- g. Data generation - procedures to assure the generation of reliable data,

h. Data processing - collection, reduction, validation, and storage of data,

i. Data quality assessment - accuracy, precision, completeness, representativeness, and comparability of data to be assessed,

j. Corrective action - QA reporting and feedback channels established to ensure early and effective corrective action, and

k. Implementation requirements and schedule.

4. Plans should be submitted through normal channels for review and/or approval.

#### 1.4.22.2 DISCUSSION

QA Program Plan is an orderly assembly of management policies, objectives, principles, and general procedures by which an agency or laboratory outlines how it intends to produce quality data. The content of the plan (outlined in 1.4.22.1) is briefly described below; eleven essential elements should be considered and addressed.

1. Identification - Each plan should have a cover sheet with the following information: document title, document control number, unit's full name and address, individual responsible (name, address, and telephone number), QA Officer, plan coverage, concurrences, and approval data.

2. Introduction - Brief background, purpose and scope of the program plan is set forth in this section.

3. QA policy statement - The policy statement provides the framework within which a unit develops and implements its QA program. It must emphasize the requirements and activities needed to ensure that all data obtained are of known quality.

4. QA management - This section of the plan shows the interrelationships between the functional units and subunits which generate or manage data. This includes the assignment of responsibilities, communications (organizational chart to indicate information flow), document control, QA program assessment.

5. Personnel - Each organization should ensure that all personnel performing tasks and functions related to data quality have the needed education, training, and experience; personnel qualifications and training needs should be identified.

6. Facilities, equipment, and services - The QA Program Plan should address the selection, evaluation, environmental aspects of equipment which might have an impact on data quality, maintenance requirements, monitoring and inspection procedures, for example.

7. Data generation - Procedures should be given to assure the generation of data that are scientifically valid, defensible, comparable, and of known precision and accuracy. QA Project Plans (as described in Section 1.4.23) should be prepared and followed. Standard operating procedures (SOP) should be developed and used for all routine monitoring programs, repetitive tests and measurements, and for inspection and maintenance of facilities, equipment, and services.

8. Data processing - The plan should describe how all aspects of data processing will be managed and separately evaluated in order to maintain the integrity and quality of the data. The collection, validation, storage, transfers, and reduction of the data should be described.

9. Data quality assessment - The plan should describe how all generated data are to be assessed for accuracy, precision, completeness, representativeness, and comparability.

10. Corrective action - Plans should describe the mechanism(s) to be used when corrective actions are necessary. Results from the following QA activities may initiate a corrective action: performance audits, system audits, interlaboratory comparison studies, and failure to adhere to a QA Program or Project Plan or to SOP.

11. Implementation requirements and schedule - A schedule for implementation is given in Reference 1.

1.4.22.3 REFERENCE

1. Guidelines and Specifications for Preparing Quality Assurance Program Plans, Quality Assurance Management Staff, Office of Research Development, USEPA, Washington, D.C., QAMS-004/80, September 1980. This document (EPA-600/8-83-024; NTIS PB 83-219667) may be obtained from the National Technical Information Service, 5885 Port Royal Road, Springfield, Virginia 22161.

#### 1.4.23 QUALITY ASSURANCE PROJECT PLAN<sup>1</sup>

##### 1.4.23.1 ABSTRACT

1. A QA Project Plan is an orderly assembly of detailed and specific procedures by which an agency or laboratory delineates how it produces quality data for a specific project. A given agency or laboratory would have only one QA Program Plan, but would have a project plan for each project or for each group of projects using the same measurement methods, (e.g., a laboratory service group might develop a plan by analytical instrument since the same service is provided to several projects). Every project that involves environmentally-related measurements should have a written and approved QA Project Plan.

2. Each of the 16 items listed below should be considered for inclusion in each QA Project Plan.<sup>1</sup>

1. Title page, with provision for approval signatures
2. Table of contents
3. Project description
4. Project organization and responsibilities
5. QA objectives for measurement data in terms of precision, accuracy, completeness, representativeness and comparability
6. Sampling procedures
7. Sample custody
8. Calibration procedures and frequency
9. Analytical procedures
10. Data analysis, validation, and reporting
11. Internal quality control checks and frequency
12. Performance and system audits and frequency
13. Preventive maintenance procedures and schedules

14. Specific procedures to be used to routinely assess data precision, accuracy, and completeness of specific measurement parameters involved
15. Corrective action
16. Quality assurance reports to management.

It is EPA policy that precision and accuracy of data must be assessed on all monitoring and measurement projects. Therefore, Item 14 must be described in all QA Project Plans.

#### 1.4.23.2 DISCUSSION

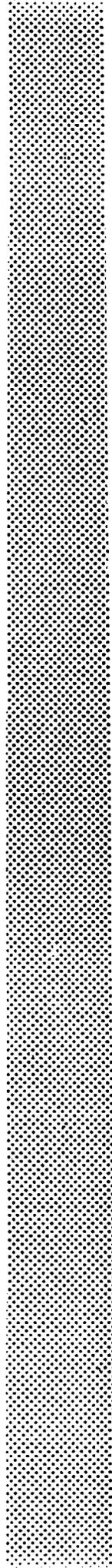
The guidelines and specifications for preparing QA Project Plans are in Appendix M. Appendix M also includes pertinent references, definition of terms, availability of performance audit materials/devices and QA technical assistance, and a model QA Project Plan.

#### 1.4.23.3 REFERENCE

1. Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans, Quality Assurance Management Staff, Office of Research Development, USEPA, Washington, D.C., QAMS-005/80, December 1980. This document (EPA-600/4-83-004; NTIS PB-83-170514) may be obtained from the National Technical Information Service, 5885 Port Royal Road, Springfield, Virginia 22161.

APPENDICES

7





APPENDIX A  
INDEX OF TERMS

	<u>Page</u>
A.1 DEFINITIONS	4
Quality Assurance	
Acceptance Sampling	4
Audit	4
Chain of Custody	4
Configuration Control	4
Data Validation	4
Document Control	5
Performance Audit	5
Quality	5
Quality Assurance	5
Quality Assurance Program Plan	6
Quality Assurance Project Plan	6
Quality Audit	6
Quality Control	6
Internal Quality Control	7
External Quality Control	7
Random Samples	7
Representative Sample	7
Sample	7
Standard Operating Procedure (SOP)	8
Statistical Control Chart	8
Stratified Sample	8
System Audit	9
Statistics	
Availability	9
Comparability	9

	<u>Page</u>
Completeness	9
Confidence Coefficient	9
Confidence Interval	10
Confidence Limits	10
Error	10
Maintainability	10
Measures of Central Tendency	10
Arithmetic Mean (Average)	10
Geometric Mean	10
Median	11
Mode	11
Measures of Dispersion or Variability	11
Range	11
Variance	11
Standard Deviation	12
Geometric Standard Deviation	12
Coefficient of Variation (Relative Standard Deviation)	13
Outlier	13
Random Error	13
Relative Error	13
Reliability (General)	14
Statistical Control Chart Limits	14
Systematic Error	14
Test Variability	14
Accuracy	14
Bias	14
Precision	14
Relative Standard Deviation	16
Repeatability	16
Replicability	16
Replicates	16
Reproducibility	17
Tolerance Limits	17

	<u>Page</u>
Testing or Measurement	
Analytical Limit of Discrimination	17
Blank or Sample Blank	17
Analytical or Reagent Blank	17
Dynamic Blank (or Field Blank)	17
Calibration	18
Dynamic Calibration	18
Static Calibration	18
Certified Reference Material (CRM) - Cylinder Gases	18
Collaborative Tests (or Studies)	18
Functional Analysis	19
Minimum Detectable Level	19
Proficiency Testing	19
Ruggedness Testing	19
Spiked Sample	19
Standards in Naturally Occurring Matrix	19
Standard Reference Material (SRM)	20
Standard Reference Sample (SRS)	20
Standards Depending upon "Purity" or Established Physical or Chemical Constants	20
Primary Standard	20
Secondary Standard	21
Standards Based upon Usage	21
Calibration Standard	21
Quality Control Reference Sample (or Working Standard)	21
Standardization	21
Traceability	21
A.2 REFERENCES	21
A.3 BIBLIOGRAPHY	22

## A.1 DEFINITIONS

### Quality Assurance

Acceptance Sampling - The procedures by which decisions to accept or reject a sampled lot or population are made based on the results of a sample inspection. In air pollution work, acceptance sampling could be used when checking a sample of filters for certain measurable characteristics such as pH, tensile strength, or collection efficiency to determine acceptance or rejection of a shipment of filters, or when checking the chemical content of a sample of vials of standard solutions from a lot of vials to be used in an interlaboratory test.

Audit - A systematic check to determine the quality of operation of some function or activity. Audits may be of two basic types: (1) performance audits in which quantitative data are independently obtained for comparison with routinely obtained data in an air pollution measurement system, or (2) system audits are of a qualitative nature and consist of an on-site review of a laboratory's quality assurance system and physical facilities for air pollution sampling, calibration, and measurement.

Chain of Custody - A procedure for preserving the integrity of a sample or of data (e.g., a written record listing the location of the sample/data at all times).

Configuration Control - A system for recording the original equipment configuration, physical arrangement and subsequent changes thereto.

Data Validation - A systematic effort to review data to identify any outliers or errors and thereby cause deletion or flagging of suspect values to ensure the validity of the data for the user. This "screening" process may be done by manual and/or computer methods, and may use any consistent technique such as pollutant concentration limits or parameter relationships to screen out impossible or unlikely values.

Document Control - A systematic procedure for indexing the original document (Revision No. 0, e.g.) and subsequent revisions (Revision No. 1, 2, ...) by number and date of revision. An example of a procedure is the one given in Section 1.4.1 and used throughout this Handbook.

Performance Audit - A quantitative analysis or check with a material or device with known properties or characteristics. The audit is performed by a person different from the routine operator/analyst using audit standards and audit equipment different from the calibration equipment. Such audits are conducted periodically to check the accuracy of a project measurement system. Some performance audits may require the identification of specific elements or compounds, in lieu of, or in addition to, a quantitative analyses. For some performance audits it may be impractical or unnecessary to have a different person than the routine operator/analyst; in these cases the routine operator/analyst must not know the concentration or value of the audit standards until the audit is completed. The other conditions of the audit must still be met, that is, the audit standards be different from the calibration standards, and the audit device be different from the calibration device.

Quality - The totality of features and characteristics of a product or service that bear on its capability to satisfy a given purpose. For air pollution measurement systems, the product is air pollution measurement data and the characteristics of major importance are accuracy, precision, completeness, and representativeness. For air monitoring systems, "completeness," or the amount of valid measurements obtained relative to the amount expected to have been obtained, is a very important measure of quality. The relative importance of accuracy, precision, and completeness depends upon the particular purpose of the user.

Quality Assurance - A system for integrating the quality planning, quality assessment, and quality improvement efforts of

various groups in an organization to enable operations to meet user requirements at an economical level. In air pollution measurement systems, quality assurance is concerned with all of the activities that have an important effect on the quality of the air pollution measurements as well as the establishment of methods and techniques to measure the quality of the air pollution measurements. The more authoritative usages differentiate between "quality assurance" and "quality control," quality control being "the system of activities to provide a quality product," and quality assurance being "the system of activities to provide assurance that the quality control system is performing adequately."

Quality Assurance Program Plan - An orderly assembly of management policies, objectives, principles, and general procedures by which an agency or laboratory outlines how it intends to produce data of acceptable quality.

Quality Assurance Project Plan - An orderly assembly of detailed and specific procedures by which an agency or laboratory delineates how it produces quality data for a specific project or measurement method. A given agency or laboratory would have only one quality assurance program plan, but would have a quality assurance project plan for each of its projects (group of projects using the same measurement methods; for example, a laboratory service group might develop a plan by analytical instrument since the service is provided to a number of projects).

Quality Audit - A systematic examination of the acts and decisions with respect to quality in order to independently verify or evaluate compliance to the operational requirements of the quality program or the specification or contract requirements of the product or service, and/or to evaluate the adequacy of a quality program.

Quality Control - The system of activities designed and implemented to provide a quality product.

Internal Quality Control - The routine activities and checks, such as periodic calibrations, duplicate analyses, use of spiked samples, included in normal internal procedures to control the accuracy and precision of a measurement process. (See Quality Control.)

External Quality Control - The activities which are performed on an occasional basis, usually initiated and performed by persons outside of normal routine operations, such as on-site system surveys, independent performance audits, interlaboratory comparisons, to assess the capability and performance of a measurement process.

Random Samples - Samples obtained in such a manner that all items or members of the lot, or population, have an equal chance of being selected in the sample. In air pollution monitoring the population is usually defined in terms of a group of time periods for which measurements are desired. For 24-h samplers, the population is usually considered as all of the 365 (or 366) 24-h calendar day periods in a year. For continuous monitors, the population is often considered as all of the hourly average values obtained (or which could have been obtained) during a particular period of time, usually a calendar year. For either 24-hour or continuous monitors, a single air pollution result from a site could be a sample of the conceptually infinite population of values that might have been obtained at the given site for all possible combinations of equipment, materials, personnel, and conditions, that could have existed at that site and time.

Representative Sample - A sample taken to represent a lot or population as accurately and precisely as possible. A representative sample may be either a completely random sample or a stratified sample depending upon the objective of the sampling and the conceptual population for a given situation.

Sample - A subset or group of objects or things selected from a larger set, called the "lot," or "population." The objects or

things may be physical such as specimens for testing or they may be data values representing physical samples. Unless otherwise specified, all samples are assumed to be randomly selected. Usually, information obtained from the samples is used to provide some indication or inference about the larger set. Samples rather than the population are examined usually for reasons of economy--the entire population under consideration is usually too large or too inaccessible to evaluate. In cases where destructive testing is performed, sampling is a must--otherwise the entire population would be consumed. In many situations, the population is conceptually infinite and therefore impossible to check or measure.

Standard Operating Procedure - (SOP) - A written document which details an operation, analysis or action whose mechanisms are thoroughly prescribed and which is commonly accepted as the method for performing certain routine or repetitive tasks.

Statistical Control Chart (Also Shewhart Control Chart) - A graphical chart with statistical control limits and plotted values (usually in chronological order) of some measured parameter for a series of samples. Use of the charts provides a visual display of the pattern of the data, enabling the early detection of time trends and shifts in level. For maximum usefulness in control, such charts should be plotted in a timely manner, that is, as soon as the data are available.

Stratified Sample (Stratified Random Sample) - A sample consisting of various portions that have been obtained from identified subparts or subcategories (strata) of the total lot, or population. Within each category or strata, the samples are taken randomly. The objective of taking stratified samples is to obtain a more representative sample than that which might otherwise be obtained by a completely random sampling. The idea of identifying the subcategories or strata is based on knowledge or suspicion of (or protection against) differences existing among

the strata for the characteristics of concern. The identification of the strata is based on knowledge of the structure of the population, which is known or suspected to have different relationships with the characteristic of the population under study. Opinion polls or surveys use stratified sampling to assure proportional representation of the various strata (e.g., geographic location, age group, sex, etc.). Stratified sampling is used in air monitoring to ensure representation of different geographical areas, different days of the week, and so forth.

System Audit - A systematic on-site qualitative review of facilities, equipment, training, procedures, recordkeeping, data validation, data management, and reporting aspects of a total (QA) system, (a) to arrive at a measure of capability of the measurement system to generate data of the required quality, and/or (b) to determine the extent of compliance of an operational QA system to the approved QA Project Plan.

### Statistics

Availability - The fraction or percentage of time that an item performs satisfactorily (in the reliability sense) relative to the total time the item is required to perform, taking into account its reliability and its maintainability, or the percentage of "up time" of an item or piece of equipment, as contrasted to its percentage of inoperative or "down time."

Comparability - A measure of the confidence with which one data set can be compared to another.

Completeness - The amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct normal operations, usually expressed as a percentage.

Confidence Coefficient - The chance or probability, usually expressed as a percentage, that a confidence interval has of including the population value. The confidence coefficients

usually associated with confidence intervals are 90, 95, and 99 percent. For a given sample size, the width of the confidence interval increases as the confidence coefficient increases.  
Confidence Interval - A value interval that has a designated probability (the confidence coefficient) of including some defined parameter of the population.

Confidence Limits - The outer boundaries of a confidence interval.

Error - The difference between an observed or measured value and the best obtainable estimate of its true value.

Maintainability - The probability that an item that has failed (in the reliability sense) can be restored (i.e., repaired or replaced) within a stated period of time.

Measures of Central Tendency - Measures of the tendency of values in a set of data to be centered at some location. Measures of central tendency are, for example, the median, the mode, the arithmetic mean, and the geometric mean.

Arithmetic Mean (Average) - The most commonly used measure of central tendency, commonly called the "average." Mathematically, it is the sum of all the values of a set divided by the number of values in the set.

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

Geometric Mean - Mathematically, the geometric mean  $\bar{X}_g$  can be expressed in two equivalent ways.

$$1) \quad \bar{X}_g = \left[ \prod_{i=1}^n X_i \right]^{\frac{1}{n}}$$

or in words, the  $n^{\text{th}}$  root of the product of all values in a set of  $n$  values.

$$2) \quad \bar{X}_g = \log^{-1} \left[ \frac{\sum_{i=1}^n \log X_i}{n} \right]$$

or in words, the antilogarithm of the arithmetic mean of the logarithms of all the values of a set of  $n$  values. (Note: the logarithms may be either natural or base 10, or any base for that matter, providing the operations are consistent, i.e., not mixed-base.) The geometric mean is generally used when the logarithms of a set of values are nearly normally (Gaussian) distributed, such as is the case for some pollution data.

Median - The middle value of a set of data when the set of data are ranked in increasing or decreasing order. If there are an even number of values in the set, the median is the arithmetic average of the two middle values.

Mode - The value or values occurring most frequently in a sample of data.

Measures of Dispersion or Variability - Measures of the differences, scatter or variability of values of a set of numbers. Commonly used measures of the dispersion or variability are the range, the standard deviation, the variance, and the coefficient of variation (or relative standard deviation).

Range - The difference between the maximum and minimum values of a set of values. When the number of values is small (i.e., 8 or less), the range is a relatively sensitive (efficient) measure of variability.

As the number of values increases above 8, the efficiency of the range (as an estimator of the variability) decreases rapidly. The range or difference between two paired values is of particular importance in air pollution measurements, since in many situations duplicate analyses or measurements are performed as a part of the quality assurance program.

Variance - Mathematically, the sample variance is the sum of squares of the differences between the individual values of a

set and the arithmetic average of the set, divided by one less than the number of values,

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1} .$$

For a finite population, the variance  $\sigma^2$  is the sum of squares of deviations from the arithmetic mean, divided by the number of values in the population.

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} ,$$

where  $\mu$  is the true arithmetic mean of the population.

Standard Deviation - For a sample, the standard deviation  $s$  is

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}} ,$$

the positive square root of the sample variance. For a finite population the standard deviation  $\sigma$  is

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$$

where  $\mu$  is the true arithmetic mean of the population and  $N$  is the number of values in the population. The property of the standard deviation that makes it most practically meaningful is that it is in the same units as the observed variable  $X$ .

Geometric Standard Deviation - In the analysis of measurements which are better approximated by a lognormal distribution, they are frequently summarized by the geometric mean ( $\bar{X}_g$ ) and geometric standard deviation ( $s_g$ ). These two statistics are calculated by first transforming the data by taking logs, obtaining the mean ( $\bar{X}$ ) and standard deviation( $s$ ) of the transformed data, and then calculating the antilogs of  $\bar{X}$  and  $s$  as indicated in the following equations:

$$\bar{x}_g = e \bar{X},$$

and

$$s_g = e^S.$$

Coefficient of Variation (Relative Standard Deviation) - A measure of precision calculated as the standard deviation of a set of values divided by the average. It is usually multiplied by 100 to be expressed as a percentage.

$$CV = RSD = \frac{s}{\bar{x}} \times 100 \text{ for a sample, or}$$

$$CV = RSD = \frac{\sigma}{\mu} \times 100 \text{ for a population.}$$

Examples of the computations for range, standard deviation, variance and relative standard deviation are presented in Appendix C.

Outlier - An extreme value that questionably belongs to the group of values with which it is associated. If the chance probability of its being a valid member of the group is very small, the questionable value is thereby "detected" and may be eliminated from the group based on further investigation of the data.

Random Error - Variations of repeated measurements that are random in nature and individually not predictable. The causes of random error are assumed to be indeterminate or nonassignable. The distribution of random errors is generally assumed to be normal (Gaussian).

Relative Error - An error expressed as a percentage of the true value or accepted reference value. All statements of precision or accuracy should indicate clearly whether they are expressed in absolute or relative sense. (This gets complicated when the absolute value is itself a percentage as is the case for many chemical analyses.)

Reliability (General) - The capability of an item or system to perform a required function under stated conditions for a stated period of time. (Specific) - The probability that an item will perform a required function under stated conditions for a stated period of time.

Statistical Control Chart Limits - The limits on control charts that have been derived by statistical analysis and are used as criteria for action, or for judging whether a set of data does or does not indicate lack of control.

Systematic Error - The condition of a consistent deviation of the results of a measurement process from the reference or known level. The cause for the deviation, or bias, may be known or unknown, but is considered "assignable." By assignable is meant that if the cause is unknown, it should be possible to determine the cause. See Bias.

#### Test Variability

Accuracy - The degree of agreement of a measurement,  $X$ , with an accepted reference or true value,  $T$ , usually expressed as the difference between the two values,  $X - T$ , or the difference as a percentage of the reference or true value,  $100(X-T)/T$ , and sometimes expressed as a ratio,  $X/T$ .

Bias - A systematic (consistent) error in test results. Bias can exist between test results and the true value (absolute bias, or lack of accuracy), or between results from different sources (relative bias). For example, if different laboratories analyze a homogeneous and stable blind sample, the relative biases among the laboratories would be measured by the differences existing among the results from the different laboratories. However, if the true value of the blind sample were known, the absolute bias or lack of accuracy from the true value would be known for each laboratory. See Systematic Error.

Precision - A measure of mutual agreement among individual measurements of the same property, usually under prescribed

similar conditions. Precision is most desirably expressed in terms of the standard deviation but can be expressed in terms of the variance, range, or other statistic. Various measures of precision exist depending upon the "prescribed similar conditions." (See Replicability, Repeatability, Reproducibility.)

Measures of precision must be qualified or explained in terms of possible sources of variability to make them most meaningful and useful. This is particularly true for repeatability. For example, the following tabulation reflects the requirements of the above definition:

Source of variability	Replicability	Repeatability	Reproducibility
Specimen (subsample)	Same or different	Same or different	Most likely different
Sample	Same	Same	Same
Analyst	Same	Same or different <sup>a</sup>	Different
Apparatus	Same	Same or different <sup>a</sup>	Different
Day	Same	Same or different <sup>a</sup>	Same or different
Laboratory	Same	Same	Different

<sup>a</sup>At least one of these must be different.

In the above tabulation, the essential requirement for repeatability is that the same sample must be analyzed by the same laboratory but under different conditions. The situation may be single analyst or multianalyst, single apparatus or multiapparatus, and single day or multiday, or any of the seven possible combinations involving at least one multifactor, each of which would result in different measures of precision. Also, for replicability, repeatability, and reproducibility, the situation may be single specimen or multispecimen, depending usually upon the physical limitations involved. For further detailed discussion, see ASTM Method E177-71.<sup>1</sup>

Dr. John Mandel<sup>2</sup> defines repeatability and reproducibility in the specific sense of an upper probability limit on differences between two test values. In the case of repeatability, the differences are those between two test values at the same laboratory, and in the case of reproducibility, the difference between two test values--one from one laboratory and the second from another laboratory. It is important that the distinction be made between precision measured as a standard deviation and precision expressed as an upper probability limit of differences between two values as both are frequently used. There is, however, a definite relationship between the two measures. For example, the upper 95% probability limit on differences between two values is 2.77 times the standard deviation. The preferred means of presenting the data would be to use the estimated standard deviations, thus minimizing the possibility of misinterpretation.

Relative Standard Deviation - See coefficient of variation.

Repeatability - The precision, usually expressed as a standard deviation, measuring the variability among results of measurements at different times of the same sample at the same laboratory. The unit of time should be specified, since within-day repeatability would be expected to be smaller than between-day repeatability.

Replicability - The precision, usually expressed as a standard deviation, measuring the variability among replicates.

Replicates - Repeated but independent determinations of the same sample, by the same analyst, at essentially the same time and same conditions. Care should be exercised in considering replicates of a portion of an analysis and replicates of a complete analysis. For example, duplicate titrations of the same digestion are not valid replicate analyses, although they may be valid replicate titrations. Replicates may be performed to any degree (e.g., duplicates, triplicates).

Reproducibility - The precision, usually expressed as a standard deviation, measuring the variability among results of measurements of the same sample at different laboratories.

Tolerance Limits - A particular type of confidence limit used frequently in quality control work where the limits apply to a percentage of the individual values of the population.

#### Testing or Measurement

Analytical Limit of Discrimination - A concentration above which one can, with relative certainty, ascribe the net result from any analysis to the atmospheric particulate and below which there is uncertainty in the result. One approach to determining a statistical limit is to use a one-sided tolerance limit for the analytical discrimination limit, that is a level (limit) below which a specified percentage (e.g., 99%) of blank filters analyses fall with a prescribed confidence (e.g., 95%). In addition, Reference 4 contains a detailed discussion of limits of detection.

Blank or Sample Blank - A sample of a carrying agent (gas, liquid, or solid) that is normally used to selectively capture a material of interest, and that is subjected to the usual analytical or measurement process to establish a zero baseline or background value, which is used to adjust or correct routine analytical results.

Analytical or Reagent Blank - A blank used as a baseline for the analytical portion of a method. For example, a blank consisting of a sample from a batch of absorbing solution used for normal samples, but processed through the analytical system only, and used to adjust or correct routine analytical results.

Dynamic Blank (or Field Blank) - A blank that is prepared, handled, and analyzed in the same manner as normal carrying agents except that it is not exposed to the material to be selectively captured. For example, an absorbing solution that

would be placed in bubbler tube, stoppered, transported to a monitoring site, left at the site for the normal period of sampling, returned to the laboratory, and analyzed.

Calibration - Establishment of a relationship between various calibration standards and the measurements of them obtained by a measurement system, or portions thereof. The levels of the calibration standards should bracket the range of levels for which actual measurements are to be made.

Dynamic Calibration - Calibration of a measurement system by use of calibration material having characteristics similar to the unknown material to be measured. For example, the use of a gas containing sulfur dioxide of known concentrations in an air mixture could be used to calibrate a sulfur dioxide bubbler system.

Static Calibration - The artificial generation of the response curve of an instrument or method by use of appropriate mechanical, optical, electrical, or chemical means. Often a static calibration checks only a portion of a measurement system. For example, a solution containing a known amount of sulfite compound would simulate an absorbing solution through which has been bubbled a gas containing a known amount of sulfur dioxide. Use of the solution would check out the analytical portion of the pararosaniline method, but would not check out the sampling and flow control parts of the bubbler system.

Certified Reference Material (CRM), Cylinder Gases - Gases prepared by gas vendors in quantities of at least 10 cylinders for which (1) the average concentration is within 1% of an available SRM, and (2) 2 cylinders are selected at random and audited by EPA.

Collaborative Tests (or Studies) - The evaluation of a new analytical method under actual working conditions through the participation of a number of typical or representative laboratories in analyzing portions of carefully prepared homogeneous samples.

Functional Analysis - A mathematical analysis that examines each aspect of the measurement system (sampling and analysis) in order to quantitate the effect of sources of error. A functional analysis is usually performed prior to a ruggedness test in order to determine those variables which should be studied experimentally.

Minimum Detectable Level (Limit of Detection) - The limit of detection for an analytical method is the minimum concentration of the constituent or species of interest which can be observed by the instrument and distinguished from instrument noise with a specified degree of probability. For example, one approach used is to make repeated measurements of the extractant liquid (trace metal analyses) and calculating the standard deviation of the results and hence the desired statistical tolerance limit for instrumental noise (e.g., an upper 99% limit at 95% confidence).

Proficiency Testing - Special series of planned tests to determine the ability of field technicians or laboratory analysts who normally perform routine analyses. The results may be used for comparison against established criteria, or for relative comparisons among the data from a group of technicians or analysts.

Ruggedness Testing - A special series of tests performed to determine the sensitivity (hopefully, to confirm the insensitivity) of a measurement system to variations of certain factors suspected of affecting the measurement system.

Spiked Sample - A normal sample of material (gas, solid, or liquid) to which is added a known amount of some substance of interest. The extent of the spiking is unknown to those analyzing the sample. Spiked samples are used to check on the performance of a routine analysis or the recovery efficiency of a method.

Standards in Naturally Occurring Matrix - Standards relating to the pollutant measurement portions of air pollution measurement systems may be categorized according to matrix, purity, or use.

Standards in a naturally occurring matrix include Standard Reference Materials and Standard Reference Samples.

Standard Reference Material (SRM) - A material produced in quantity, of which certain properties have been certified by the National Bureau of Standards (NBS) or other agencies to the extent possible to satisfy its intended use. The material should be in a matrix similar to actual samples to be measured by a measurement system or be used directly in preparing such a matrix. Intended uses include (1) standardization of solutions, (2) calibration of equipment, and (3) auditing the accuracy and precision of measurement systems.

Standard Reference Sample (SRS) - A carefully prepared material produced from or compared against an SRM (or other equally well characterized material) such that there is little loss of accuracy. The sample should have a matrix similar to actual samples used in the measurement system. These samples are intended for use primarily as reference standards (1) to determine the precision and accuracy of measurement systems, (2) to evaluate calibration standards, and (3) to evaluate quality control reference samples. They may be used "as is" or as a component of a calibration or quality control measurement system.

Examples: An NBS certified sulfur dioxide permeation device is an SRM. When used in conjunction with an air dilution device, the resulting gas becomes an SRS. An NBS certified nitric oxide gas is an SRM. When diluted with air, the resulting gas is an SRS.

#### Standards Depending upon "Purity" or Established Physical or Chemical Constants

Primary Standard - A material having a known property that is stable, that can be accurately measured or derived from established physical or chemical constants, and that is readily reproducible.

Secondary Standard - A material having a property that is calibrated against a primary standard.

Standards Based upon Usage

Calibration Standard - A standard used to quantitate the relationship between the output of a sensor and a property to be measured. Calibration standards should be traceable to Standard Reference Materials (SRM), Certified Reference Materials (CRM) or a primary standard.

Quality Control Reference Sample (or Working Standard) - A material used to assess the performance of a measurement or portions thereof. It is intended primarily for routine intra-laboratory use in maintaining control of accuracy and would be prepared from or traceable to a calibration standard.

Standardization - A physical or mathematical adjustment or correction of a measurement system to make the measurements conform to predetermined values. The adjustments or corrections are usually based on a single-point calibration level.

Traceability - A documented chain of comparisons connecting a working standard (in as few steps as is practical) to a national (or international) standard such as a standard maintained by NBS.

A.2 REFERENCES

1. Use of Terms Precision and Accuracy as Applied to Measurement of a Property of a Material, ASTM Method E177-71.
2. Mandel, John. Repeatability and Reproducibility, Materials Research Standards, Vol. 11, No. 8, pp. 8-16, August 1971.
3. Assessment of Arsenic Losses During Ashing: A Comparison of Two Methods Applied to Atmospheric Particulates, Journal of the Air Pollution Control Association, Vol. 28, No. 11, pp. 1134-1136, November 1978.
4. Currie, L. A. Limits for Qualitative Selection and Quantitative Determination. Analytical Chemistry, Vol. 40, No. 3, pp. 586-593, March 1968.

### A.3 BIBLIOGRAPHY

The following sources were reviewed in the development of Appendix A. The definitions have been prepared with particular meaning and application to air pollution measurement systems, and may not agree in some details with definitions cited in the bibliography. The purpose of Appendix A is to promote a more common understanding and usage of these basic terms for air pollution measurement activities.

1. Akland, G. G., Environmental Protection Agency, Quality Assurance and Environmental Monitoring Laboratory, Research Triangle Park, North Carolina, document prepared for Symposium of Trace Element Analysis, May 1973.
2. The Analyst, Vol. 90, No. 1070, p. 252. Analytical Methods Committee, report prepared by the Analytical Standards Sub-Committee (Sodium Carbonate as a Primary Standard in Acid-Base Titrimetry).
3. Environmental Protection Agency, 999-AP-15 (or 999-WP-15), Environmental Measurements Symposium, Valid Data and Logical Interpretation.
4. Environmental Protection Agency, APTD-0736, Field Operations Guide for Automatic Air Monitoring Equipment.
5. Environmental Protection Agency, APTD-1132, Quality Control Practices in Processing Air Pollution Samples.
6. ANSI/ASQC Standard A1-1978, Definitions, Symbols, Formulas, and Tables for Control Charts.
7. ANSI/ASQC Standard A2-1978, Terms, Symbols, and Definitions for Acceptance Sampling.
8. ANSI/ASQC Standard A3-1978, Quality Systems Terminology.
9. ANSI/ASQC Standard Z1.15-1980, Generic Guidelines for Quality Systems.
10. ASTM, Glossary of ASTM Definitions ASTM Committee E-8 on Nomenclature and Definitions, 2nd Ed., 1973, ASTM, 1916 Race Street, Philadelphia, Pennsylvania 19103.
11. ASTM Standard Recommended Practice E117-71. Use of the Terms Precision and Accuracy as Applied to Measurement of a Property of a Material.

12. ASTM Standard Recommended Practice E180-67. Developing Precision Data on ASTM Methods for Analysis and Testing of Industrial Chemicals.
13. Duncan, A.J., "Quality Control and Industrial Statistics," 3rd Ed., 1965, Richard D. Irwin Inc., Homewood, Illinois.
14. "Glossary of Terms Used in Quality Control," European Organization for Quality Control, Rotterdam, Netherlands, 1972, 3rd Ed.
15. Federal Register, Vol. 36, No. 84, April 30, 1971, "National Primary and Secondary Ambient Air Quality Standards."
16. Federal Register, "Ambient Air Monitoring Equivalent and Reference Methods," October 12, 1973, Vol. 38, No. 197, Part II.
17. Feigenbaum, A.V., Total Quality Control, Engineering and Management. McGraw-Hill (1961).
18. Department of Health, Education and Welfare, Public Health Service, National Institute for Occupational Safety and Health, Division of Laboratories and Criteria Development, Cincinnati, Ohio 45202.
19. National Institute for Occupational Safety and Health, Cincinnati, Ohio 45202, Industrial Hygiene Service Laboratory Quality Control Manual, Technical Report No. 78 (Draft).



APPENDIX B  
NOMENCLATURE

This appendix contains a list of the symbols which are used throughout the Appendices.

- a,b - intercept, slope of best fit linear equation by the method of least squares,  $Y = a + bX$ .
- $A_1$  - factor for computing the control chart for  $\bar{X}$  given  $\bar{s}$  and  $\bar{X}$ , i.e.,  $UCL_{\bar{X}} = \bar{X} + A_1\bar{s}$ ,  $LCL_{\bar{X}} = \bar{X} - A_1\bar{s}$ .
- $A_2$  - factor used in constructing control chart for  $\bar{X}$ , given  $\bar{R}$  and  $\bar{X}$ , i.e.,  $UCL_{\bar{X}} = \bar{X} + A_2\bar{R}$ ,  $LCL_{\bar{X}} = \bar{X} - A_2\bar{R}$ .
- A - availability (only in Appendix L).
- b' - slope of best fit line through the origin.
- CV (or RSD) - coefficient of variation (relative standard deviation) of the sample =  $100 s/\bar{X}$ .
- CV' (or RSD') - coefficient of variation (or relative standard deviation) of the population =  $100 \sigma/\mu$ .
- c - acceptance number for a single sample plan when sampling by attributes; i.e., if d is less than or equal to c, the lot is accepted.
- D - downtime (only in Appendix L).
- D - (signed) difference between two measurements =  $X_1 - X_2$ .
- d - number of defectives observed in a sample of n measurements.

d - signed % difference between measurements =

$$\frac{X_1 - X_2}{(X_1 + X_2)/2} \times 100 \text{ (i.e., the signed dif-}$$

ference divided by the average). In some applications, the absolute % difference is used instead of the signed % difference.

d - the allowable relative margin of error in % (Appendix E only).

DF (or df) - degrees of freedom.

$d_2$  - factor to estimate  $\sigma$  given the mean range, values are given in Table C.2.

$D_3, D_4, D_5, D_6$  - factors used in constructing control chart for R, i.e.,  $UCL_R = D_4 \bar{R}$  and  $LCL_R = D_3 \bar{R}$ ,  $UWL_R = D_6 \bar{R}$ ,  $LWL_R = D_5 \bar{R}$ .

$f_i$  - frequency of the ith group, cell, or interval.

k - number of samples or sets of data averaged.

L - number of laboratories.

$\log_b X$  - logarithm of X using base b, normally  $b = 10$  or  $b = e = 2.7183$  (natural base).

$\overline{\log X}$  - the mean of the logarithms of the X's in a sample

$$\text{i.e., } \overline{\log X} = \frac{\sum \log X}{n} .$$

M - maintainability (see Appendix L for more details).

$M_R$  - repair time for maintenance.

$M_D$  - diagnostic time for maintenance.

n - number in the sample or number of items in test.

$n-1$  - number of degrees of freedom associated with estimate  $s^2$  of  $\sigma^2$  based on a sample of size  $n$ .

$N$  - population size, if finite, or lot size in acceptance sampling problems.

$n!$  -  $n$  factorial =  $n(n-1)(n-2) \dots 2 \cdot 1$  (e.g.,  $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ ).

$\binom{n}{r}$ ,  $C_r^n$  (or  $n^C_r$ ) - the number of combinations of  $n$  items taken  $r$  at a time.

$\binom{n}{r} = \frac{n!}{r!(n-r)!}$  (e.g.,  $C_0^5 = \frac{5!}{0!5!} = 1$ ,  $0! = 1$  by

definition,  $C_2^5 = \binom{5}{2} = \frac{5!}{2!3!} = 10$ ).

$p$  - fraction of defects (or defective measurements) in the sample.

$\bar{p}$  - fraction of defects in the population of measurements sampled.

$P$  - number of intervals for grouped data.

$P(X)$  - the probability of the event  $X$ .

$P(a < X < b)$  - the probability that  $X$  falls between  $a$  and  $b$ ,  $a < b$  ( $a$  less than  $b$ ).

$r_{11}$  - test statistic for the largest value, a possible outlier.

$R$  - range of a data set or sample,  $R =$  largest value less the smallest value of a set of measurements.

$\bar{R}$  - average range for  $k$  samples of the same sample size  $n$ ,  $\bar{R} = \Sigma R/k$ .

$\Sigma R$  - sum of the ranges for  $k$  samples.

RSD - relative standard deviation (see CV).

$s^2$  - sample variance =  $\Sigma(X - \bar{X})^2 / (n-1)$ .

$s$  - sample standard deviation =  $\sqrt{s^2}$ .

$\bar{s}$  - average  $s$  for  $k$  groups of data =  $\Sigma s / k$ .

$s_g$  - geometric standard deviation = antilog  $\{s(\log X)\}$ .

$s_D$  - standard deviation of a set of differences of two paired values.

$s(\log X)$  - the standard deviation of the logarithms of the  $X$ 's in a sample.

$s_{Y|X}$  - standard deviation of the observed response from the fitted line (or curve in general),  $s$  is frequently used if it is clearly understood from the context that  $s_{Y|X}$  is the standard deviation of the discussion.

$t$  - time.

$t_{n-1, \alpha}$  - tabulated  $t$  value for specified degrees of freedom (DF =  $n-1$ ) and for which the fraction  $\alpha$  of the absolute values of  $t$  exceed  $t_{n-1, \alpha}$ .

$T_1$  - test statistic for the smallest value  $X_1$ , a suspect outlier =  $(\bar{X} - X_1) / s$ .

$T_n$  - test statistic for the largest value  $X_n$  in a sample, a suspect outlier,  $T_n = (X_n - \bar{X}) / s$ .

$U$  - uptime (see Appendix L for more details).

$UCL$  ( $LCL$ ) - upper (lower) control limit.

$UWL$  ( $LWL$ ) - upper (lower) warning limit.

$UCL_R$ ,  $UCL_{\bar{X}}$  - upper control limits for  $R$ ,  $\bar{X}$ , respectively, the subscript denotes the variables used in the chart.

$w$  - width of confidence interval (used only in Appendix C).

- $X_i$  -  $i$ th measurement (also used as the  $i$ th smallest measurement of a set of measurements arranged in ascending order, see Appendix F).
- $\bar{X}$  - sample mean =  $\Sigma X/n = (X_1 + X_2 + \dots + X_n)/n$ .
- $\hat{X}$  - median of a sample.
- $\bar{X}_g$  - geometric mean of a sample of measurements =  
antilog ( $\overline{\log X}$ ) =  $\log^{-1}(\overline{\log X})$ .
- $X$  - random variable or measured value.
- $X_n$  - largest value in a sample of size  $n$ , see Appendix F.
- $X_1$  - smallest value in a sample of size  $n$ , see Appendix F.
- $\bar{\bar{X}}$  - grand average for  $k$  data sets or  $k$  samples =  $\Sigma \bar{X}/k$  (if all samples are of equal size).
- $X$  - independent or controlled variable such as the concentration of  $\text{NO}_2$  (Appendix J).
- $X_p$  - predicted value of  $X$  for an observed value of  $Y$  (e.g., an analyzer reading  $Y$ ).
- $Y$  - dependent variable or response variable.
- $\bar{Y}$  - mean of the  $Y$ 's for the sample =  $\Sigma Y/n$ .
- $Y_p$  - predicted mean response.
- $Z$  (or  $u$ ) - standard normal variable =  $(X - \mu)/\sigma$  where  $\mu$  and  $\sigma$  are the mean and standard deviation of the normal distributed variable  $X$ .

#### GREEK NOTATION

Letters of the Greek alphabet are commonly used in statistical texts and literature to denote the parameters of the conceptual population of measurements. These are typically unknown

values to be estimated on the basis of a sample of measurements taken from the conceptual population. The estimates are denoted by letters of the English alphabet, for example,

$\bar{X}$  is an estimate of  $\mu$

$s$  is an estimate of  $\sigma$ .

Occasionally the Greek letter with a caret or "hat" is used to denote an estimate (e.g.,  $\hat{\mu}$  or  $\hat{\sigma}$  are estimates of  $\mu$  and  $\sigma$ , respectively). These will be used when considering more than one estimate or an estimate different from the standard one.

$\alpha, \beta$  - (alpha, beta) parameters (intercept and slope) of the true linear relationship between the response variable  $Y$  and the independent variable  $X$  (i.e.,  $Y = \alpha + \beta X + \epsilon$ ).

$\delta$  - allowable (absolute) margin of error in estimating the mean  $\mu$ .

$\epsilon$  (epsilon) - random error of measurement associated with the response variable  $Y$ .

$\mu$  (mu) - mean of the population of measurements

$$= \frac{\sum_{i=1}^N X_i}{N} \text{ if a finite population.}$$

$\mu_g$  - geometric mean of population = antilog ( $\mu\{\log X\}$ ).

$\prod X$  (pi) -  $\prod_{i=1}^n X_i = X_1 \cdot X_2 \cdots X_n$ , that is, the product of the  $X_i$ 's of the sample.

$\sigma^2$  - population variance =  $\sum (X - \mu)^2 / N$  if  $N$  is finite.  
(sigma squared)

$\sigma$  - population standard deviation =  $\sqrt{\sigma^2}$ .

$\sigma_X = \sigma\{X\}$  - standard deviation of the variable  $X$ ; often it is necessary to discuss more than one measure of standard deviation; in this case the variable is denoted by a subscript or in braces  $\{ \}$ .

- $\sigma_{\bar{X}}$  - standard deviation of the sample mean of n independent measurements =  $\sigma/\sqrt{n}$ , (i.e., the population standard deviation divided by  $\sqrt{n}$ ).
- $\sigma_g$  - geometric standard deviation of population = antilog ( $\sigma\{\log X\}$ ).
- $\sigma_r^2$  - variance among replicates within a day and within a laboratory.
- $\sigma_d^2$  - variance among days within a laboratory.
- $\sigma_l^2$  - variance among laboratories.
- $\hat{\sigma}$  ( $\sigma$  "hat") - estimate of  $\sigma$ , such as based on R (range) (i.e.,  $\hat{\sigma} = R/d_2$  or  $\bar{R}/d_2$ ).
- $\chi_i$  (chi) - mid-value of the ith interval (used only in Appendix C).
- $\chi_{P,DF}^2$  - tabulated value of  $\chi^2$  distribution for specified (chi-squared) DF and 100 Pth percentile.
- $\nu$  (nu) - degrees of freedom for students t (alternative notation of DF).



APPENDIX C  
COMPUTATIONAL EXAMPLES  
OF  
DESCRIPTIVE STATISTICS

C.1 INTRODUCTION

Statistical methods dealing with procedures for the collection, analysis and interpretation of data can be grouped into two classes: (1) those used to summarize a body of data to make them more meaningful and (2) those used to make generalizations about a large body of possible data from a small body of available data. These two classes can be referred to as descriptive statistics and statistics used to make inferences. This appendix is devoted to a discussion of the more frequently used descriptive statistics.

C.2 BASIC CONCEPTS

Data to which statistical methods may be applied may be either measurements made on individual elements or counts of the number of elements that possess specific attributes. The totality of measurements of all individual elements or the count of the number of elements with all possible attributes is referred to as the population (or aggregate). The population may consist of a very large number of elements such as the one-hour concentrations of SO<sub>2</sub> for several years or a small number, for example, the number of sources in a county that emit more than 10,000 tons of SO<sub>2</sub> in a year.

A statistical sample is a collection of elements selected in some way from the population. Depending upon the way in which the sample is selected, it may or may not provide data that can be used to make useful inferences about the population.

When a value such as the arithmetic mean is calculated from all possible data for the population, it is referred to as a parameter and identified by the Greek letter  $\mu$  (mu). The arithmetic mean for a sample is referred to as a statistic and identified by the symbol  $\bar{X}$ , and called the average. Similarly, the standard deviation for a sample is  $s$  and that for a population is  $\sigma$  (sigma).

C.3 DATA PRESENTATION

C.3.1 Frequency Distribution

Data recorded in the order in which they are received are referred to as raw data. Using raw data, it is most difficult to perceive the information contained in a large set of data. A very useful method of presenting a body of data is to group the data into intervals and then count the number of data values that fall within each interval. Table C.1 shows the frequency distribution of 162 measurements of the ambient concentration of suspended particulates made over a period of time at a sampling station.

TABLE C.1. CONCENTRATION OF SUSPENDED PARTICULATES AT STATION X

Concentration, $\mu\text{g}/\text{m}^3$	Number of values
25 < X < 50	3
50 < X < 75	10
75 < X < 100	14
100 < X < 125	24
125 < X < 150	33
150 < X < 175	31
175 < X < 200	18
200 < X < 225	19
225 < X < 250	8
250 < X < 275	2
	<u>162</u>

These same data may be presented graphically as a bar chart or histogram as shown in Figure C.1, where the height of the bar is the frequency of values in the corresponding interval.

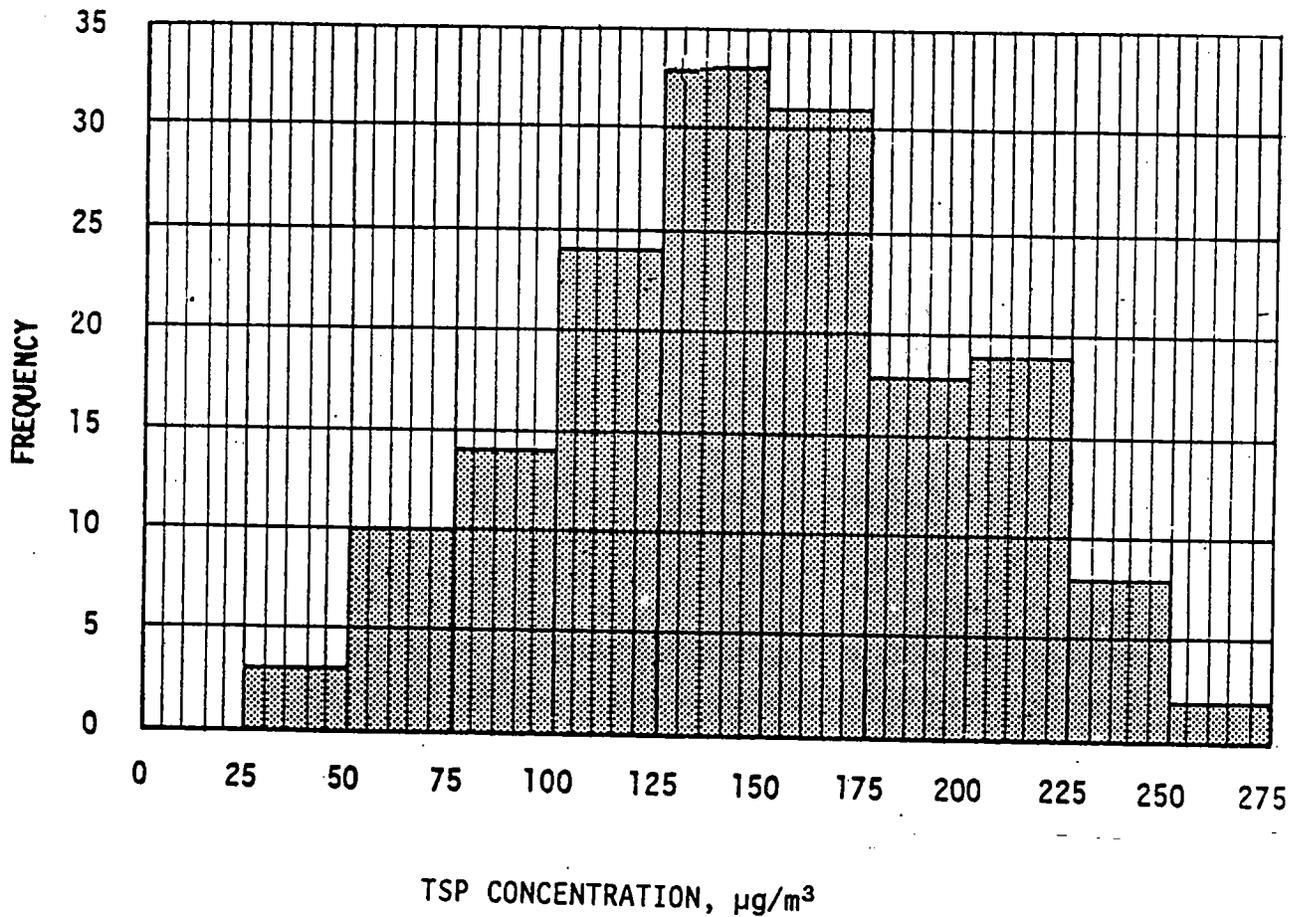


Figure C.1. Frequency distribution of concentrations of TSP.

From Figure C.1, it can now be seen that the ambient concentrations range between 25 and 275  $\mu\text{g}/\text{m}^3$ . Also, it can be seen that the most frequently occurring concentrations are in the range of 125-175  $\mu\text{g}/\text{m}^3$ .

### C.3.2 Measures of Central Tendency

It is often desirable to select a single value to represent a body of data. Such values are referred to as measures of central tendency (or location parameters). Included as measures of central tendency are such parameters as the arithmetic mean, the median, the geometric mean, the mode, and the harmonic mean. Several of the more frequently used location parameters are discussed below.

C.3.2.1 Arithmetic Mean - Perhaps the most widely used location parameter is the arithmetic mean. If the frequency distribution for a set of data is nearly symmetrical (as is the case for the data shown in Figure C.1), the arithmetic mean may be the most representative location parameter.

The equations for the arithmetic mean of a finite population and a sample selected from the population are given by,

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i, \quad (1)$$

$$\bar{X} = \frac{1}{n} \sum X, \quad (2)$$

where

- $\mu$  = population mean,
- $\bar{X}$  = sample average,
- $N$  = number of elements in population,
- $n$  = number of elements in sample,
- $X$  = the individual data values, and

$$\sum X = \sum_{i=1}^n X_i = X_1 + X_2 + \cdots + X_n. \quad (3)$$

Throughout the text, the notation  $\sum_{i=1}^n X_i$  will be replaced by  $\sum X$ , with the summation of  $X$  for all values in the sample being implied. Very simply, the average is the sum of the individual values divided by the number of values.

Two examples are presented below to illustrate the computation of the average for ungrouped data and for grouped data presented as a frequency distribution.

Example C.1 Measurements of the ambient concentration of suspended particulates were made by 12 laboratories as part of a collaborative testing program.<sup>1</sup> Because this is assumed to be a sample, the sample average  $\bar{X}$  is calculated.

Laboratory	X, $\mu\text{g}/\text{m}^3$
1	138
2	125
3	128
4	126
5	127
6	128
7	128
8	108
9	126
10	125
11	125
12	131

$$\Sigma X = 1515$$

$$\begin{aligned} \bar{X} &= \frac{1}{12} \Sigma X \\ &= \frac{1}{12} (1515) \\ &= 126.2 \mu\text{g}/\text{m}^3. \end{aligned}$$

Example C.2 The computation of the average for data presented as a frequency distribution is illustrated using the data presented in Ref. 1.

Concentration $\mu\text{g}/\text{m}^3$	Mid-value X	No. of values f	fX
25 < X < 50	37.5	3	112.5
50 < X < 75	62.5	10	625.0
75 < X < 100	87.5	14	1225.0
100 < X < 125	112.5	24	2700.0
125 < X < 150	137.5	33	4537.5
150 < X < 175	162.5	31	5037.5
175 < X < 200	187.5	18	3375.0
200 < X < 225	212.5	19	4037.5
225 < X < 250	237.5	8	1900.0
250 < X < 275	262.5	2	525.0
		n = 162	$\Sigma fX = 24075.0$

The equation for the average for data in a frequency distribution is

$$\bar{X} = \frac{1}{n} \sum_{i=1}^P f_i \chi_i = \frac{1}{n} \sum f \chi \quad (4)$$

where

f = the number of values in the indicated interval,

$\chi$  = mid-value of the indicated interval,

P = number of intervals,

$$\bar{X} = \frac{1}{162} (24075.0) = 148.6 \mu\text{g}/\text{m}^3.$$

C.3.2.2 Median - There are situations in which the average may not be the best location parameter to represent a set of data. Consider the following situation in which 6 measurements of the ambient concentration of suspended particulates were 65, 90, 70, 82, 96 and 485  $\mu\text{g}/\text{m}^3$ , respectively. The average of this set of data, 148  $\mu\text{g}/\text{m}^3$ , is not truly representative of the typical concentration of suspended particulates. In a situation like this, the median (i.e., the "middle" value) may be a more meaningful location parameter.

To determine the median for ungrouped data, it is first necessary to arrange the data in order of magnitude such that  $X_1 \leq X_2 \leq \dots \leq X_n$ . When the number of data values is even, the median value is equivalent to the half-way point between the two middle values, (i.e., between the  $\frac{n}{2}$  and the  $\frac{n+2}{2}$  values); and where the number of data values is odd, it is the middle value. With the 6 values above, the median concentration is half-way between the  $\frac{6}{2} = 3\text{rd}$  and the  $\frac{6+2}{2} = 4\text{th}$  values.

$X_1$	65
$X_2$	70
$X_3$	82
$X_4$	90
$X_5$	96
$X_6$	485

$$\text{Median} = \frac{82 + 90}{2} = 86 \mu\text{g}/\text{m}^3.$$

For this set of data,  $86 \mu\text{g}/\text{m}^3$  is a more representative value than  $148 \mu\text{g}/\text{m}^3$ .

C.3.2.3 Geometric Mean - With most air quality measurement data, the frequency distributions tend to be nonsymmetrical. The frequency distribution of the 162 values in Example C.4 is "skewed" to the right as shown in Figure C.2.

Under conditions to be discussed in Appendix D in which a frequency distribution is skewed as shown in Figure C.2 the geometric mean may be the most representative location parameter. The geometric mean is defined as the  $n$ th root of the product of  $n$  values.

$$\text{where } \bar{X}_g = (\prod X)^{1/n} \tag{5}$$

$\bar{X}_g$  = geometric mean  
 $X$  = the individual data values

and  $\prod X = X_1 \cdot X_2 \cdots X_n$ .

The geometric mean may also be defined as the antilog of the average of the logarithms of the data values;

$$\bar{X}_g = \text{antilog}_b \left[ \frac{1}{n} \sum \log_b X \right]. \tag{6}$$

Either common logarithms ( $\log_{10}$ ) or natural logarithms ( $\log_e$ ) can be used to calculate the geometric mean.

Example C.3 Calculate the geometric mean for the data presented in Example C.1.

X	log X
	10
138	2.1398
125	2.0969
128	2.1071
126	2.1003
127	2.1037
128	2.1071
128	2.1071
108	2.0334
126	2.1003
125	2.0969
125	2.0969
131	2.1173

$$\begin{aligned} \bar{X}_g &= \text{antilog} \left[ \frac{1}{n} \sum \log_{10} X \right] \\ &= \text{antilog} \frac{25.2068}{12} \\ &= \text{antilog} 2.10057 \\ &= 126.1 \mu\text{g}/\text{m}^3 \end{aligned}$$

$$\sum \log_{10} X = 25.2068$$

The average for this sample is  $126.2 \mu\text{g}/\text{m}^3$

Example C.4 Calculate the geometric mean for the following data.

Concentration, $\mu\text{g}/\text{m}^3$	Mid-value		Frequency	$f \log_{10} X$
	$X$	$\log_{10} X$	$f$	
25 < X < 50	37.5	1.57403	3	4.72209
50 < X < 75	62.5	1.79588	18	32.32584
75 < X < 100	87.5	1.94201	37	71.85437
100 < X < 125	112.5	2.05115	31	63.58565
125 < X < 150	137.5	2.13830	27	57.73410
150 < X < 175	162.5	2.21085	14	30.95190
175 < X < 200	187.5	2.27300	17	38.64100
200 < X < 225	212.5	2.32736	8	18.61888
225 < X < 250	237.5	2.37566	5	11.87830
250 < X < 275	262.5	2.41913	2	4.83826
			$\Sigma f \log_{10} X = 335.15039$	

$$\begin{aligned} \bar{X}_g &= \text{antilog}_{10} \left[ \frac{1}{n} \Sigma f \log_{10} X \right] \\ &= \text{antilog}_{10} \left[ \frac{1}{162} (335.15039) \right] \\ &= \text{antilog}_{10} 2.06883 \\ &= 117.2 \mu\text{g}/\text{m}^3 \end{aligned}$$

The average for this set of data is 126.2  $\mu\text{g}/\text{m}^3$ .

### C.3.3 Measures of Dispersion

In addition to presenting a location parameter that is representative of a set of data, it is generally important to know the amount of scatter or dispersion of the individual data values. The more widely used measures of dispersion include the range, the standard deviation, and the variance. In the case of air pollution data the geometric standard deviation is also used.

C.3.3.1 Range - The range is defined as the difference between the largest and the smallest values in a set of data. For Example C.1, the range is 138-108 = 30  $\mu\text{g}/\text{m}^3$ . By definition the range makes use of only two values out of a set of data. As such, the range is very sensitive to extreme values. The range

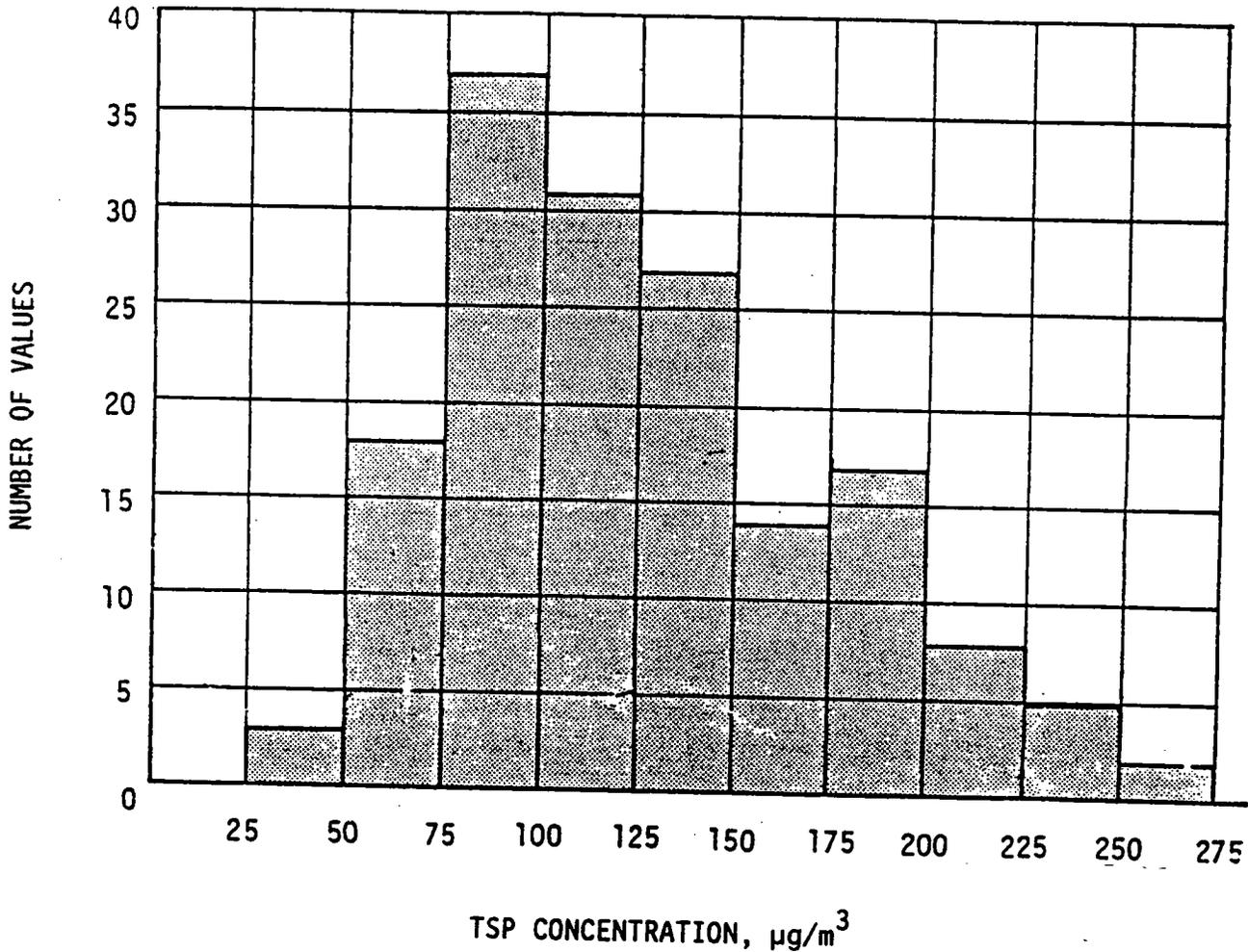


Figure C.2. Frequency distribution of measurements of the concentration of suspended particulates, data for Example C.4.

has relatively good efficiency compared to the standard deviation when the sample size is small ( $2 \leq n \leq 8$ ). With larger sample sizes, the standard deviation is considerably more efficient than the range and is preferred.

C.3.3.2 Variance and Standard Deviation - The variance of a finite population is defined as the sum of the squares of the deviations of the individual values from the mean divided by the number of values in the population. The variance of the finite population is given by

$$\sigma^2 = \frac{1}{N} \sum (X - \mu)^2. \quad (7)$$

The variance of a sample is defined as  $s^2$  and given by

$$s^2 = \frac{1}{n-1} \sum (X - \bar{X})^2 \quad (8)$$

The divisor  $n-1$  is used, rather than  $n$ , so that the value of  $s^2$ , the sample variance, is an unbiased estimate of  $\sigma^2$ , the population variance (i.e., on the average the sample variance will be equal to the population variance).

For computational purposes, with ungrouped data, the equation for  $s^2$  can also be written as

$$s^2 = \frac{\sum X^2 - (\sum X)^2/n}{n-1} \quad (9)$$

The equation for  $s^2$  in this form allows one to accumulate the sum  $(\sum X)$  and the sums of squares  $(\sum X^2)$  very easily on a desk or mini-portable calculator. Currently, many calculators are programmed to obtain the sample average and variance and the use of Equation (9) is not necessary.

Because of the process of squaring, the units of the variance are actually the square of the units of measurement. In order to obtain a statistic with the same units as the original data, the standard deviation, which is defined as the positive square root of the variance, is used more frequently.

Example C.5 The variance and standard deviation of the data presented in Example C.1 are calculated below.

Laboratory	X	
1	138	$s^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}$ $s^2 = \frac{191,777 - \frac{(1515)^2}{12}}{11}$ $s^2 = 46.20$ $s = 6.8 \mu\text{g}/\text{m}^3$
2	125	
3	128	
4	126	
5	127	
6	128	
7	128	
8	108	
9	126	
10	125	
11	125	
12	131	
$\overline{\sum X}$	$\overline{1515}$	
$\overline{\sum X^2}$	$\overline{191,777}$	

C.3.3.3 Geometric Standard Deviation - When the data are skewed to the right as illustrated in Figure C.2 and when they are distributed according to the lognormal frequency distribution as described in Appendix D, then the standard geometric deviation  $s_g$  is used as a measure of dispersion instead of the standard deviation  $s$ . This is defined as

$$\begin{aligned}
 s_g &= \text{antilog of the standard deviation of the logarithms of the measurements, that is,} \\
 &= \text{antilog } \left[ \frac{\sum (\log X - \overline{\log X})^2}{(n-1)} \right]^{1/2} \\
 &= \text{antilog } \left[ \frac{\sum (\log X) - \frac{(\sum \log X)^2}{n}}{n-1} \right]^{1/2}
 \end{aligned}
 \tag{10}$$

where the logarithms are taken to any convenient base, preferably the common base 10. Thus the standard deviation of the logarithms of the measurements is computed in the usual manner after transforming each value  $X$  (or center value in the case of a frequency distribution such as in Example C.4) to its corresponding log value or  $\log X$ .

Example C.6 Use the data of Example C.3 to calculate the geometric standard deviation,  $s_g$ . For these data,

$$\sum \log_{10} X = 25.2075$$

$$\sum (\log_{10} X)^2 = 52.9579$$

$$s(\log X) = 0.02433$$

$$s_g = \text{antilog}(0.02433) = 1.058$$

$$\bar{X}_g = 126.1 \text{ as obtained in Example C.3.}$$

C.3.3.4 Use of Range to Estimate the Standard Deviation - The range is frequently used to estimate the standard deviation, particularly in control chart applications where the simplicity in calculating the range is desired. Table C.2 gives factors  $d_2$  for dividing the range by (i.e.,  $R/d_2$ ) to estimate the standard deviation. Thus for  $n = 12$ , the range would be divided by  $d_2 = 3.258$  to estimate  $\sigma$ . For the Example C.1, this estimate  $\hat{\sigma}$  of  $\sigma$

would be  $30/3.258 = 9.21$ , slightly larger than the estimate  $s = 6.8$  of  $\sigma$  which is based on all of the data. This relationship between the standard deviation and the range is based on the assumption that the measurements are normally distributed and it can also be used as a quick check on the calculated standard deviation  $s$ , that is, the standard deviation is approximated by

$$\hat{\sigma} = \frac{\text{Range of } n \text{ measurements}}{d_2} \cong s. \quad (11)$$

It is recommended that a quick rule of thumb be used in all cases to check the calculated  $s$ , for example, one might use the following rough approximation.

For n between,	divide the range by $d_2$ to estimate s
$2 \leq n \leq 5$	$d_2 \cong 2$
$6 \leq n \leq 15$	$d_2 \cong 3$
$16 \leq n \leq 50$	$d_2 \cong 4$
$51 \leq n \leq 200$	$d_2 \cong 5$
$n > 200$	$d_2 \cong 6$

TABLE C.2. FACTORS ASSOCIATED WITH THE RANGE  
 $\hat{\sigma} = \text{range}/d_2$

n	$d_2$
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078
11	3.173
12	3.258

C.3.3.5 Relative Standard Deviation - The relative standard deviation (RSD) is a frequently used measure of dispersion in the air pollution literature (the RSD is also referred to as the coefficient of variation (CV) in the statistical literature). The RSD of the sample is computed by

$$\text{RSD} = \frac{s}{\bar{X}} \times 100, \quad (12)$$

the ratio of the standard deviation to the average and multiplied by 100 to convert to a percent of the average. After some experience with data in a particular field of measurements, typical values of the RSD are determined, for example, 5% to 20% is a reasonable range of values in many of the measurement processes used in measuring pollutant concentrations.

There is one caution that must be kept in mind when stating dispersion in terms of  $s$  (absolute terms) or in terms of the RSD. The use of the RSD can and does imply that the absolute standard deviation changes with the value  $\bar{X}$ , whereas stating  $s$  implies that it does not change with  $\bar{X}$  unless explicitly indicated otherwise. In practice for air pollution measurements the standard deviation does tend to depend on the level of the measurement, but not necessarily with constant proportion over all  $\bar{X}$ . In many practical problems the RSD is essentially constant over the range of interest, and in this case it is the most useful measure of variation. For example, if one were to make replicate analyses of a filter media for lead concentration, the standard deviation of the replicate analyses would tend to increase as the concentration of lead increased and the RSD would remain essentially constant.

C.3.3.6 Absolute % Difference (or Relative Range) - This measure of dispersion is a very useful measure of variation for the special case in which  $n = 2$  (i.e., the sample size is two). It is defined by

$$d^* = \frac{|X_1 - X_2|}{(X_1 + X_2)/2} \times 100 \quad (13)$$

that is, the range of two measurements divided by their average and multiplied by 100 to express the result as a percentage. In the case of two observations the range is directly related to the standard deviation (i.e.,  $R = \sqrt{2} s$ ) thus the percent difference is  $\sqrt{2}$  times the RSD. Because of the ease of computation and the frequency of using repeat or duplicate measurements in air pollution applications, the percent difference is a useful measure of variation. Furthermore, control charts may be applied to this measure when the error of measurement increases with the concentration, say, in preference to an ordinary range chart which assumes that the measurement variation remains constant for the period of application of a control chart and for all levels of concentration.

**C.3.3.7 Signed % Difference** - In some applications (e.g., in the presentation of performance audit results) the signed percent difference is used instead of the absolute value in order to emphasize the direction (+ or -) of the measurement bias. That is, the difference between the routinely measured response (Y) and the audited response (X) is divided by the audited response (presumed to be correct); then multiplied by 100 to convert to a %. Hence,

$$d = 100 \frac{Y-X}{X} \quad (14)$$

#### C.4 NUMBER OF PLACES TO BE RETAINED IN COMPUTATION AND PRESENTATION OF DATA

The following working rule is recommended in the ASTM manual<sup>2</sup> in carrying out computations of  $\bar{X}$ ,  $s$ , and confidence limits based on a set of  $n$  observed values of a variable quantity:

---

\*  $d$  might also be defined for sample sizes other than  $n = 2$  (e.g.,  $d = R/\bar{X}$ ).  
 $|X_1 - X_2|$  is the absolute value of the difference of two measurements.

"In all intermediate operations on the set of  $n$  observed values, such as adding, subtracting, multiplying, dividing, squaring, extracting square root, retain the equivalent of at least two more places of figures than in the single observed values. For example, if observed values are read or determined to the nearest 1 lb., carry numbers to the nearest 0.01 lb. in the computations; if observed values are read or determined to the nearest 10 lb., carry numbers to the nearest 0.1 lb. in the computations.

Rejecting places of figures should be done after computations are completed, in order to keep the final results substantially free from computation errors. In rejecting places of figures the actual rounding off procedure should be carried out as follows:<sup>2</sup>

1. When the figure next beyond the last figure or place to be retained is less than 5, do not change the figure in the last place retained.

2. When the figure next beyond the last figure or place to be retained is greater than 5, increase by 1 the figure in the last place retained.

3. When the figure next beyond the last place to be retained is 5, and

a. there are no figures, or only zeros, beyond this 5, increase by 1 the figure in the last place to be retained if it is odd, leave the figure unchanged, if it is even, but

b. if the 5 next beyond the figure in the last place to be retained is followed by any figures other than zero, increase by 1 the figure in the last place retained whether it is odd or even.

For example, if, in the following numbers, the places of figures in parenthesis are to be rejected:

39 4(49) becomes 39 400  
39 4(50) becomes 39 400,  
39 4(51) becomes 39 500, and  
39 5(50) becomes 39 600.

The number of places of figures to be retained in presentation depends on what use is to be made of the results. No general rule, therefore, can safely be laid down. The following working rule has, however, been found generally satisfactory in presenting the results of testing in technical investigations and development work:

1. For averages, retain the number of places shown below:

Single values obtained to the nearest	Number of observed values, n		
0.1, 1, 10, etc., units		2-20	21-200
0.2, 2, 20, etc., units	less than 4	4-40	41-400
0.5, 5, 50 etc., units	<u>less than 10</u>	<u>10-100</u>	<u>101-1000</u>
Number of places of figures to be retained in the average.....	} Same number of places as in single values	1 more place than in single values	2 more places than in single values

2. For standard deviations, retain three places of figures.

3. If confidence limits are presented, retain the same places of figures as are retained for the average.

For example, if n = 10, and if observed values were obtained to the nearest 1 lb., present averages and confidence "limits" to the nearest 0.1 lb., and present the standard deviation to three places of figures."

C.5 SUMMARY

In summary, given a set of measurements, they can be summarized by the following quantities or statistics.

	<u>Location</u>
Average $\bar{X} = \frac{\sum X}{n}$	
$\bar{X}_g = \text{Antilog}_b \left[ \frac{\sum \log_b X}{n} \right]$	
Median = $\tilde{X} =$	{ Middle value (n odd) { Average of two middle values (n even)

Dispersion

---

$$s^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}$$

$$s = \sqrt{s^2}$$

$$s_g = \text{antilog} \left[ \frac{\sum (\log X)^2 - \frac{(\sum \log X)^2}{n}}{n-1} \right]^{1/2}$$

R = largest less the smallest value

$\hat{\sigma}$  = R/d<sub>2</sub> (See Table C.2)

$$\text{RSD (or CV)} = \frac{s \times 100}{\bar{X}}$$

$$d = \frac{100 |X_1 - X_2|}{(X_1 + X_2)/2}, \text{ absolute \% difference}$$

or

$$d = 100 (Y - X)/X, \text{ signed \% difference}$$


---

The use of particular statistics will depend on assumptions concerning the frequency distribution of the measurements as described in the following section. However, the (arithmetic) average  $\bar{X}$  and estimated standard deviation  $s$  have properties which make them generally useful as measures of the central location and of dispersion of the data and thus as estimates of these same characteristics or parameters of the population. Additional references<sup>3,4,5,6</sup> which are recommended on the subject of descriptive statistics are given at the end of this appendix.

#### C.6 REFERENCES

1. McKee, H. C., Childers, R. E., and Saena, O. S., Jr. Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method), Contract CPA 70-40, SWRI Project 21-2811, June 1971.

2. ASTM Manual on Quality Control of Materials, Prepared by ASTM Committee E-11 on Quality Control of Materials, Special Technical Publication 15-C, American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA, January 1951. The successor to publication 15-C is the following: ASTM Manual on Presentation of Data and Control Chart Analysis, Special Technical Publication 15-D, (1976).
3. Bauer, E. L., A Statistical Manual For Chemists, Second Edition, Academic Press, New York, 1971.
4. Dixon, W. J., Introduction to Statistical Analysis, McGraw-Hill Book Company, Inc., New York, 1951.
5. Daniel, W. W., Biostatistics: A Foundation for Analysis in the Health Sciences, John Wiley and Sons, Inc., New York, 1974.
6. Hald, A., Statistical Theory with Engineering Applications, John Wiley and Sons, Inc., New York, 1952.

## APPENDIX D

### PROBABILITY DISTRIBUTIONS

#### D.1 INTRODUCTION

In Appendix C a frequency distribution was used as a means of presenting a large quantity of data in a meaningful way. If the number of data values become quite large and if the width of the intervals of the frequency distribution is allowed to tend toward zero, the midpoints of the tops of the bar graph will tend to describe a smooth curve. Three major types of continuous frequency distributions are used to describe air pollution data, namely, the normal, lognormal, and Weibull distributions. All of these will be briefly described in this appendix. There are also applications in which the measurement of interest can take on a limited number of distinct values, as for example the number of times during a year when the 3 h air quality standard for SO<sub>2</sub> was exceeded. In this case the number of such occurrences among  $n$  measurements can only be 0, 1, 2, ...,  $n$ . The relative frequency of each such occurrence would be an example of a discrete frequency distribution. The discrete frequency distribution will not be discussed in this appendix; the reader is referred to other texts on the subject.<sup>1,2</sup>

#### D.2 NORMAL DISTRIBUTION

The most widely used continuous frequency distribution is the normal (or Gaussian) function. The normal distribution is described by two parameters, the mean ( $\mu$ ) and the standard deviation ( $\sigma$ ). Referring to Figure D.1, one can observe that changing the value of  $\sigma$  causes the curve to become more spread out or more peaked. Changing the value of  $\mu$  merely shifts the curve to the right or left on the horizontal axis.

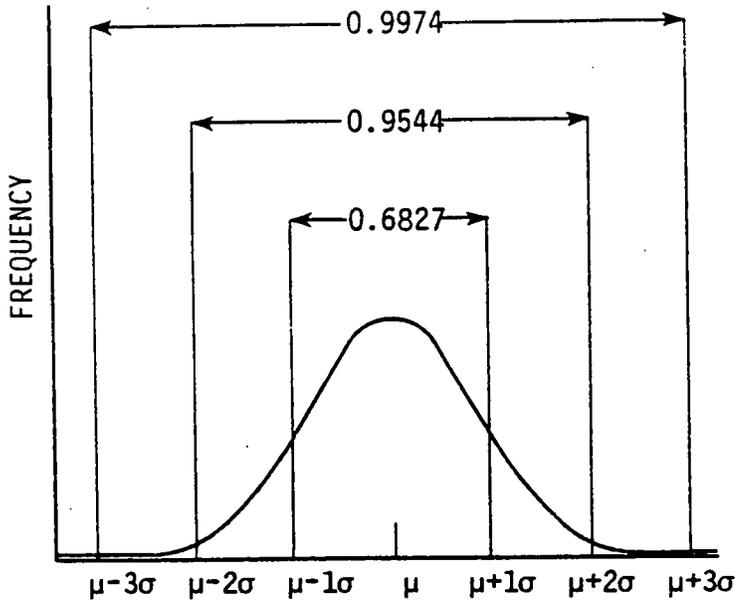


Figure D.1. Area under normal curve between specified limits.

The area under the normal curve between two specified ordinates can be used to express the probability that a measurement from a normal population would fall in the interval bounded by the two ordinates. Probabilities for selected intervals specified in units of standard deviation are also shown in Figure D.1. Thus, it can be seen that the probability is 0.9544 that a value  $X$  selected at random from the standard normal population (i.e.,  $\mu = 0$  and  $\sigma = 1$ ) will fall in the interval between  $-2$  and  $+2$ . This statement of probability can also be written in the form

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544$$

for a normal population with mean  $\mu$  and standard deviation  $\sigma$ . It should also be obvious from Figure D.1 that

$$P(X \geq \mu) = 0.5$$

Since the normal curve in a particular application depends upon the values of  $\mu$  and  $\sigma$ , there are an infinite number of possible normal curves. Standard tables of probabilities for the normal curve are constructed for the special case where

$\mu = 0$  and  $\sigma = 1$ . To use such tables it is necessary to rescale the variable of measurement by the following transformation

$$Z = \frac{X - \mu}{\sigma} . \quad (1)$$

The quantity  $Z$  is usually referred to as a standard normal variate or the "normal deviate". The probabilities associated with positive values of  $Z$  are presented in Table D.1. This table gives the probability that a value selected at random from the standard normal distribution will fall in the interval  $Z = 0$  to  $Z = Z_1$ .

Example D.1 Suppose the measurement of concentration of a certain pollutant is normally distributed with  $\mu = 75$  and  $\sigma = 25 \mu\text{g}/\text{m}^3$ . What is the probability that a measurement made at random will be in the interval between 56 and 118?

$$\begin{aligned} Z_1 &= \frac{X - \mu}{\sigma} \\ &= \frac{118 - 75}{25} \\ &= 1.72. \end{aligned}$$

That is,  $Z_1$  is 1.72 standard deviations larger than the mean value.

$$\begin{aligned} Z_2 &= \frac{56 - 75}{25} \\ &= -0.76. \end{aligned}$$

From Table D.1 the probability for the interval bounded by  $Z_1 = 1.72$  is 0.4573 and for  $Z_2 = -0.76$  is 0.2764 (the negative sign indicates that the  $Z_2$  lies to the left of the mean). Thus the probability statement for this example is

$$\begin{aligned} P(56 \leq X \leq 118) &= 0.4573 + 0.2764 \\ &= 0.7337. \end{aligned}$$

Thus there is a 73% chance that the measured value will fall between 56 and 118 given that the measurements are normally distributed with  $\mu = 75$  and  $\sigma = 25$ .

TABLE D.1<sup>a</sup>

CUMULATIVE NORMAL FREQUENCY DISTRIBUTION (Area under the standard normal curve from 0 to Z)										
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000									

<sup>a</sup>Reproduced with permission from NBS Handbook 91, Experimental Statistics.

Example D.2 For the distribution in Example D.1, what is the probability that a value will lie in the interval between 78 and 96?

$$Z_1 = \frac{96 - 75}{25}$$

$$= 0.84$$

$$Z_2 = \frac{78 - 75}{25}$$

$$= 0.12$$

Referring to Table D.1 to obtain the probabilities for  $Z_1$  and  $Z_2$ , the following can be written

$$P(78 \leq X \leq 96) = 0.2995 - 0.0478 \\ = 0.2517.$$

The normal distribution is not necessarily the preferred distribution for approximating the distribution of pollutant concentrations, and thus other distributions are described in the following sections. However, quality control/quality assurance data are typically approximated by the normal distribution (e.g., the difference between measurements of split samples by two laboratories or repeated measurements of a working standard).

### D.3 USE OF PROBABILITY GRAPH PAPER

There are several statistical procedures for checking whether data may be considered to be normally, lognormally, or Weibull distributed. One of the most common and useful procedures in practice is to use the corresponding probability graph paper, on which the cumulative frequency function of a sample of  $n$  observations will be approximately a straight line if the data may be considered to be a random sample from the corresponding distribution. An example is given herein to illustrate how to obtain the cumulative frequency function and to plot it on the graph paper. The data used are those given in Example C.2. They are repeated here for convenience.

Example D.3 Using the data in Example C.2, obtain the cumulative frequency distribution of the concentration of suspended particulates at station X.

The cumulative frequencies are obtained by cumulating or adding the frequencies in Example C.2 to obtain the total frequency of observed concentrations at or below a particular value, in this case the upper limit of the corresponding class interval. The relative frequency is then expressed as a percentage of  $n$  by dividing the cumulative frequencies by  $n$  and multiplying by 100.<sup>2</sup> These relative cumulative frequencies are plotted as the ordinates of the points and the abscissae are the corresponding concentrations. There are discussions in the literature which recommend plotting one of the following:

1.  $\frac{\text{cumulative frequency (cf)}}{n+1} \times 100,$
2.  $\frac{cf - 1/2}{n} \times 100,$  (2)
3.  $\frac{cf - 3/8}{n+(1/4)} \times 100.$

The use of a particular method depends on the particular use to be made of the data.<sup>3</sup> For small samples Method 1 is recommended for cases in which it is desired to draw inferences concerning the extreme values.

In Example C.2, the cumulative frequency distribution follows approximately a straight line on normal probability paper and thus one might assume normality for practical applications. This does not imply that the data are normally distributed, in fact, there is considerable evidence that data on concentration of total suspended particulates tend to follow a lognormal distribution. Lognormal probability paper would be the same as that in Figure D.2 except that the concentration scale would be in equal steps of the logarithms (i.e., a log scale).

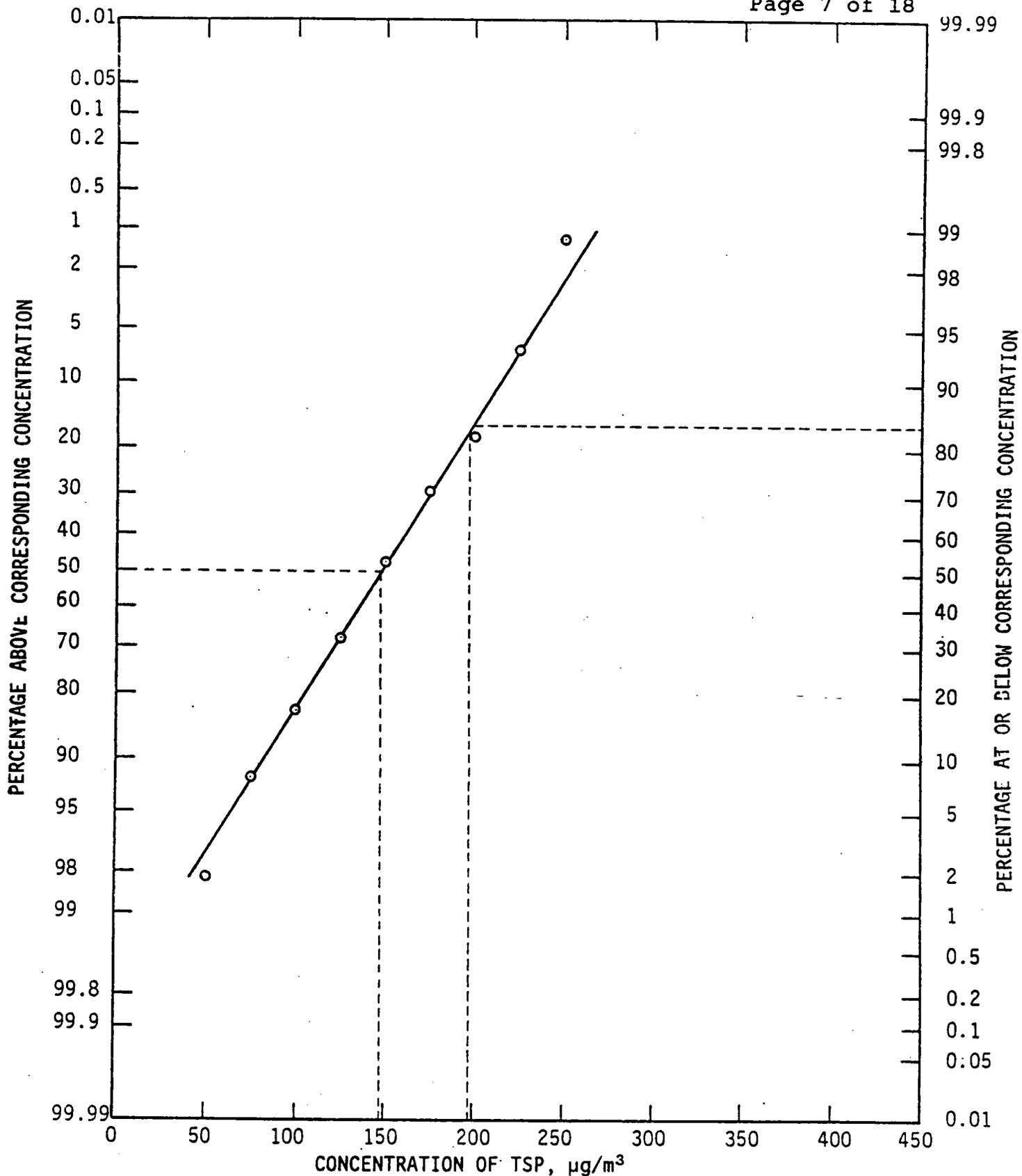


Figure D.2. Cumulative frequency distribution of concentration of TSP ( $\mu\text{g}/\text{m}^3$ ).

Concentration X	Cumulative frequency (cf) = number of values less than or equal to X	Relative cumulative frequency, %
25	0	0
50	3	1.8
75	13	8.0
100	27	16.7
125	51	31.5
150	84	51.8
175	115	71.0
200	133	82.1
225	152	93.8
250	160	98.8
275	162	100.0

Using a straight line fit to the data (an eye-fitted line in this case), it is possible to estimate the mean and standard deviation of the data in the sample as follows:

1. An estimate of the mean is obtained by reading the concentration corresponding to the 50th percentile (median value) as shown in the figure, in this case  $147 \mu\text{g}/\text{m}^3$ . Actually  $\bar{X} = 149 \mu\text{g}/\text{m}^3$  from Example C.2.

2. An estimate of the standard deviation is obtained by reading the concentration corresponding to the 84th percentile and subtracting from this the concentration at the mean or 50th percentile, in this case about  $197 - 147$  or about  $50 \mu\text{g}/\text{m}^3$  is an estimate of  $\sigma$ . From Table D.1 the area under the standard normal curve between the mean and one standard deviation above the mean,  $Z = 1$ , is 0.34; or approximately 84% of the area lies below the value  $\mu + \sigma$  for any normal curve. The sample standard deviation for these data is  $49.8 \mu\text{g}/\text{m}^3$ .

These two graphical estimates compare very well with the computed values, in fact, much more closely than one expects in practice. The degree of closeness depends, of course, on how easily one can fit a line to the data as shown in Figure D.2.

#### D.4 LOGNORMAL DISTRIBUTION<sup>4</sup>

Frequency distributions of measurements of ambient concentrations of air contaminants have been studied extensively in recent years. Most investigators agree that such distributions are not necessarily normal but they tend to disagree somewhat as to which distribution best describes such data. In the case of concentrations of total suspended particulate, for example, the logarithm of the daily measurement does tend to be nearly normally distributed. In this situation the distribution of  $Y = \text{Log } X$  is described by the mean  $\mu_Y$  and standard deviation  $\sigma_Y$ . In order to obtain values with units consistent with the measured data, the antilogs of the results are used. Thus, a variable which is lognormally distributed is usually described in terms of the geometric mean ( $\mu_g$ ) and the geometric standard deviation ( $\sigma_g$ ) where

$$\mu_g = \text{antilog}(\mu_Y) \quad (3)$$

$$\sigma_g = \text{antilog}(\sigma_Y). \quad (4)$$

**Example D.4** Suppose that the logarithm of the concentration of total suspended particulates is normally distributed with mean  $\mu_Y = 1.8751$  ( $\mu_g = 75$ ) and the standard deviation  $\sigma_Y = 0.1$  ( $\sigma_g = 1.26$ ). What is the probability that a measurement made at random will fall between 65 and 85  $\mu\text{g}/\text{m}^3$ ?

To answer this question, it is first necessary to obtain  $\log_{10} 65 = 1.8129$  and  $\log_{10} 85 = 1.9294$ . Hence the probability that a measurement  $X$  falls between 65 and 85  $\mu\text{g}/\text{m}^3$  is equal to the probability that the  $Y = \log X$  falls between 1.8129 and 1.9294, given that the mean  $\mu_Y = 1.8751$  and  $\sigma_Y = 0.1$  in log units. Then calculate  $Z_1$  and  $Z_2$  as in examples D.1 and D.2 to obtain

$$Z_1 = \frac{1.9294 - 1.8751}{0.1} = 0.543$$

and

$$Z_2 = \frac{1.8129 - 1.8751}{0.1} = -0.622$$

For these values, the areas for the standard normal curve from Table D.1 are approximately 0.233 (by interpolation) and 0.206. Hence, the probability that a measurement at random falls between 65 and 85 is given by the sum of these two values or about 0.44. See Figure D.3 for a graphical explanation of the above steps.

Example D.5 Use the data of Example C.4 and plot the sample distribution function on lognormal probability paper.

The data of Example C.4 are repeated here with additional calculations needed for plotting the cumulative frequencies on lognormal probability paper (Figure D.4).

Concentration interval, $\mu\text{g TSP}/\text{m}^3$	Frequency (f)	Cumulative frequency (f)	cf/100, %
25 < X < 50	3	3	1.9
50 < X < 75	18	21	13
75 < X < 100	37	58	35.8
100 < X < 125	31	89	54.9
125 < X < 150	27	116	71.6
150 < X < 175	14	130	80.2
175 < X < 200	17	147	90.7
200 < X < 225	8	155	95.7
225 < X < 250	5	160	98.8
250 < X < 275	2	162	100.0

Note that the cumulative frequency in % is plotted versus the upper value of the concentration interval (e.g., 90.7% of the values fall below 200  $\mu\text{g TSP}/\text{m}^3$ ). This probability paper is so constructed that if the data are truly lognormally distributed, then the plot will be a straight line. These data are closely approximated by a straight line as shown.

#### D.5 WEIBULL DISTRIBUTION<sup>5,6</sup>

Another distribution which has received extensive application in the analysis of air pollution data, particularly hourly averages of ozone,  $\text{NO}_2$ , and  $\text{CO}$ , is the Weibull distribution. In

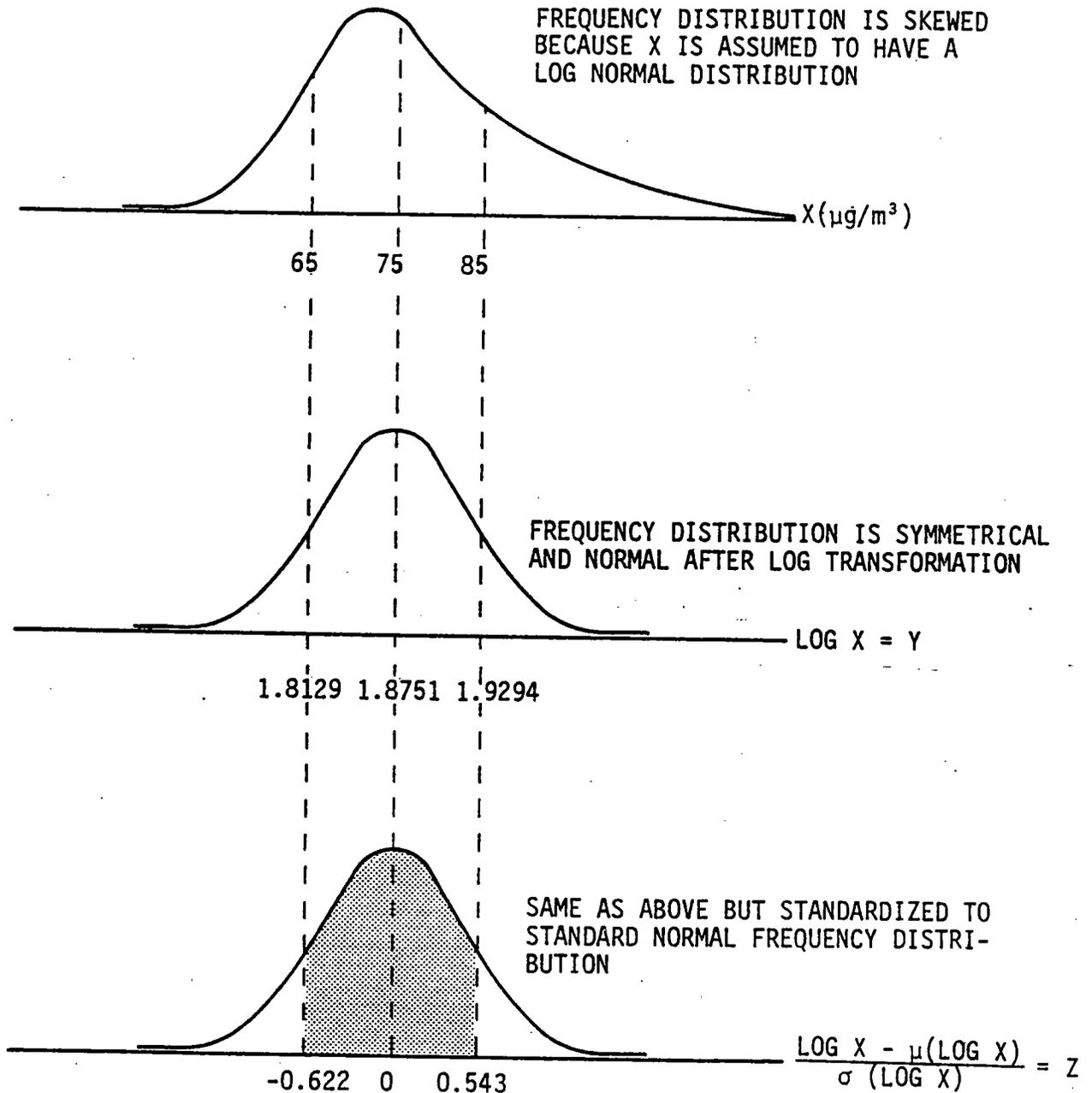


Figure D.3. Illustration of computation of Example D.5.

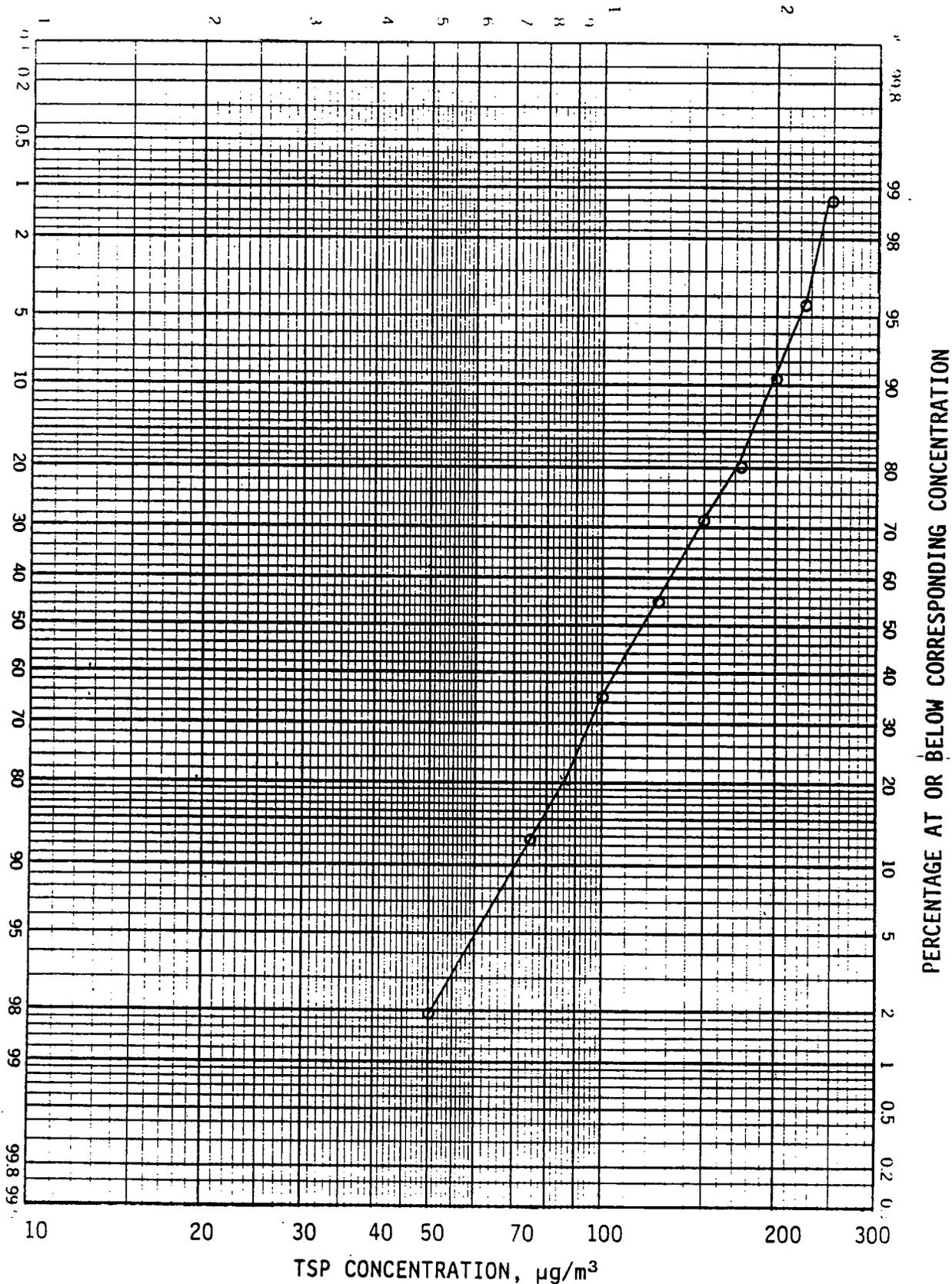


Figure D.4. Cumulative frequency of TSP concentrations on lognormal graph paper (Data of Example D.5).

1951 Walodde Weibull<sup>6</sup> suggested the Weibull distribution to describe experimental data. The probability that a Weibull variable is less than or equal to X ( $\leq X$ ) is given by

$$F(X) = 1 - \exp [-(X/\delta)^k]. \quad (5)$$

The two parameters  $\delta$  and  $k$  are the scale factor and shape parameters, respectively. The parameter  $k$  determines the degree and direction of the curvature of the frequency curve. When  $k = 1$ , the Weibull is equivalent to the exponential distribution (a distribution frequently used in the reliability literature). For  $k > 1$  the distribution is "heavy-tailed" and for  $k < 1$ , "light-tailed."<sup>4</sup> (A distribution is "heavy-tailed" when  $G(X) = 1 - F(X)$  is concave upward at the tail when plotted on semilog paper; a distribution is "light-tailed" when concave downward. The Weibull can fit either heavy-or light-tailed distributions, whereas the lognormal distribution is always heavy-tailed.)

Methods of fitting the Weibull distribution to data are given in References 4 and 5. An example taken from Reference 5 is given herein.

**Example D.6** The following cumulative frequency table (actually 100 minus the cumulative frequency %) is based on ozone data for the year 1975 at a site in Memphis, Tennessee. The data are 1 hour averages and in units of ppb. For example, we read from the following table that 11.4% of the concentrations exceed 50 ppb and that there are 260 values between 50 and 55 ppb.

These upper tail data are then plotted on Weibull graph paper to determine the adequacy of fit of this distribution to these data. In Reference 5 these data are also plotted on lognormal graph paper and the fit was not as good as that by the Weibull.<sup>4,5</sup>

From the plot in Figure D.5 it is obvious that the data points are fitted very well by a straight line and hence the Weibull approximation is an adequate approximation to at least the upper 80% of the data. For example, one infers from the

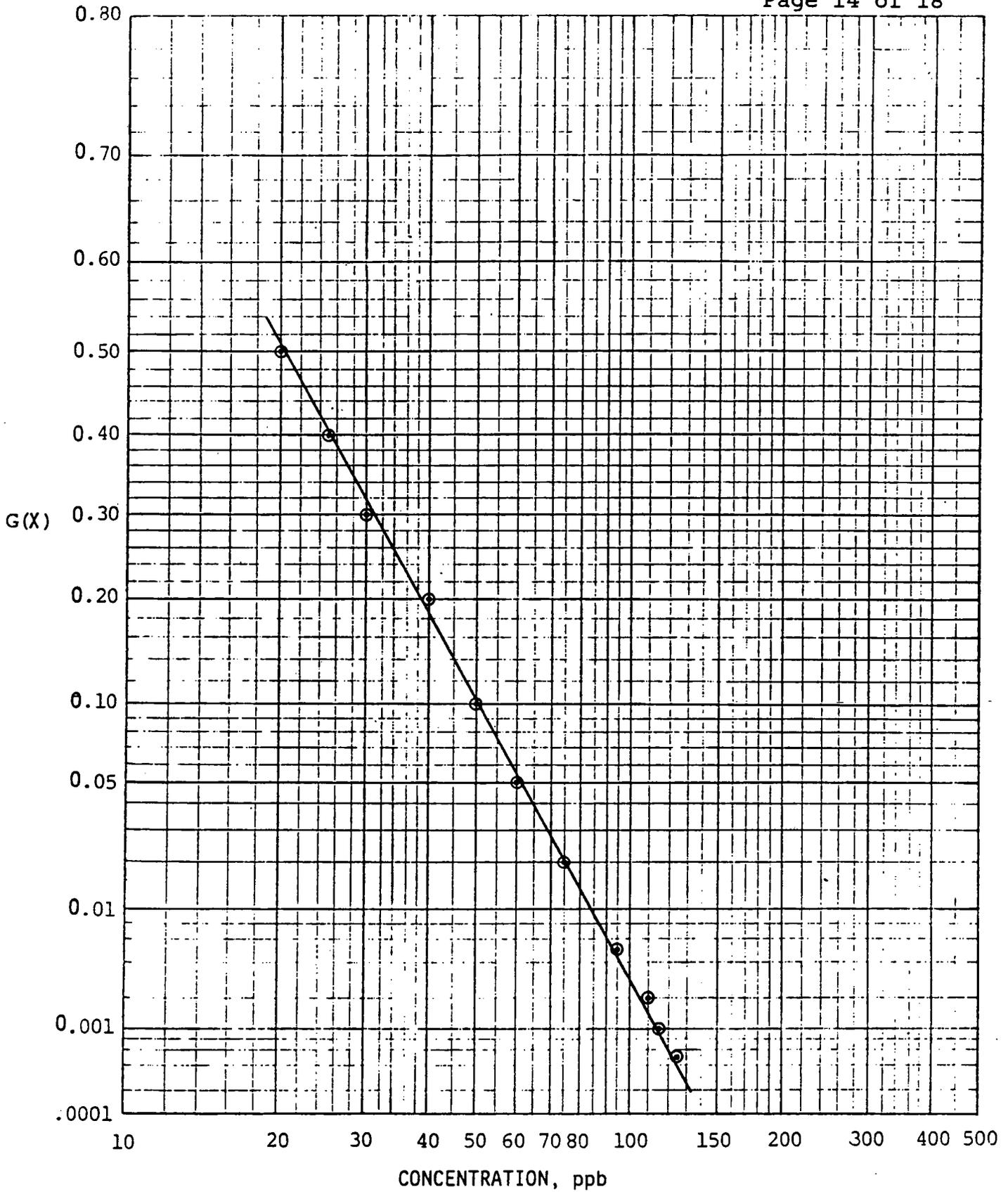


Figure D.5. Cumulative frequency distribution for data of Example D.6 on Weibull graph paper.

TABLE D.2. CUMULATIVE FREQUENCY DISTRIBUTION FOR 1 HOUR AVERAGE  
OZONE CONCENTRATIONS

Concentration, ppb	Frequency	Cumulative frequency, %
135	1	0.0126
130	1	0.0251
125	3	0.0628
120	2	0.0879
115	8	0.1883
110	6	0.2636
105	6	0.3389
100	10	0.4645
95	9	0.5775
90	17	0.7909
85	21	1.0545
80	36	1.5064
75	77	2.4730
70	69	3.3392
65	101	4.6071
60	148	6.4650
55	137	8.1848
50	260	11.4487
45	263	14.7502
40	439	20.2611
35	517	26.7512
30	705	35.6013
25	804	45.6942
20	880	56.7411
15	886	67.9890
10	993	80.4544
6	1	80.4670
5	567	87.5847
2	1	87.5973
0	988	100.0000

plot that approximately 10% of the concentrations exceed 50 ppb, 1% exceed 85 ppb, and 0.1% exceed 115 ppb.

#### D.6 DISTRIBUTION OF SAMPLE MEANS (NORMAL POPULATION)

Quite often the process of sampling is used to estimate the mean ( $\mu$ ) of the population. The sample mean or average ( $\bar{X}$ ) is used as an estimate of the population mean ( $\mu$ ).<sup>2</sup> A major result from statistical theory is that almost regardless of the shape of the frequency distribution of the original population, the

frequency distribution of  $\bar{X}$  in repeated samples of size  $n$  tends to become normal as  $n$  increases. The standard normal distribution can be used to determine probabilities related to the average by the following equation

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (6)$$

Example D.7 Suppose a sample of  $n = 25$  measurements of concentration of air pollutants are made on a population which is normally distributed with mean 60 and standard deviation 15. What is the probability that the average  $\bar{X}$  will lie between 55 and 65?

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z_1 = \frac{65 - 60}{\frac{15}{\sqrt{25}}} = 1.67.$$

(It should be noted that Greek letters (e.g.,  $\mu$ ,  $\sigma$ ) are generally reserved for parameters of the population and English letters, or Greek letters with ^'s (e.g.,  $\bar{X}$ ,  $s$ ,  $\hat{\sigma}$ ) for statistics of the sample).

$$Z_2 = \frac{55 - 60}{\frac{15}{\sqrt{25}}} = -1.67.$$

From Table D.1 the probability associated with  $Z = 1.67$  is 0.4525. Therefore,

$$\begin{aligned} P(55 \leq \bar{X} \leq 65) &= 0.4525 + 0.4525 \\ &= 0.9050. \end{aligned}$$

Example D.8 Using the population mean ( $\mu$ ), standard deviation ( $\sigma$ ), and sample size given in Example D.7, what is the probability that the average  $\bar{X}$  would be equal to or greater than 68?

$$Z = \frac{68 - 60}{\frac{15}{\sqrt{25}}}$$

$$= 2.67.$$

$$P(\bar{X} \geq 68) = 0.5000 - 0.4962 \\ = 0.0038.$$

Example D.9 Suppose that the average of two measurements of a standard reference sample is used as means of checking on the measurement process by quality control as described in Appendix H. Suppose that the mean and standard deviation of the concentration of the standard reference sample are  $\mu = 75$  and  $\sigma = 10 \mu\text{g}/\text{m}^3$ , respectively. Between what two values will 95% of the averages of samples of size two ( $n = 2$ ) fall?

The standard deviation of the average of a sample of two is given by  $\sigma/\sqrt{n} = 10/\sqrt{2} = 7.1 \mu\text{g}/\text{m}^3$ . The mean of the averages is  $75 \mu\text{g}/\text{m}^3$ ; this does not change with sample size. From Table D.1, 47.5% of the area under the curve falls between  $Z = 0$  and  $Z = 1.96$ . Hence, the values are determined by

$$\mu \pm Z \sigma/\sqrt{n} \tag{7} \\ = 75 \pm 1.96 (10/\sqrt{2})$$

or

$$61.1 \text{ and } 88.9 \mu\text{g}/\text{m}^3.$$

This is the process by which quality control limits for averages are determined. For simplicity, they are usually taken to be  $2\sigma$  or  $3\sigma$  limits, corresponding to areas of 0.9544 and 0.9974 as given in Figure D.1. In this case the limits for averages based upon  $n$  observations each are determined by, for example

$$\mu \pm 2(\sigma/\sqrt{n}) \tag{8}$$

or

$$\mu \pm 3 (\sigma/\sqrt{n}).$$

The  $2\sigma$  limits are referred to as warning limits and the  $3\sigma$  as the control limits. As the sample size  $n$  increases the limits become more narrow. In practical applications in air

pollution measurements the sample sizes,  $n = 1$  or  $2$ , are most common. It should be noted that in the use of  $n = 1$  or  $2$  measurements, the distribution assumption is critical.

#### D.7 REFERENCES

1. Juran, J.M., Quality Control Handbook, Third Edition, McGraw-Hill Book Co., Inc., New York, 1974.
2. Dixon, W.J., and Massey, F.J., Introduction to Statistical Analysis, McGraw-Hill Book Co., Inc., New York, 1951.
3. Hald, A., Statistical Theory with Engineering Applications, John Wiley and Sons, Inc., New York, 1952.
4. Curran, T. C. and N. H. Frank. "Assessing the validity of the lognormal model when predicting maximum air pollutant concentrations," Paper No. 75-51.3, 68th Annual Meeting of the Air Pollution Control Association, Boston, Massachusetts (1975).
5. Johnson, T. A Comparison of the Two-Parameter Weibull and Lognormal Distributions Fitted to Ambient Ozone Data.
6. Weibull, W. "A Statistical Distribution Function of Wide Applicability." J. of Applied Mechanics. 18:392-297 (September 1951).

#### D.8 BIBLIOGRAPHY

1. Pollack, R. I. Studies of Pollutant Concentration Frequency Distributions, Environmental Monitoring Series, EPA-650/4-75-004, 1975.

APPENDIX E  
ESTIMATION PROCEDURES

E.1 INTRODUCTION

The problems to which statistical methods of analysis are most often applied fall into one of two classes: (1) estimation of one or more unknown parameters for the population from which the sample was selected, and (2) testing hypotheses concerning the population parameters or the validity of the model assumed for the population. Problems of estimation can be further subdivided into those involving point estimates and those involving interval estimation. The problems of testing hypotheses will not be included in this appendix. The reader may refer to texts referenced at the end of this appendix for information concerning testing hypotheses and further information on estimation.<sup>1,2,3,4</sup>

E.2 ESTIMATION

E.2.1 Point Estimates

It is often necessary to obtain a single value estimate of a population parameter. For example, the average ( $\bar{X}$ ) of a sample of concentrations of TSP for  $n$  equal to 60 days is used to estimate the annual mean ( $\mu$ ). Similarly, if one is concerned with variability in a set of data, the sample standard deviation ( $s$ ) is used to estimate the standard deviation ( $\sigma$ ) of the population. The sample variance ( $s^2$ ) is used to estimate the variance  $\sigma^2$  of the population, as described in Appendix C.

Further, in the situation where a linear calibration curve is used to express the relationship between an instrument reading ( $Y$ ) and the concentration of a standard sample ( $X$ ), the slope ( $b$ ) and intercept ( $a$ ) of the fitted line are used as point estimates of the parameters  $\beta$  and  $\alpha$  in the model. This procedure is described in Appendix J.

A point estimate is determined from the data for a sample of n observations selected from the population. The procedure for selecting a sample is very important. The sample must be a representative subset of the total population for which an inference is to be made. Procedures for the selection of a sample have been developed through extensive study.<sup>5</sup> Obviously if the sample is not representative of the population, the point estimate may be a biased estimate of the population parameter. For example, if it is desired to estimate the annual mean (or geometric mean) daily concentration of a pollutant, a sample of days must be selected from all days in the year in such a manner as to represent the entire year in terms of daily, weekly and seasonal variations. See Appendix I concerning procedures for selecting a sample.

Quality control procedures are generally based on samples of fewer than 8 observations. In this situation the range, R = max. value - min. value, is occasionally used to derive a point estimate of the population standard deviation rather than s (Appendix C). In statistical language, R is said to be nearly as efficient as the sample standard deviation for small samples. The equation for estimating  $\sigma$  from the range is

$$\hat{\sigma} = \frac{R}{d_2} .$$

Values of  $d_2$  for selected sample sizes are presented below:<sup>6</sup>

<u>n</u>	<u>d<sub>2</sub></u>
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078

### E.2.2 Confidence Interval

It should be immediately obvious that the sample averages  $\bar{X}_1$  and  $\bar{X}_2$ , for two independent samples of size  $n$  selected at random from a population will likely not have the same numerical value. Similarly, neither  $\bar{X}_1$  nor  $\bar{X}_2$  would be expected to be equal to the population mean  $\mu$ . In fact, if one were to select a large number of samples of size  $n$ , one could construct a frequency distribution of the sample averages. The average of the sample averages ( $\bar{\bar{X}}$ ), if the number of independent samples is quite large, would be essentially equal to the population mean  $\mu$ . The standard deviation of an average is given by  $\sigma/\sqrt{n}$ , where  $\sigma$  is the standard deviation of the observations comprising the population.

Because the sample average is not likely to be equivalent in value to the population mean, it is common practice to calculate two values,  $A$  and  $B$ , such that there is a given confidence that the interval ( $A \leq \mu \leq B$ ) will include the unknown value of the population mean  $\mu$ . The interval so specified is referred to as a confidence interval. The probability statement for the confidence interval estimate for the population mean ( $\mu$ ) is

$$P (A \leq \mu \leq B) = 1 - \alpha$$

$$P \left( \bar{X} - \frac{st_{n-1,\alpha}}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{st_{n-1,\alpha}}{\sqrt{n}} \right) = 1 - \alpha \quad (1)$$

where  $\bar{X}$  = sample mean

$s$  = sample standard deviation

$n$  = sample size

$\alpha$  = risk level (usually 0.10, 0.05, or 0.01)

$t_{n-1,\alpha}$  = value of the Student "t" distribution for  $n-1$  degrees of freedom and risk level  $\alpha$  (See Table E.1).

The risk level ( $\alpha$ ) is determined by the consequence which may result from an incorrect decision.

TABLE E.1. PERCENTILES OF THE 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
$\infty$	.253	.524	.842	1.282	1.645	1.960	2.326	2.576

Adapted by permission of R. A. Fisher and F. Yates, Statistical Tables for Biological, Agricultural and Medical Research, published by Longman Group Ltd., London, (previously published by Oliver and Boyd, Edinburgh) and by permission of the authors and publishers.

\*For two-tailed tests (or symmetrical confidence intervals), use, for example,  $t_{0.95}$  for obtaining a 90% confidence interval,  $t_{0.975}$  for 95% confidence, etc.

Example E.1 Construct a 95% confidence interval estimate of the population mean concentration based upon a sample of 16 observations for which the sample average  $\bar{X}$  and standard deviation  $s$  are 165 and 20  $\mu\text{g}/\text{m}^3$ , respectively:

The value of  $\alpha$  is 0.05, and from Table E.1 the value  $t_{n-1, \alpha}$  is  $t_{15, 0.05} = 2.131$ .

$$\frac{st_{15, 0.05}}{\sqrt{n}} = \frac{(20) 2.131}{\sqrt{16}} = 11$$

Thus, there is 95% confidence that the following interval includes  $\mu$ ,

$$165 - 11 \leq \mu \leq 165 + 11$$

or

$$154 \leq \mu \leq 176 \mu\text{g}/\text{m}^3.$$

The interval 154 to 176  $\mu\text{g}/\text{m}^3$  is defined to be a 95% confidence interval for  $\mu$ .

Example E.2 Using the information for Example E.1, construct a 99% confidence interval estimate for  $\mu$ .

$$t_{15, 0.01} = 2.947$$

$$\frac{st_{15, 0.01}}{\sqrt{n}} = \frac{(20) 2.947}{\sqrt{16}} = 15$$

Thus, there is 99% confidence that the inequality is satisfied,

$$150 \leq \mu \leq 180 \mu\text{g}/\text{m}^3.$$

Example E.3 Suppose that five measurements are made of a standard sample and found to be 44, 50, 47, 50, and 53  $\mu\text{g}/\text{m}^3$ . Assuming no bias in these laboratory measurements, estimate the mean concentration of the standard with a 95% confidence interval.

The confidence limits are given by

$$\bar{X} \pm \frac{st_{4, 0.05}}{\sqrt{n}}$$

where  $\bar{X} = 48.8 \mu\text{g}/\text{m}^3$  and  $s = 3.42 \mu\text{g}/\text{m}^3$ .

The 95% confidence interval is given by

$$\bar{X} \pm \frac{st_{4,0.05}}{\sqrt{n}}$$

or

$$48.8 \pm \frac{3.42}{\sqrt{5}} (2.776)$$

$$48.8 \pm 4.3 \mu\text{g}/\text{m}^3.$$

As the level of confidence increases from 95 to 99%, the width of the confidence interval also increases. Similarly, if the level of confidence was to be reduced to 90%, the width of the confidence interval would decrease.

The procedures for constructing confidence limits for the standard deviation, for regression parameters, and for other parameters are somewhat more complex than that for the mean. These procedures are presented in many elementary texts on statistical methods.<sup>1,4,6,7</sup>

**Example E.4** Suppose that the information is provided later by the supplier of standard gas cylinder that the true value of the concentration of the standard sample of Example E.3 is  $50.1 \pm 0.2 \mu\text{g}/\text{m}^3$ . Are the measurements given in Example E.3 consistent with this information?

It is obvious that they are consistent, but in general, it is necessary to determine if the confidence interval contains the given or reference value, in this case  $50.1 \pm 0.2 \mu\text{g}/\text{m}^3$ . This range of values falls within the interval given by the solution to Example E.3.

### E.3 DETERMINATION OF SAMPLE SIZE

In the previous example the width of the confidence interval was given by

$$\frac{2s}{\sqrt{n}} t_v \quad (2)$$

where  $v$  is the number of degrees of freedom equal to  $n-1$  for applications in this section.

The width of the confidence interval varies from sample to sample according to  $s$  and  $n$ . Ideally it is desirable to estimate  $n$  to yield a confidence interval which has a practical width, such as a specified % of the measured value, say 10%. In order to determine  $n$  precisely, it is necessary to know the standard deviation beforehand. Although this information is not available before the sample is taken, one usually has some previous experience which indicates that the standard deviation is, for example, approximately 5% of the mean. The computation is given in Example E.5 for estimating the sample size  $n$ .

**Example E.5** How many measurements should be made of a standard reference sample to obtain the sample concentration  $\bar{X}$  within 2% of the true value  $\mu$ , assuming  $\sigma = 5\%$  of the true value, ( $\sigma/\mu = 0.05$ ), and 95% confidence is desired. It is further assumed that there is no measurement bias.

Solution to E.5: The width of the confidence interval is  $2(0.02\mu) = 0.04\mu$ , and the standard deviation is  $0.05\mu$ , where  $\mu$  is the mean. Thus the width of the 95% confidence interval is given by

$$\frac{2(0.05\mu)}{\sqrt{n}} t_{n-1,0.05} = 0.04\mu$$

or

$$\frac{0.10 t_{n-1,0.05}}{0.04} = \sqrt{n}$$

or

$$n = 6.25 t_{n-1,0.05}^2$$

For  $n$  large,  $t_{n-1,0.05} \cong 2$ , and hence  $n \cong 25$ . Thus if 25 measurements are made of a standard sample and the confidence interval determined, the observed average would fall within about 2% of the mean. For example, if  $\mu = 50 \mu\text{g}/\text{m}^3$ , the average

should fall between 49 and 51  $\mu\text{g}/\text{m}^3$ . In general, an approximation for n large is given by

$$n = 2^2 \left( \frac{\hat{\sigma}}{\delta} \right)^2 \left( \frac{\text{RSD}}{d} \right)^2 \quad (3)$$

(Ratio of the estimated or guessed standard deviation (s or  $\hat{\sigma}$ ) to the halfwidth of the confidence interval both expressed as a percentage of the mean).<sup>2</sup>

that is,

$$n = 4 \left( \frac{\hat{\sigma}}{\delta} \right)^2 = 4 \left( \frac{\text{RSD}}{d} \right)^2 \quad (3)$$

where

RSD =  $100 \hat{\sigma}/\hat{\mu}$  = estimated relative standard deviation, %

$\hat{\sigma}$  = guessed or estimated standard deviation

$\hat{\mu}$  = guessed mean level

$\delta$  = half-width of the confidence interval in absolute units

d = half-width of the confidence interval as a percentage of the mean (the relative margin of error), %.

A second approach to determining the sample size depends on the availability of a preliminary sample, from which it is desired to estimate how much additional data are required to obtain an estimate of  $\mu$  with a specified precision. This procedure is in Reference 7. Again  $\delta$  is the allowable (absolute) margin of error, s is the sample standard deviation (based on preliminary data), RSD is the sample relative standard deviation, and t is the tabulated value for the degrees of freedom associated with s, and hence

$$n = \frac{t^2 s^2}{\delta^2} \quad (4)$$

This approach can be applied as a two stage procedure as follows:

- (1) choose the allowable margin or error  $\delta$  and the risk  $\alpha$  that the estimate  $\bar{X}$  of  $\mu$  will be off by  $\delta$  or more

(2) use  $\hat{\sigma}$  (a guessed value) to compute  $n'$  (first estimate of total sample size required)

$$(3) \text{ compute } n' = \left[ \frac{z_{1-\alpha/2} \hat{\sigma}}{\delta} \right]^2 \quad (5)$$

(4) use  $n_1$ , the size for the first sample at about  $0.5n'$

(5) make the observations and compute  $s_1$ , the standard deviation for the first sample

(6) use  $s_1$  to compute  $n$ , where

$$n = \frac{t^2 s_1^2}{\delta^2}, \quad (6)$$

(7) the sample size for the second stage  $n_2$  is  $n_2 = n - n_1$ .

This approach ensures that the final confidence interval satisfies the conditions specified, where as the previous approach gives a guessed value for  $n$  and the resulting statement may not satisfy the prescribed margin of error. One should refer to Reference 7 if it is desired to compute sample sizes for other applications such as comparing the means of two populations (e.g., a control treatment vs. a standard).

Example E.6 Suppose that for Example E.5 it is decided to make  $n = 12$  measurements in the initial sample and then to obtain additional measurements to ensure that the margin of error will not exceed  $\delta = 0.02$  with risk  $\alpha = 0.05$ . Assume the standard deviation of the first sample ( $s_1$ ) to be 0.035 ppm and then determine the sample size for the second stage ( $n_2$ ).

Solution to E.6

$$n = \frac{t^2 s_1^2}{\delta^2} = \frac{(2.201)^2 (0.035)^2}{(0.02)^2} = 14.84 \text{ or } 15$$

to ensure a margin of error  $\leq 0.02$  ppm. Hence  $n_2 = n - n_1 = 15 - 12 = 3$  additional measurements would be required to meet the specified conditions.

E.4 REFERENCES

1. Youden, W. J., Statistical Methods for Chemists, John Wiley & Sons, Inc., New York, 1951.
2. Bauer, E. L., A Statistical Manual for Chemists, Second Edition, Academic Press, New York, 1971.
3. Mandel, John, The Statistical Analysis of Data, Interscience Publishers, a division of John Wiley & Sons, Inc., New York, 1964.
4. Hald, A., Statistical Theory with Engineering Applications, John Wiley and Sons, Inc., New York, 1952.
5. Cochran, W. G., Sampling Techniques, John Wiley & Sons, Inc., New York, 1959.
6. Bennett, C. A. and N. L. Franklin, Statistical Analysis On Chemistry and Chemical Industry, John Wiley & Sons, Inc., New York, 1954, (Table 10.1, p. 636).
7. Ordnance Engineering Design Handbook, Experimental Statistics, Section 1, Basic Concepts and Analysis of Measurement Data, ORDP 20-110, June 1962.

## APPENDIX F

### OUTLIERS

#### F.1 INTRODUCTION

An unusually large (or small) value or measurement in a set of observations is usually referred to as an outlier. Some of the reasons for an outlier in data are:

Faulty instrument or component part  
Inaccurate reading of record, dial, etc.

Error in transcribing data

Calculation errors

Actual value due to unique circumstances under which the observation(s) was obtained--an extreme manifestation of the random variability inherent in the data.

It is desired to have some statistical procedure to test the presence of an outlier in a set of measurements. The purpose of such tests would be to:

1. Screen data for outliers and hence to identify the need for closer control of the data generating process.

2. Eliminate outliers prior to analysis of the data. For example, in developing control charts the presence of outliers would lead to limits which are too wide and would make the use of the control charts of minimal, if any, value. In most statistical analysis of data (e.g., regression analysis and analysis of variance) the presence of outliers violate a basic assumption of the analysis. Incorrect conclusions are likely to result if the outliers are not eliminated prior to analysis. Outliers should be reported, and their omission from analysis should be noted.

3. Identify the real outliers due to unusual conditions of measurement (e.g., a TSP concentration which is abnormally

large due to local environmental conditions during the time of sample collection). Such observations would not be indicative of the usual concentrations of TSP, and may be eliminated depending on the use of the data. Ideally, these unusual conditions should be recorded on the field data report. Failure to report complete information and unusual circumstances surrounding the collection and analysis of the sample often can be detected by outlier tests. Having identified the outliers using one or more tests, it is necessary to determine, if possible, the cause of the outlier and then to correct the data if appropriate.

It will be assumed in this discussion that the measurements are normally distributed and that the sample of  $n$  measurements is being studied for the possibility of one or two outliers. If the measurements are lognormally distributed, such as for concentration of TSP, then the logarithm of the data should be taken prior to application of the tests given herein.

## F.2 PROCEDURE(S) FOR IDENTIFYING OUTLIERS

Let the set of  $n$  measurements be arranged in ascending order and denoted by

$$X_1, X_2, \dots, X_n$$

where  $X_i$  denotes the  $i$ th smallest measurement. Suppose that  $X_n$  is suspected of being too large, and that a statistical test is to be applied to the particular measurement to determine whether  $X_n$  is consistent with the remaining data in the sense that it is reasonable that it is part of the same population of measurements from which the sample is taken. Consider the following TSP data from a specific monitoring site during August 1978.

Example F.1	<u>TSP, <math>\mu\text{g}/\text{m}^3</math></u>	<u>ln TSP</u>
	40	3.69
	88	4.48
	71	4.26
	175	5.16
	85	4.44

One test procedure for questionable data is to use a test by Dixon,<sup>1</sup> see Table F.1,

$$r_{10} = \frac{X_n - X_{n-1}}{X_n - X_1} = \frac{175-88}{175-40} = \frac{87}{135} = 0.655. \quad (1)$$

Referring to Table F.1 the 5% significance level for  $r_{10}$  is 0.642 and we would thus declare that the value 175 appears to be an outlier. The value should be flagged for further investigation. We do not automatically remove data because a statistical test indicates the value(s) to be questionable.

Suppose that we know that the data are lognormally distributed (or at least that the log normal distribution is a very good approximation), then we should examine the Dixon Ratio for this example. Using the logarithm, the Dixon ratio is

$$r_{10} = \frac{5.16 - 4.48}{5.16 - 3.69} = 0.46,$$

and this value is not significant at the 5% level. Hence on this basis the extreme value 175 is not questionable.

We still may wish to investigate the value further (data permitting) and we compare the data with those at a neighboring site. The corresponding data are given below.

<u>Site 20</u> <u>TSP, <math>\mu\text{g}/\text{m}^3</math></u>	<u>Site 14</u> <u>TSP, <math>\mu\text{g}/\text{m}^3</math></u>
40	42
88	53
71	56
175	129
85	64

Thus we see that the value 175 does not appear to be questionable in view of the corresponding value for a neighboring site. Both sites have high values on the same day, suggesting a common source of the high values. The only means to investigate these values further is to go to the source of the data collection and review the meteorological factors, comments in the site logbooks relative to local construction activity, daily traffic, and other possible causation factors.

TABLE F.1. DIXON CRITERIA FOR TESTING OF EXTREME OBSERVATION (SINGLE SAMPLE)\*

n	Criterion.	Significance level		
		10%	5%	1%
3 4 5	$r_{10} = \frac{x_2 - x_1}{x_n - x_1}$ if* smallest value is suspected; $= \frac{x_n - x_{n-1}}{x_n - x_1}$ if largest value is suspected.	.886	.941	.988
6 7		.679 .557	.765 .642	.889 .780
8 9 10	$r_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1}$ if smallest value is suspected; $= \frac{x_n - x_{n-1}}{x_n - x_2}$ if largest value is suspected.	.479	.554	.683
		.441 .409	.512 .447	.635 .597
11 12 13	$r_{31} = \frac{x_3 - x_1}{x_{n-1} - x_1}$ if smallest value is suspected. $= \frac{x_n - x_{n-2}}{x_n - x_2}$ if largest value is suspected.	.517	.576	.679
		.490 .467	.546 .521	.642 .615
14 15 16 17 18	$r_{22} = \frac{x_3 - x_1}{x_{n-2} - x_1}$ if smallest value is suspected. $= \frac{x_n - x_{n-2}}{x_n - x_3}$ if largest value is suspected;	.492	.546	.641
19 20 21 22 23 24 25		.472 .454 .438 .424	.525 .507 .490 .475	.616 .595 .577 .561

\*Reproduced with permission from W. J. Dixon, "Processing Data for Outliers," Biometrics, March 1953, Vol. 9, No. 1, Appendix, Page 89. (Reference [1])

$$x_1 \leq x_2 \leq \dots \leq x_{n-2} \leq x_{n-1} \leq x_n$$

Criterion  $r_{10}$  applies for  $3 \leq n \leq 7$

Criterion  $r_{11}$  applies for  $8 \leq n \leq 10$

Criterion  $r_{21}$  applies for  $11 \leq n \leq 13$

Criterion  $r_{22}$  applies for  $14 \leq n \leq 25$

This example points out several considerations in validating data and in particular in detecting and flagging outliers.

1. The use of a statistical procedure for detecting an outlier is a first step and the result should not be to throw out the value(s) if the statistic is significant but to treat the value(s) as suspect until further information can be obtained.

2. The statistical procedures depend on specific assumptions, particularly concerning the distribution of the data--normal, lognormal, and Weibull--and the result should be checked using the distribution which best approximates the data.

3. Often there are values at neighboring sites which can be used to compare the values. If the values at the two sites are correlated, as in the Example F.1, this approach can be very helpful.

4. The final resolution of the suspect values can be made by the collection agency, thus the importance of performing the data validation at the local agency.

Another commonly used test procedure,<sup>2</sup> requires additional computation and is given by

$$T_n = (X_n - \bar{X})/s \quad (2)$$

where:  $X_n$  is the largest observed value among n measurements,  
 $\bar{X}$  is the sample average,  
s is the sample standard deviation (i.e.,  
 $s = \{\sum(X - \bar{X})^2 / (n-1)\}^{1/2}$ ).

For the data set previously given,

$$\begin{aligned} X_n &= 175 \\ \bar{X} &= 91.8 \\ s &= 50.2 \end{aligned}$$

and hence  $T_n = 1.66$ , which is not significant at the 0.05 level, that is, it is less than 1.672 which is the tabulated value for this level from Table F.2. This test result is not in agreement with the previous one, however, both test results are borderline

TABLE F.2. TABLE OF CRITICAL VALUES FOR T(ONE-SIDED TEST OF  $T_1$  OR  $T_n$ ) WHEN THE STANDARD DEVIATION IS CALCULATED FROM THE SAME SAMPLE

Number of Observations n	Upper .1% Significance Level	Upper .5% Significance Level	Upper 1% Significance Level	Upper 2.5% Significance Level	Upper 5% Significance Level	Upper 10% Significance Level
3	1.155	1.155	1.155	1.155	1.153	1.148
4	1.499	1.496	1.492	1.481	1.463	1.425
5	1.780	1.764	1.749	1.715	1.672	1.602
6	2.011	1.973	1.944	1.887	1.822	1.729
7	2.201	2.139	2.097	2.020	1.938	1.828
8	2.358	2.274	2.221	2.126	2.032	1.909
9	2.492	2.387	2.323	2.215	2.110	1.977
10	2.606	2.482	2.410	2.290	2.176	2.036
11	2.705	2.564	2.485	2.355	2.234	2.088
12	2.791	2.636	2.550	2.412	2.285	2.134
13	2.867	2.699	2.607	2.462	2.331	2.175
14	2.935	2.755	2.659	2.507	2.371	2.213
15	2.997	2.806	2.705	2.549	2.409	2.247
16	3.052	2.852	2.747	2.585	2.443	2.279
17	3.103	2.894	2.785	2.620	2.475	2.309
18	3.149	2.932	2.821	2.651	2.504	2.335
19	3.191	2.968	2.854	2.681	2.532	2.361
20	3.230	3.001	2.884	2.709	2.557	2.385
21	3.266	3.031	2.912	2.733	2.580	2.408
22	3.300	3.060	2.939	2.758	2.603	2.429
23	3.332	3.087	2.963	2.781	2.624	2.448
24	3.362	3.112	2.987	2.802	2.644	2.467
25	3.389	3.135	3.009	2.822	2.663	2.486
26	3.415	3.157	3.029	2.841	2.681	2.502
27	3.440	3.178	3.049	2.859	2.698	2.519
28	3.464	3.199	3.068	2.876	2.714	2.534
29	3.486	3.218	3.085	2.893	2.730	2.549
30	3.507	3.236	3.103	2.908	2.745	2.563
31	3.528	3.253	3.119	2.924	2.759	2.577
32	3.546	3.270	3.135	2.938	2.773	2.591
33	3.565	3.286	3.150	2.952	2.786	2.604
34	3.582	3.301	3.164	2.965	2.799	2.616
35	3.599	3.316	3.178	2.979	2.811	2.628
36	3.616	3.330	3.191	2.991	2.823	2.639
37	3.631	3.343	3.204	3.003	2.835	2.650
38	3.646	3.356	3.216	3.014	2.846	2.661
39	3.660	3.369	3.228	3.025	2.857	2.671
40	3.673	3.381	3.240	3.036	2.866	2.682
41	3.687	3.393	3.251	3.046	2.877	2.692
42	3.700	3.404	3.261	3.057	2.887	2.700
43	3.712	3.415	3.271	3.067	2.896	2.710
44	3.724	3.425	3.282	3.075	2.905	2.719
45	3.736	3.435	3.292	3.085	2.914	2.727
46	3.747	3.445	3.302	3.094	2.923	2.736
47	3.757	3.455	3.310	3.103	2.931	2.744
48	3.768	3.464	3.319	3.111	2.940	2.753
49	3.779	3.474	3.329	3.120	2.948	2.760
50	3.789	3.483	3.336	3.128	2.956	2.768

Reproduced with permission from American Statistical Association.

Use  $T_1 = \frac{\bar{X} - X_1}{s}$  when testing the smallest value,  $X_1$ .

Use  $T_n = \frac{X_n - \bar{X}}{s}$  when testing the largest value,  $X_n$  in a sample of n observations. Unless one has prior information about largest values (or smallest values) the risk levels should be multiplied by two for application of the test.

TABLE F.2 (continued)

Number of Observations n	Upper .1% Significance Level	Upper .5% Significance Level	Upper 1% Significance Level	Upper 2.5% Significance Level	Upper 5% Significance Level	Upper 10% Significance Level
51	3.798	3.491	3.345	3.136	2.964	2.775
52	3.808	3.500	3.353	3.143	2.971	2.783
53	3.816	3.507	3.361	3.151	2.978	2.790
54	3.825	3.516	3.368	3.158	2.986	2.798
55	3.834	3.524	3.376	3.166	2.992	2.804
56	3.842	3.531	3.383	3.172	3.000	2.811
57	3.851	3.539	3.391	3.180	3.006	2.818
58	3.858	3.546	3.397	3.186	3.013	2.824
59	3.867	3.553	3.405	3.193	3.019	2.831
60	3.874	3.560	3.411	3.199	3.025	2.837
61	3.882	3.566	3.418	3.205	3.032	2.842
62	3.889	3.573	3.424	3.212	3.037	2.849
63	3.896	3.579	3.430	3.218	3.044	2.854
64	3.903	3.586	3.437	3.224	3.049	2.860
65	3.910	3.592	3.442	3.230	3.055	2.866
66	3.917	3.598	3.449	3.235	3.061	2.871
67	3.923	3.605	3.454	3.241	3.066	2.877
68	3.930	3.610	3.460	3.246	3.071	2.883
69	3.936	3.617	3.466	3.252	3.076	2.888
70	3.942	3.622	3.471	3.257	3.082	2.893
71	3.948	3.627	3.476	3.262	3.087	2.897
72	3.954	3.633	3.482	3.267	3.092	2.903
73	3.960	3.638	3.487	3.272	3.098	2.908
74	3.965	3.643	3.492	3.278	3.102	2.912
75	3.971	3.648	3.496	3.282	3.107	2.917
76	3.977	3.654	3.502	3.287	3.111	2.922
77	3.982	3.658	3.507	3.291	3.117	2.927
78	3.987	3.663	3.511	3.297	3.121	2.931
79	3.992	3.669	3.516	3.301	3.125	2.935
80	3.998	3.673	3.521	3.305	3.130	2.940
81	4.002	3.677	3.525	3.309	3.134	2.945
82	4.007	3.682	3.529	3.315	3.139	2.949
83	4.012	3.687	3.534	3.319	3.143	2.953
84	4.017	3.691	3.539	3.323	3.147	2.957
85	4.021	3.695	3.543	3.327	3.151	2.961
86	4.026	3.699	3.547	3.331	3.155	2.966
87	4.031	3.704	3.551	3.335	3.160	2.970
88	4.035	3.708	3.555	3.339	3.163	2.973
89	4.039	3.712	3.559	3.343	3.167	2.977
90	4.044	3.716	3.563	3.347	3.171	2.981
91	4.049	3.720	3.567	3.350	3.174	2.984
92	4.053	3.725	3.570	3.355	3.179	2.989
93	4.057	3.728	3.575	3.358	3.182	2.993
94	4.060	3.732	3.579	3.362	3.186	2.996
95	4.064	3.736	3.582	3.365	3.189	3.000
96	4.069	3.739	3.586	3.369	3.193	3.003
97	4.073	3.744	3.589	3.372	3.196	3.006
98	4.076	3.747	3.593	3.377	3.201	3.011
99	4.080	3.750	3.597	3.380	3.204	3.014
100	4.084	3.754	3.600	3.383	3.207	3.017

Source: Grubbs, F. E., and Beck, G., Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations, *Technometrics*, Vol. 14, No. 4, Nov. 1972, pp. 847-854.

situations. If the  $T_n$  is applied to the logarithms, the result is  $T_n = \frac{5.16-4.41}{0.527} = 1.42$ , which is not significant and which agrees with the Dixon ratio test. In many examples it will be obvious that a particular value is an outlier, whereas in Example F.1 this is not the case. A plot of the data is often helpful in examining a set of data.

After rejecting one outlier using either  $T_n$  or  $T_1$  the analyst may be faced with the problem of considering a second outlier. In this case the mean and standard deviation may be re-estimated and either  $T_{n-1}$  or  $T_1$  applied to the sample of  $n-1$  measurements. However, the user should be aware that the test  $T_n$  or  $T_1$  is not theoretically based on repeated use.

Grubbs<sup>2</sup> gives a test procedure (including tables for the critical values) for simultaneously testing the two largest or two smallest values. This procedure is not given here.

The use of the procedures given in Table F.1 requires very little computation and would be preferable on a routine basis. Grubbs<sup>3</sup> gives a tutorial discussion of outliers and is a very good reference to the subject. A recent text on outliers is also recommended to the reader with some statistical background.<sup>4</sup>

One other procedure for data validation which has an advantage relative to the previous two procedures (Dixon and Grubbs) is the use of a statistical control chart.<sup>5,6</sup> The control chart is discussed in Appendix H and the reader is referred to that Appendix for details in application. The TSP data for a specific site for the years 1975 to 1977 for which there are five measurements per month are used as a historical data base for the control chart and the data for 1978 are plotted on the chart to indicate any questionable data. These data are shown in Table F.3 (historical data) and in Table F.4 (1978 data). Figure F.1 (upper part) is the control chart with both  $2\sigma$  and  $3\sigma$  limits for the averages.

$$\bar{\bar{X}} \text{ (average of the } \bar{X}'\text{s)} = 56.5 \text{ } \mu\text{g}/\text{m}^3$$

TABLE F.3. TSP DATA FROM SITE 397140014H01 SELECTED AS HISTORICAL DATA BASE FOR SHEWHART CONTROL CHART (1975-1977)

Month-year	Mean ( $\bar{X}$ ), $\mu\text{g}/\text{m}^3$	Range (R), $\mu\text{g}/\text{m}^3$	Month-year	Mean ( $\bar{X}$ ), $\mu\text{g}/\text{m}^3$	Range (R), $\mu\text{g}/\text{m}^3$
1-75	54.6	67	10-76	34.6	50
5-75	63.8	39	11-76	53.4	29
6-75	59.0	25	12-76	52.2	44
7-75	63.0	23	3-77	40.4	28
8-75	68.2	54	4-77	63.6	57
10-75	41.8	26	6-77	45.4	31
11-75	68.4	81	7-77	53.4	19
12-75	57.6	39	8-77	58.6	26
1-76	82.4	87	9-77	46.0	12
4-76	90.2	117	10-77	45.6	33
5-76	43.8	48	11-77	49.8	54
7-76	72.6	80	12-77	30.4	22
9-76	73.4	83			

TABLE F.4. TSP DATA FROM SITE 397140014H01 FOR CONTROL CHART (1978)

Data set	Month	Mean	Range	s
1	1	30.6	27	10.4
2	2	47.4	60	21.7
3	3	54.4	39	17.2
4	4	31.8	29	13.6
5	5	53.6	46	21.8
6	6	64.8	46	19.0
7	8	68.8	87	34.6
8	9	43.2	31	11.3
9	10	52.4	59	24.2
10	11	60.8	71	29.0
11	12	31.6	22	9.8

$\bar{X}$  AND S CHART

JECT Validation of TSP Data. MEASUREMENT Daily TSP at Site 39774001444 UNITS  $\mu\text{g}/\text{m}^3$

Section No. F  
Revision No. 1  
Date January 9, 1984  
Page 10 of 14

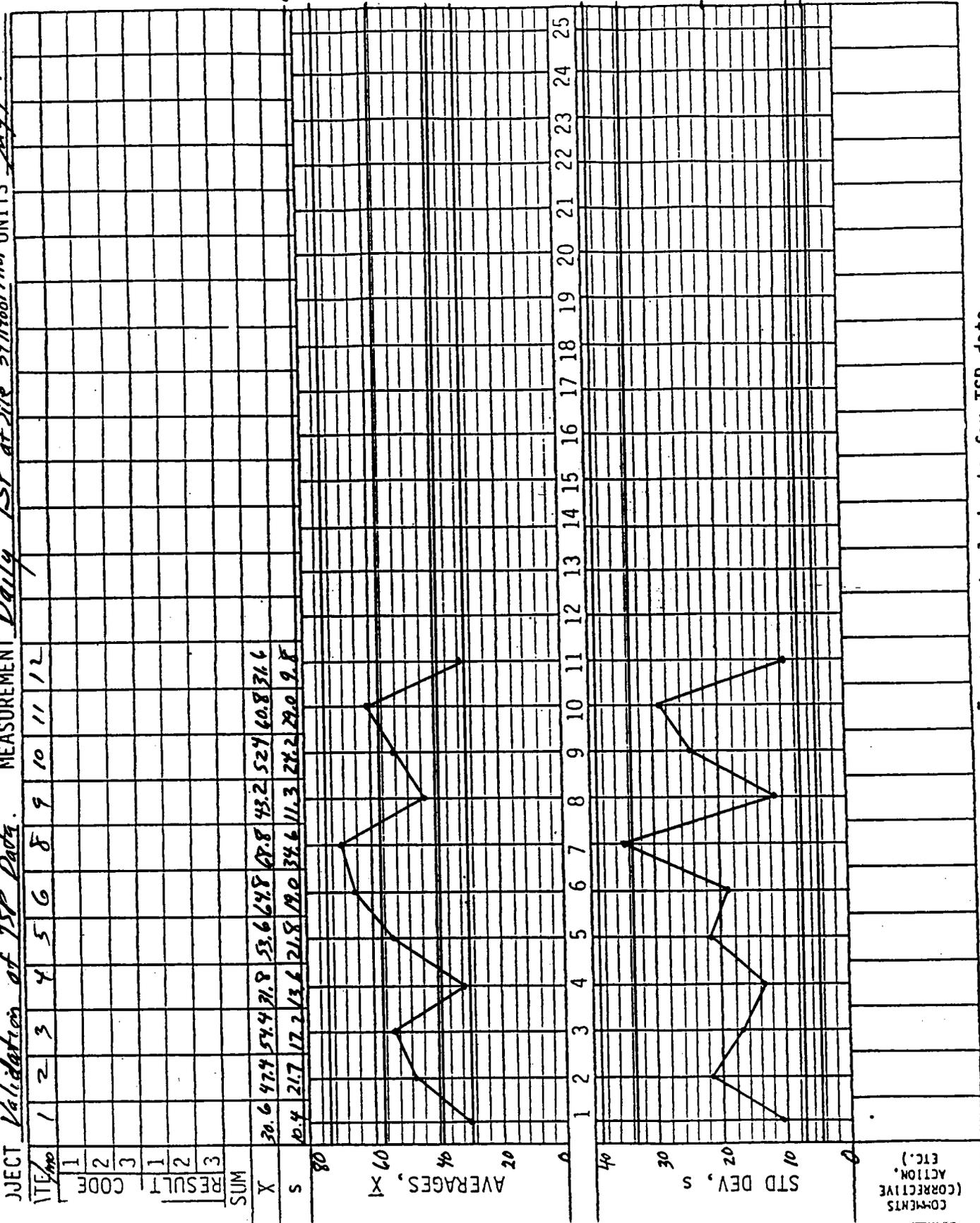


Figure F.1.  $\bar{X}$  and s control charts for TSP data.

$$\begin{aligned}\sigma_{\bar{X}} \text{ (standard deviation of the mean)} &= 9.0 \text{ } \mu\text{g}/\text{m}^3 \\ \text{UWL}_{\bar{X}} \text{ (upper } 2\sigma \text{ limit)} &= 74.5 \text{ } \mu\text{g}/\text{m}^3 \\ \text{LWL}_{\bar{X}} \text{ (lower } 2\sigma \text{ limit)} &= 38.5 \text{ } \mu\text{g}/\text{m}^3 \\ \text{UCL}_{\bar{X}} \text{ (} 3\sigma \text{)} &= 83.5 \text{ } \mu\text{g}/\text{m}^3 \\ \text{LCL}_{\bar{X}} \text{ (} 3\sigma \text{)} &= 29.5 \text{ } \mu\text{g}/\text{m}^3\end{aligned}$$

Figure F.1 shows three averages below the  $\text{LWL}_{\bar{X}}$  ( $2\sigma$  limit) and no values above the  $\text{UWL}_{\bar{X}}$  ( $2\sigma$  limit). No values are below the  $3\sigma$  limit  $\text{LCL}_{\bar{X}}$  ( $3\sigma$ ). Hence we do not suspect any averages to be significantly different from the historical average and which would suggest further investigation.

Figure F.1 (lower part) is the control chart for the standard deviation.

$$\begin{aligned}\bar{R} \text{ (average range)} &= 47.0 \\ \hat{\sigma} &= 0.43 (47.0) = 20.2 \\ \text{UWL}_s \text{ (upper } 2\sigma \text{ limit for } s \text{)} &= 33.7 \\ \text{LWL}_s \text{ (lower } 2\sigma \text{ limit for } s \text{)} &= 7.0 \\ \text{UCL}_s \text{ (99.5 percentile)} &= 38.9 \\ \text{LCL}_s \text{ (0.5 percentile)} &= 4.6\end{aligned}$$

There is a single outlier on this chart and this sample (one month of data--5 values) should be checked for factors which might explain the high value for the standard deviation. See Example F.1 for further discussion of this example relative to action taken after the flagging or identification of the questionable value. The same data were used in that example.

The advantage of the quality control chart approach is that not only are questionable values within a month detected, but also if all of the values for a month are high relative to values for other months, they will be flagged. The latter can result from personnel changes, instrument problems, calculation errors, and such changes will go undetected when comparing a single possible outlier within a data set. It is recommended that both test procedures (Dixon or Grubbs and the control chart) be used if resources permit, if not use the control chart technique.

### F.3 GUIDANCE ON SIGNIFICANCE LEVELS

The problem of selecting an appropriate level of significance in performing statistical tests for outliers is one of comparing two resulting costs. If the significance level is set too high (e.g., 0.10 or 0.20) there is the cost of investigating the data identified as questionable a relatively large proportion of the time that, in fact, the data are valid.<sup>1</sup> On the other hand, if the significance level is set too low (e.g., 0.005 or 0.001) invalid data may be missed and these data may be subsequently used in making incorrect decisions. This cost can also be large but is difficult to estimate. The person responsible for data validation must therefore seek an appropriate level based on these two costs. If the costs of checking the questionable data are small, it is better to err on the safe side and use  $\alpha = 0.05$  or 0.10 say. Otherwise, a value of  $\alpha = 0.01$  would probably be satisfactory for most applications. After experience is gained with the validation procedure, the  $\alpha$  value should be adjusted as necessary to minimize the total cost (i.e., the cost of investigating outliers plus that of making incorrect decisions).

### F.4 REFERENCES

1. Dixon, W. J., Processing Data for Outliers, Biometrics, Vol. 9, No. 1, March 1953, pp. 74-89.
2. Grubbs, F. E. and Beck, G., Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations, Technometrics, Vol. 14, No. 4, November 1972, pp. 847-854.
3. Grubbs, F. E., Procedures for Detecting Outlying Observations in Samples, Technometrics, Vol. 11, No. 1, February 1969, pp. 1-21.
4. Barnett, V. and T. Lewis, Outliers in Statistical Data, John Wiley and Sons, New York, 1978.
5. US Environmental Protection Agency, Screening Procedures for Ambient Air Quality Data, EPA-450/2-78-037, July 1978.

6. Nelson, A. C., D. W. Armentrout, and T. R. Johnson. Validation of Air Monitoring Data. EPA-600/4-80-030, June 1980.

#### F.5 BIBLIOGRAPHY

1. Curran, T. C., W. F. Hunt, Jr., and R. B. Faoro. Quality Control for Hourly Air Pollution Data. Presented at the 31st Annual Technical Conference of the American Society for Quality Control, Philadelphia, May 16-18, 1977.
2. Data Validation Program for SAROAD, Northrup Services, EST-TN-78-09, December 1978, (also see Program Documentation Manual, EMSL).
3. Faoro, R. B., T. C. Curran, and W. F. Hunt, Jr., "Automated Screening of Hourly Air Quality Data," Transactions of the American Society for Quality Control, Chicago, Ill., May 1978.
4. Hunt, Jr., W. F., J. B. Clark, and S. K. Goranson, "The Shewhart Control Chart Test: A Recommended Procedure for Screening 24-Hour Air Pollution Measurements," J. Air Poll. Control Assoc. 28:508, 1979.
5. Hunt, Jr., W. F., T. C. Curran, N. H. Frank, and R. B. Faoro, "Use of Statistical Quality Control Procedures in Achieving and Maintaining Clean Air," Transactions of the Joint European Organization for Quality Control/International Academy for Quality Conference, Vernice Lido, Italy, September 1975.
6. Hunt, Jr., W. F., R. B. Faoro, T. C. Curran, and W. M. Cox, "The Application of Quality Control Procedures to the Ambient Air Pollution Problem in the USA," Transactions of the European Organization for Quality Control, Copenhagen, Denmark, June 1976.
7. Hunt, Jr., W. F., R. B. Faoro, and S. K. Goranson, "A Comparison of the Dixon Ratio Test and Shewhart Control Test Applied to the National Aerometric Data Bank," Transactions of the American Society for Quality Control, Toronto, Canada, June 1976.
8. Rhodes, R. C., and S. Hochheiser. Data Validation Conference Proceedings. Presented by Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, EPA-600/9-79-042, September 1979.

9. US Department of Commerce. Computer Science and Technology: Performance Assurance and Data Integrity Practices. National Bureau of Standards, Washington, D. C., January 1978.
10. 1978 Annual Book of ASTM Standards, Part 41. Standard Recommended Practice for Dealing with Outlying Observations, ASTM Designation: E 178-75. pp. 212-240.

## APPENDIX G

### TREATMENT OF AUDIT DATA

#### G.1 AUDIT DATA

One means of checking on the performance of a measurement system or process is to conduct an independent audit of a pertinent portion of the system or of the entire system if possible. In conducting an audit, there will result a set of data collected by the standard test method and a second set of data collected by an audit procedure. The latter may be performed, for example, by an independent operator using the same or different measuring instruments. It is desirable that the two sets of measurements be made, in so far as possible, independently of one another. However, the audit must measure the same characteristic as the standard test measurement. One example of an audit would be to challenge an SO<sub>2</sub> analyzer with at least one gas of known concentration between 0.40 and 0.45 ppm SO<sub>2</sub> and to compare the analyzer response with the known concentration.

An audit is usually performed on a sampling basis, for example, by checking every tenth filter or one sampled at random from each set of ten. A rate of one out of about fourteen was suggested in the guideline documents.<sup>1</sup> The audit data are then used to infer if the measurement process is biased. This appendix will discuss the types and uses of audit data and the types of inferences which may be made from audit results.

#### G.2 DATA QUALITY ASSESSMENT

In accordance with 40 CFR 58, Appendix A,<sup>2</sup> an analyzer is challenged with at least one audit gas (of known concentration) from each of the specified ranges which fall within the measurement range of the analyzer being audited. The percentage difference ( $d_i$ ) between the concentration of the audit test gas

( $X_i$ ) and the concentration indicated by the analyzer ( $Y_i$ ) is used to assess the accuracy of the monitoring data, that is,

$$d_i = 100 \frac{Y_i - X_i}{X_i} . \quad (1)$$

The accuracy of a single analyzer is determined by the  $d_i$  for each audit concentration. If the  $d_i$  is within acceptable limits the analyzer is considered accurate; if not, corrective action is necessary. The accuracy for the reporting organization is calculated by averaging the  $d_i$  for each audit concentration level,

$$\bar{D} = \frac{1}{k} \sum_{i=1}^k d_i , \quad (2)$$

where there are  $k$  analyzers audited per quarter and  $\bar{D}$  is the average % difference for the  $k$  analyzers. (See Sections 2.0.8 and 2.0.9 of Volume II of this Handbook for further details).

If there is a consistent bias for the  $k$  analyzers within an agency,  $\bar{D}$  and  $S_D$  (the standard deviation of the differences  $d_i$ ) will reveal this because  $t = \sqrt{k}\bar{D}/S_D$  has a  $t$  distribution with  $k-1$  degrees of freedom (see Subsection G.4 for an example computation). If  $t$  is significantly large (positively or negatively) then there is a consistent bias for all the analyzers used by the agency. If on the other hand  $t$  is not large then we can infer that the biases vary among the analyzers. They may be large or small for individual analyzers. The individual values for  $d_i$  must be studied for making further conclusions.

### G.3 EPA AUDIT PERFORMANCE

Measurement principles for  $SO_2$ ,  $NO_2$ , CO, sulfate, nitrate, and Pb are audited on a semiannual basis. Blind samples, the concentrations of which are known only to EPA, are sent to participating laboratories. The analytical results are returned to EPA, Quality Assurance Division (QAD) for evaluation. After processing the data, an individual report is returned to each participant (laboratory). In addition, a summary report of the

audit results is prepared by EPA, QAD.<sup>3,4</sup> Some results for a CO audit are given in Appendix K. These data provide a measurement of the bias, precision, and accuracy of the audit data for measurement methods used by the participating laboratories for each of several (usually 3 to 5) concentration levels. The bias for all laboratories is given by the deviation of the median value from the true value, expressed as a percent. This is determined for each concentration level along with other statistics describing the variation of the data (e.g., range, relative standard deviation). The ultimate purpose of these audits is to provide information to the participants relative to the accuracy of their measurement method and hence to improve overall data quality by means of corrective actions taken by participants with respect to questionable data. See Appendix K for further discussion of these audits.

#### G.4 ANALYSIS OF AUDIT DATA

Consider the set of data given below.

No.	NO <sub>3</sub> Analysis (mg/filter)		
	Lab 1 (test data)	Lab 2 (audit data)	Difference (D)
1	1.7	2.0	-0.3
2	2.2	2.4	-0.2
3	3.9	3.7	0.2 (possible outlier)
4	3.3	3.6	-0.3
5	2.7	3.3	-0.6
6	3.5	3.8	-0.3
7	0.9	1.5	-0.6
8	1.3	1.5	-0.2
9	6.1	6.4	-0.3
10	2.9	3.2	-0.3

Lab 2 data are audits or checks on the Lab 1 test data and are to be used to determine if the test data are valid based on the following three criteria and problem types:

1. From past experience a maximum (absolute) difference between audit and test results of 1 mg has been suggested. What can one infer concerning the test data?

2. No standard (such as in (1) above) is available, but a statistical analysis is to be conducted to compare the two sets of data with significance level of 0.95 or a risk level of 0.05.

3. Assume that the audit data are unbiased and that it is desired to report the results of the test data as an estimated bias and expected range of variation using  $3\sigma$  limits.

#### G.4.1 Criterion (1)

Based on criterion (1) above, all of the differences (absolute) are less than 1 mg/filter and hence the test data would be considered to be unbiased. This analysis does not check the suggested standard of 1 mg/filter. This will be done with respect to the second criterion.

Suppose further that these 10 audits represent a random sample of 10 test measurements selected from 50 which are checked for validity. What can one infer about the set of 50 measurements? The answer to this question requires some further background in statistical sampling than that given in these appendices, including Appendix I. However, with appropriate tables on sampling<sup>1</sup> one can, for example, infer that:

"there is 50% confidence that the percent of good test measurements exceeds 90%,"

or

"there is 95% confidence that the percent of good test measurements exceeds 75%."

These are examples of the types of statements that can be made on just this one data set. As additional data are obtained, one's confidence in a given percent of good test data should increase if the test data are actually satisfactory.

#### G.4.2 Criterion (2)

In this case no prior information is assumed about the expected deviation between a test measurement and an audit value. Thus the comparison is made on the basis of the behavior of the two sets of data or, really, the differences in the corresponding data pairs. For example, one wishes to determine if there is a significant bias in the measurements, and secondly what is a reasonable difference or standard to suggest for acceptance of test data?

G.4.2.1 Paired t-test - A statistical check on the bias is provided by a paired t-test<sup>5,6</sup> which is described in almost any elementary text on statistical techniques and briefly herein. This is a very useful test for comparing paired data sets obtained by making two related measurements on the same sample or equivalent samples under the same conditions except, for example, a change in the operator and/or instrument. After taking the differences, the test is conducted ignoring the original data and using only the differences. The average  $\bar{D}$  and standard deviation of the differences  $s_D$  are obtained.

$$\bar{D} = -0.291$$

$$s_D = 0.22.$$

Using these values, a value of t (with 9 degrees of freedom) is calculated as follows,

$$t = \frac{\bar{D} - 0}{s_D/\sqrt{n}} = \frac{-0.291}{0.22/\sqrt{10}} = 4.18. \quad (3)$$

(See Table G.1 for a computational form for t.) This t value is then checked against the value in Table E.1 to determine if it is unusually small or large. Assuming 95% confidence (or 5% risk) it is obvious that this value is larger in absolute value than expected and hence one infers a bias exists.

Next consider the implications for a suggested standard of 1 mg/filter for differences between a test measurement and an

TABLE G.1. PAIRED SAMPLES, SPLIT SAMPLES OR DUPLICATES--  
 COMPARISON OF METHODS, LABS, OR REPEAT MEASUREMENTS

Sample number	X <sub>1</sub>	X <sub>2</sub>	Difference (D)	Sample number	X <sub>1</sub>	X <sub>2</sub>	X <sub>1</sub> - X <sub>2</sub> = D
1	—	—	—	11	—	—	—
2	—	—	—	12	—	—	—
3	—	—	—	13	—	—	—
4	—	—	—	14	—	—	—
5	—	—	—	15	—	—	—
6	—	—	—	16	—	—	—
7	—	—	—	17	—	—	—
8	—	—	—	18	—	—	—
9	—	—	—	19	—	—	—
10	—	—	—	20	—	—	—

Total =  $\Sigma D$  \_\_\_\_\_

Calculations:

Average difference =  $\bar{D} = \Sigma D/n =$  \_\_\_\_\_

Standard Deviation of Differences:

$\Sigma D^2 =$  \_\_\_\_\_

$(\Sigma D)^2/n =$  \_\_\_\_\_

Difference = \_\_\_\_\_

Divide by n-1

$s_D^2 =$  \_\_\_\_\_

$s_D =$  \_\_\_\_\_

$s_D/\sqrt{n} =$  \_\_\_\_\_

Is the mean difference equal to 0?

Calculate

$t = \frac{\bar{D}-0}{s_{\bar{D}}} = \frac{\sqrt{n}\bar{D}}{s_D} =$  \_\_\_\_\_  $=$  \_\_\_\_\_

Compare to tabulated t value in Table E.1 with n-1 degrees of freedom (n is the number of differences) for the selected level of significance or risk (e.g., for n-1 = 9 DF and 95% level of significance t = 2.262).

audit value for this particular analysis. In answering this question, the standard deviation of the difference,  $s_D = 0.22$  mg/filter, is a measure of the variation of the differences about their own mean difference. Hence  $3s_D = 0.66$  mg/filter can serve as an expected limit which would be exceeded a relatively small percentage of the time just as one would use  $3\sigma$  limits in developing control chart limits. However, there are two limitations to this approach which must be considered in developing reasonable limits, (1) only ten data pairs (differences) were available and this does not meet the usual recommendations for setting control limits, say  $n = 20$  pairs would be a preferred number of values, and (2) the bias is not considered. In practice some bias between labs, audit and test values is reasonable and an acceptable magnitude of bias must be determined. To determine an acceptable level, a number of further data sets like those given in the example must be analyzed.

G.4.2.2 Sign Test<sup>5</sup> - One simple test of a significant bias is to check the sign of the differences. If all ten differences are negative, then one has considerable doubt about the lack of a bias as it would be expected that on the average five would be positive and five negative if no bias were present. The chances that all ten are of one sign is like flipping an unbiased coin ten times and obtaining ten heads or ten tails; since this is a very small chance, one usually infers that there is a bias. In the example, there are nine negative differences among ten. The chances of 9 or 10 negative or positive differences, if there were no bias present, is given by the computation,<sup>5</sup>

$$2 \left[ 10 \left(\frac{1}{2}\right)^{10} + 1 \left(\frac{1}{2}\right)^{10} \right] = \frac{22}{1024} = 0.0215.$$

The first term in square brackets is the product of the number of ways of getting 1 tail and 9 heads (or vice versa) in 10 tosses of a coin, multiplied by  $1/2^{10}$  which is 1 divided by the number of arrangements of heads and tails for ten coins (i.e., two for each coin and  $2^{10}$  for ten coins). Similarly the

second term is the number of ways of obtaining 10 heads in 10 tosses of coin, or 1, multiplied by  $1/2^{10}$  as for the first term. The entire bracket is multiplied by 2 to take into account, getting all heads or all tails, or 9 heads, 1 tail or 9 tails, 1 head. Since this probability is very small, say less than 0.05, it is inferred that one set of data is biased with respect to the other set of data. Note that either the test data (lab 1) or the audit data (lab 2) or both may be biased. Unless some outside check of the results is available (e.g., against some reference standard) it is not possible to assume that one data set is not biased and the other set is biased.

#### G.4.3 Criterion (3)

In this case it is assumed further that the audit data (lab 2) are not biased, and that it is desired to present the test data (lab 1) in terms of the bias and variation. In this case, the bias is estimated to be  $\bar{D} = -0.29$  mg/filter, that is, the test data are biased low on the average by 0.29 mg/filter. Hence, based on these data, it is inferred, for a single test value that the true measurement in mg/filter is given by the following:

$$\begin{aligned} \text{Test measurement (X) - Bias} \pm 3s_D & \quad (4) \\ X + 0.29 \pm 3(0.22) & \end{aligned}$$

or between  $X - 0.37$  and  $X + 0.95$  mg/filter.

#### G.5 PRESENTATION OF AUDIT RESULTS

There are several ways in which the data can be presented to compare the routine measurements versus the audit measurements. One method is to plot the routine measurements versus the corresponding audit measurements. If there is good agreement the plotted points should follow a  $45^\circ$  line (assuming equivalent scales on both axes). If there is a systematic error in the results the data may follow a line with a slope very different from unity. Two examples are given in Figure G.1, one with good agreement G.1a and one with poor agreement G.1b between the routine and audited results.

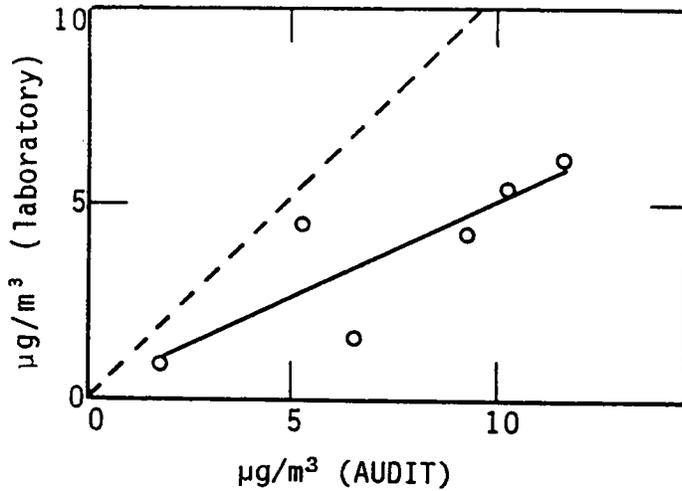


Figure G.1a. Nitrate comparison between laboratory and an audit data.

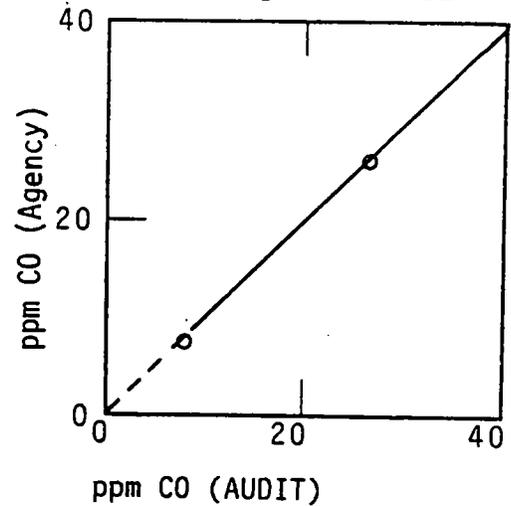


Figure G.1b. CO comparison between an agency and audit data.

Figure G.1. Examples of poor and good agreement between routine and audit results.

A second means of presenting the results is a plot of the  $d_j$ 's as a function of time. One can also add the upper and lower probability limits shown in Figure G.2. These data may also be presented in tabular form as in Table G.1. These data are for one agency, 5 audits; the tabulation contains the average difference  $\bar{d}_j$ , the standard deviation of the percent differences  $S_j$ , and the slope and intercept of the line relating the agency reported value to the audited value.

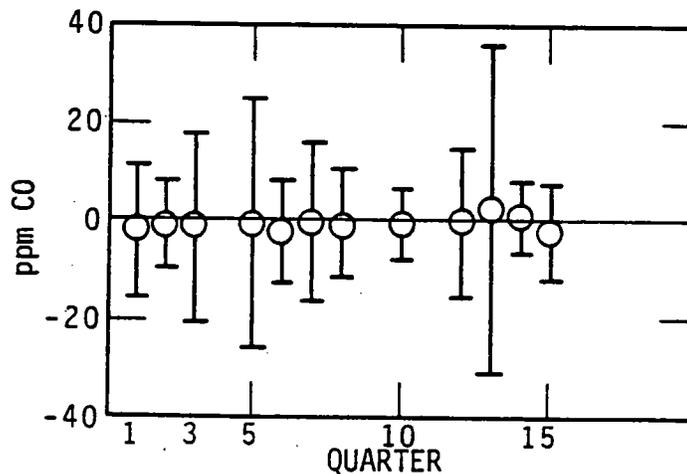


Figure G.2. CO performance evaluation for agencies as a function of time. The data at quarter No. 1 are actually 4th quarter 1976. The vertical axes show  $d_j$  and the 95% upper and lower probability limits in units of %.

TABLE G.1. SUMMARY OF PERFORMANCE FOR  $SO_4^{=}$  SURVEYS FOR ONE AGENCY

Quarter/ year	Average % diff. $d_j$	Standard deviation $S_i$	Slope	Intercept $\mu g/m^3$
2/77	7.3	4.5	1.005	0.333
3/77	6.8	3.5	1.042	0.194
4/77	5.4	5.8	1.009	0.843
1/78	4.9	15.5	1.175	-0.383
2/78	12.5	7.2	1.036	0.287

#### G.6 SUMMARY

In summary, some of the possible uses and methods of presentation of test and audit data are described in this appendix. If standard reference samples were available, they could be used to audit measurements made by analytical methods and the lab biases determined. Interlaboratory tests aid in estimating the within-lab and among-lab variation and the use of these tests is described in Appendix K.

#### G.7 REFERENCES

1. Smith, F. and Nelson, A.C., Guidelines for Development of Quality Assurance Programs and Procedures, Final Report to EPA on Contract No. EPA-Durham 68-02-0598, August 1973.
2. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS). Federal Register, Vol. 44, No. 92, 00. 27574-81. May 10, 1979.
3. Bromberg, S. M., R. L. Lampe, and B. I. Bennett, Summary of Audit Performance: Measurement of  $SO_2$ ,  $NO_2$ , CO, Sulfate, Nitrate, Lead, Hi-Vol Flow Rate - 1977, U.S. Environmental Protection Agency, EPA-600/5-79-014, February 1979.
4. Bromberg, S. M., R. L. Lampe, and B. I. Bennett, Summary of Audit Performance: Measurement of  $SO_2$ ,  $NO_2$ , CO, Sulfate, Nitrate, Lead, and Hi-Vol Flow Rate - 1978, U.S. Environmental Protection Agency, 1980.
5. Dixon, W.J. and Massey, F.J., Introduction to Statistical Analysis, McGraw-Hill Book Co., Inc., New York, 1951.

6. Youden, W.J., Statistical Methods for Chemists, John Wiley and Sons, Inc., New York, 1951.

G.8 BIBLIOGRAPHY

1. Cher, M. Quality Assurance in Support of Energy Related Monitoring Activities. EPA-600/7-79-136, June 1979.
2. Performance Audit Publication of Research Triangle Institute, Section 2.0.12.



## APPENDIX H CONTROL CHARTS

### H.1 DESCRIPTION AND THEORY

The control chart provides a tool for distinguishing the pattern of indeterminate (random) variation from the determinate (assignable cause) variation. The chart displays data from a process or method in a form which graphically compares the variability of all test results with the average and the expected variability of small groups of data.

The control charts in this appendix are constructed on standardized forms. Blank copies of these forms are included on the following two pages. The Handbook user should copy these forms and use them for constructing control charts for all routine measurement systems. The important features of the standard forms follow:

1. Measurement performed - Record the pollutant, or parameter, measured and the method of measurement, for example, SO<sub>2</sub> analysis of aqueous sodium sulfite standards, ...Method.
2. Measurement units - Metric units.
3. Date - Write year next to the date. Write the month and day in the appropriate column.
4. Measurement code - A number assigned to the measurement to permit easy reference to a more complete description of measurement conditions and results. This number, for example, should be traceable to an analyst notebook.
5. Measurement results - Numerical results for the measurement code.
6. Comments - Note important observations and/or corrective actions taken, for example "instrument recalibrated." Comments should be entered when the measurement system is out of control and subsequent corrective action is taken.





The determination of appropriate control limits can be based on the capability of the procedure itself as known from past experience or on the specified requirements of the measurement procedure. Common practice sets control limits at the mean  $\pm 3$  standard deviations. Since the distribution of averages, and many distributions of individual values, exhibit a normal form, the probability of results falling outside the control limits can be readily calculated.

The control chart is actually a graphical presentation of quality control efficiency. If the procedure is "in control," the results will almost always fall within the established control limits. Further, the chart will disclose trends and cycles resulting from assignable causes which can be corrected promptly. Chances of detecting small changes in the process average are improved when the average of several values is used for a single control point, (an  $\bar{X}$  chart). As the sample size increases (for a single  $\bar{X}$  point), the chance that small changes in the average will be detected is increased, provided the subgroup size is not altered as described in the following paragraph.

The basic procedure of the control chart is to compare "within group" variability to "between group" variability. For a single analyst running a procedure, the "within group" may well represent one day's output and the "between group" represents between days or day-to-day variability. When several analysts or several instruments or laboratories are involved, the selection of the subgroup unit is important. Generally speaking, subgroups should be selected in a way that makes each subgroup as homogeneous as possible and that gives the maximum opportunity for variation from one subgroup to another. Assignable causes of variation should then show up as "between group" and not "within group" variability. Thus, if the differences between analysts may be assignable causes of variation, their results should not be lumped together in a "within group" subgrouping. The size of the subgroup is also important. Shewhart

suggested 4 as the ideal size but subgroups of sizes less than 4 are often used in air pollution applications.

## H.2 APPLICATION AND LIMITATIONS

In order for quality control to provide a method for separating the determinate (systematic) from indeterminate (random) sources of variation, the analytical method must clearly emphasize those details which should be controlled to minimize variability. A check list would include:

- Sampling procedures
- Preservation of the sample
- Aliquoting methods
- Dilution techniques
- Chemical or physical separations and purifications
- Instrumental procedures
- Calculating and reporting results.

The next step to be considered is the application of control charts for evaluation and control of the more important of these unit operations. Decisions relative to the basis for construction of a chart are required:

1. Select the variables (unit operations) to be measured
2. Choose method of measurement
3. Select the objective
  - a. Control of variability and/or precision
  - b. Control of bias and/or accuracy of measurements
  - c. Control of completeness of reported data
  - d. Control of percentage of invalid measurements.
4. Select the size and frequency of subgroup samples:
  - a. Size--The analysis will often be dealing with samples of 2 in air pollution applications; process changes are detected with increased probability as the sample size is increased.
  - b. Frequency of subgroup sampling--changes are detected more quickly as the sampling frequency is increased.

5. Control limits can be calculated, but judgment must be exercised in determining whether or not these limits satisfy criteria established for the method, that is, are the limits properly identifying "out of control" points? The control limits (CL's) can be calculated and control charts constructed as described in the following section.

Some of the types of data for which QC charts should be maintained include the following:

1. Zero and span data
2. Repeated analyses of a standard or a control sample
3. Repeated analyses of blank samples
4. Results of audit samples (should separate the results by concentration level)
5. Results for analytical and/or data processing audits (percent difference)
6. Split sample analyses from two labs (if routinely performed)
7. Percent recovery analyses, if routinely performed
8. Percentage of missing data (e.g., percentage of hourly SO<sub>2</sub> data missing relative to total number of hours of data to be obtained)
9. Percentage (or number) of invalid data
10. Average and range/standard deviation of pollutant concentrations for which a QC chart is used as one validation technique
11. Quality cost data (e.g., monthly costs with respect to missing and invalid data, quality control and data validation costs); purpose is to relate prevention costs and "defective" costs.

### H.3 CONSTRUCTION OF CONTROL CHARTS

#### H.3.1 Control Charts for Precision and/or Variability

The use of range (R) in place of sample standard deviation (s) has been justified for sample size  $n \leq 8$  since R is nearly

as efficient as  $s$  for use in estimating  $\sigma$ , and  $R$  is easier to calculate. The latter justification no longer applies, particularly with the availability of the pocket size calculator to potential users of control charts. Hence in this section the use of  $s$  is recommended for sample sizes larger than 2 ( $n > 2$ ); and  $s$  should always be used for  $n > 4$ .

Control Charts Using the Range (R)

The average range ( $\bar{R}$ ) can be calculated from accumulated results, or from a known or assumed  $\sigma$  as ( $d_2\sigma$ ). Values of  $d_2$  are tabulated vs. sample size  $n$  in Table H.1. This table is restricted to  $n \leq 4$  because  $s$  should always be used for larger  $n$ .

TABLE H.1. FACTORS FOR ESTIMATING THE STANDARD DEVIATION  $\sigma$  FROM THE RANGE  $R$

Size of sample	$d_2$	$\frac{1}{d_2}$
2	1.13	0.886
3	1.69	0.591
4	2.01	0.486

If  $\hat{\sigma}$  is given,  $\bar{R}$  can be calculated using  $\bar{R} \approx d_2\hat{\sigma}$ .

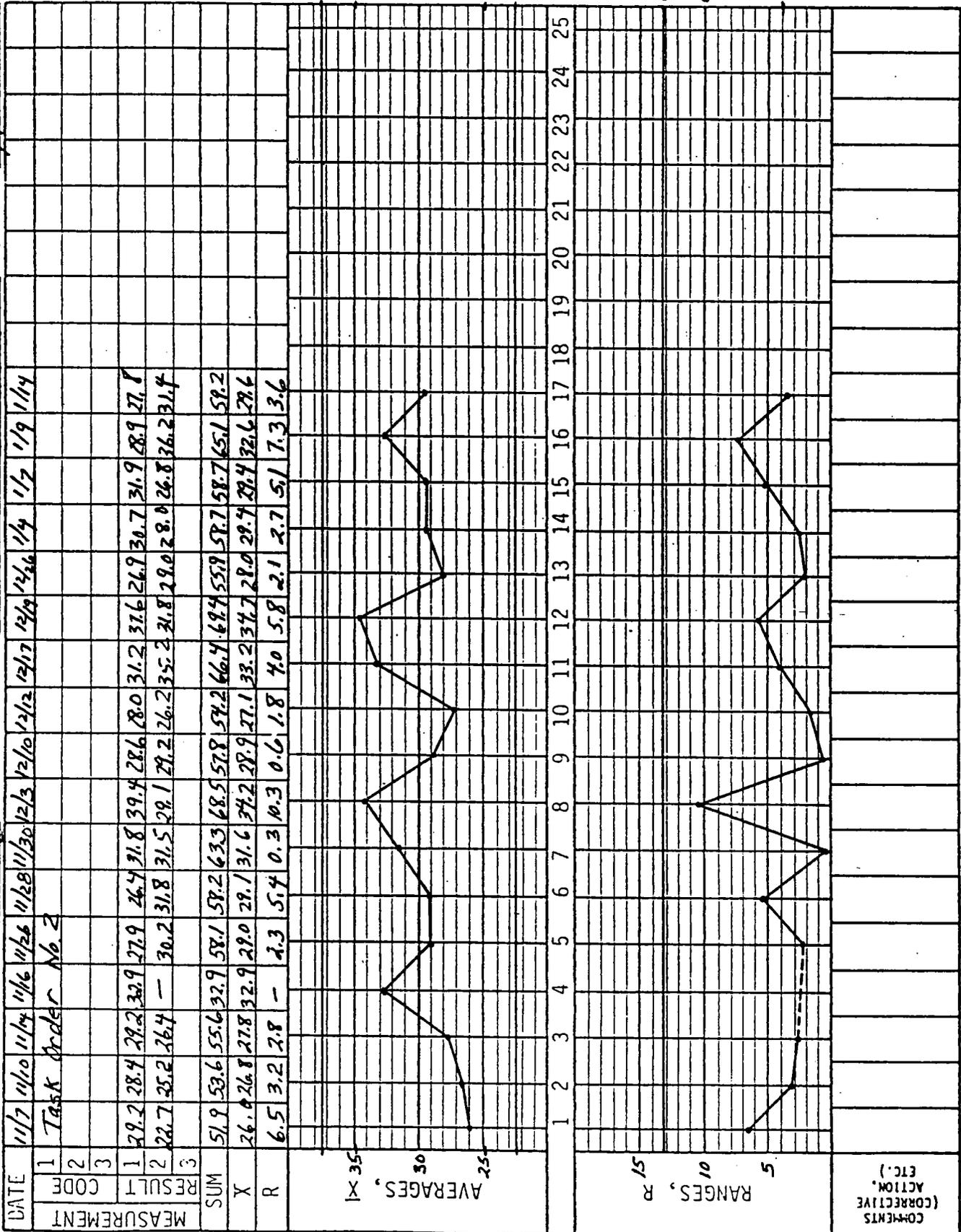
If  $\bar{R}$  is known, an estimate of the standard deviation is  $\hat{\sigma} = \bar{R}/d_2$ .

Example H.1            If  $n = 3$ ,  $\hat{\sigma} = 5$ ,  
 $\bar{R} \approx 1.69(5) = 8.45$ .

Example H.2            If  $n = 2$ ,  $\bar{R} = 3$   
 $\hat{\sigma} = 0.886(3) = 2.66$ .

The steps employed in the construction of a precision control chart using the range are given below and illustrated in Figure H.1, utilizing data on measurements of  $SO_2$  concentrations in Table H.2.

$\bar{X}$  AND R CHART  
 PROJECT Evaluation of Portable  $SO_2$  METERS MEASUREMENT. Barium Chloride Method UNITS PPM



UCL = 37.4  
 UWL = 37.9  
 $\bar{X}$  = 30  
 LWL = 21.9  
 LCL = 22.4

UCL = 13.1  
 UWL = 10.0  
 $\bar{R}$  = 4.0

Figure H.1. Laboratory quality control worksheet.

TABLE H.2. MEASUREMENTS OF SO<sub>2</sub> CONCENTRATIONS - BARIUM  
CHLORANILATE METHOD<sup>a</sup>

Day	Duplicate measurements, ppm		$\bar{X}$	R
1	29.2	22.7	25.95	6.5
2	28.4	25.2	26.80	3.2
3	29.2	26.4	27.80	2.8
4	32.9	----	32.90	---
5	27.9	30.2	29.05	2.3
6	26.4	31.8	29.10	5.4
7	31.8	31.5	31.65	0.3
8	39.4	29.1	34.25	10.3
9	28.6	29.2	28.90	0.6
10	28.0	26.2	27.10	1.8
11	31.2	35.2	33.20	4.0
12	37.6	31.8	34.70	5.8
13	26.9	29.0	27.95	2.1
14	30.7	28.0	29.35	2.7
15	31.9	26.8	29.35	5.1
16	28.9	36.2	32.55	7.3
17	27.8	31.4	29.60	3.6

Subtotals 516.8 470.7 ----- $\Sigma R = 63.8$

$$\Sigma X = 987.5$$

$$\bar{X} = 29.92 \text{ or } 30 \text{ ppm} \quad \bar{R} = 3.988 \text{ or } 4.0 \text{ ppm}$$

$$\hat{\sigma} = \bar{R}/d_2 = 3.988/1.13 = 3.53$$

<sup>a</sup>Source: Parker, Carl D., Research Triangle Institute, NIOSH Report, Evaluation of Portable SO<sub>2</sub> Meters, April 1974.

1. Calculate R for each set of analyses (subgroup)
2. Calculate  $\bar{R}$  from the sum of R values divided by the number (k) of sets (subgroups)
3. Calculate the upper and lower control limits for the range:

$$UCL_R = D_4 \bar{R}, \quad LCL_R = D_3 \bar{R}$$

$$LCL_R = 0 \text{ when } n \leq 6$$

Since the analyses are in duplicates,  $D_4 = 3.27$ ,  $D_3 = 0$ , from Table H.3.

4. Calculate the upper and lower warning limits:

$$UWL_R = \bar{R} + 2\sigma_R = \bar{R} + (2/3) (D_4 \bar{R} - \bar{R}) = D_6 \bar{R}$$

(from Table H.3).

$$LWL_R = 0.$$

5. Chart  $\bar{R}$ ,  $UWL_R$  and  $UCL_R$  on an appropriate scale which will permit addition of new results on a plot such as shown in Figure H.1.

6. Plot results (R) and take action on out-of-control points. (See Subsection H.4).

Example H.3. Compute the  $A_2$  factor for the control limits for samples sizes of  $n = 2$  and  $n = 4$ .

The standard  $3\sigma$  limits for averages of samples of size n are given by

$$\bar{\bar{x}} \pm 3 \sigma_{\bar{x}} = \bar{\bar{x}} \pm 3 \frac{\sigma}{\sqrt{n}},$$

(i.e.,  $\sigma_{\bar{x}}$ , the standard deviation of the sample average  $\bar{x}$ , is  $\sigma/\sqrt{n}$ , the standard deviation of the sampled data divided by  $\sqrt{n}$ ).

An estimate of  $\sigma$  based on the average range  $\bar{R}$  is given by

$$\hat{\sigma} = \bar{R}/d_2$$

where  $d_2$  is obtained from Table H.1. Hence, the limits in terms of  $\bar{R}$  are

$$\bar{\bar{x}} \pm \frac{3}{\sqrt{n}} \frac{\bar{R}}{d_2}.$$

TABLE H.3. FACTORS FOR COMPUTING CONTROL CHART LINES USING  $\bar{R}^a$   
(with example calculations)

Number of observations in subgroup, n	Factor for $\bar{X}$ chart $A_2$	Factors for range chart			
		Control limits		Warning limits <sup>b</sup>	
		$D_3$	$D_4$	$D_5$	$D_6$
2	1.88	0	3.27	0	2.51
3	1.02	0	2.57	0	2.05
4	0.73	0	2.28	0.15	1.85

<sup>a</sup>All factors in Table H.3 are based on the normal distribution.  $A_2$  is used to determine the limits of the  $\bar{X}$  chart and  $D_4$ ,  $D_5$ , and  $D_6$  are used for an R chart as described below.

$$\frac{b/}{D_5} = \frac{5 - 2D_4}{3}, \quad D_6 = \frac{1 + 2D_4}{3}$$

Formulas for calculation

$$\begin{aligned} \bar{R} &= \Sigma R \div k \\ UCL_R &= D_4 \bar{R} \\ LCL_R &= 0 \text{ for } n \leq 6 \\ UWL_R &= D_6 \bar{R} \\ LWL_R &= D_5 \bar{R} \\ \bar{X} &= \Sigma \bar{X} \div k \end{aligned}$$

Example using data of Table H.1

$$\begin{aligned} \bar{R} &= \underline{63.8} \div \underline{16} = \underline{4.0} \\ UCL_R &= \underline{3.27} \times \underline{4.0} = \underline{13.08} \\ LCL_R &= 0 \\ UWL_R &= \underline{2.51} \times \underline{4.0} = \underline{10.4} \\ LWL_R &= \underline{0} \times \underline{4.0} = \underline{0} \\ \bar{X} &= \underline{\quad} \div \underline{\quad} = \underline{\quad} \end{aligned}$$

<sup>c/</sup> or  $\bar{X} = \Sigma X \div$  (total no. of measurements)

$$\begin{aligned} UCL_{\bar{X}} &= \bar{X} + A_2 \bar{R} \\ LCL_{\bar{X}} &= \bar{X} - A_2 \bar{R} \\ UWL_{\bar{X}} &= \bar{X} + \frac{2}{3} A_2 \bar{R} \\ LWL_{\bar{X}} &= \bar{X} - \frac{2}{3} A_2 \bar{R} \end{aligned}$$

$$\begin{aligned} \text{or } \bar{X} &= \underline{987.5} \div \underline{33} = \underline{29.92} \\ UCL_{\bar{X}} &= \underline{29.92} + \underline{1.88} \times \underline{4.0} = \underline{37.44} \\ LCL_{\bar{X}} &= \underline{29.92} - \underline{1.88} \times \underline{4.0} = \underline{22.40} \\ UWL_{\bar{X}} &= \underline{29.92} + \underline{5.01} = \underline{34.93} \\ LWL_{\bar{X}} &= \underline{29.92} - \underline{5.01} = \underline{24.91} \end{aligned}$$

<sup>c/</sup> Use this second form when the numbers of measurements per subgroup differ.

TABLE H.3. FACTORS FOR COMPUTING CONTROL CHART LINES USING R<sup>a</sup>  
(blank data form)

Number of observations in subgroup, n	Factor for $\bar{X}$ chart $A_2$	Factors for range chart			
		Control limits		Warning limits <sup>b</sup>	
		$D_3$	$D_4$	$D_5$	$D_6$
2	1.88	0	3.27	0	2.51
3	1.02	0	2.57	0	2.05
4	0.73	0	2.28	0.15	1.85

<sup>a</sup>All factors in Table H.3 are based on the normal distribution.  $A_2$  is used to determine the limits of the  $\bar{X}$  chart and  $D_4$ ,  $D_5$ , and  $D_6$  are used for an R chart as described below.

$$\frac{b/}{D_5} = \frac{5 - 2D_4}{3}, \quad D_6 = \frac{1 + 2D_4}{3}$$

Formulas for calculation

Example using data of Table H.1

$$\bar{R} = \Sigma R \div k$$

$$UCL_R = D_4 \bar{R}$$

$$LCL_R = 0 \text{ for } n \leq 6$$

$$UWL_R = D_6 \bar{R}$$

$$LWL_R = D_5 \bar{R}$$

$$\bar{\bar{X}} = \Sigma \bar{X} \div k$$

$$\bar{R} = \underline{\quad} \div \underline{\quad} = \underline{\quad}$$

$$UCL_R = \underline{\quad} \times \underline{\quad} = \underline{\quad}$$

$$LCL_R = 0$$

$$UWL_R = \underline{\quad} \times \underline{\quad} = \underline{\quad}$$

$$LWL_R = \underline{\quad} \times \underline{\quad} = \underline{\quad}$$

$$\bar{\bar{X}} = \underline{\quad} \div \underline{\quad} = \underline{\quad}$$

<sup>c/</sup> or  $\bar{\bar{X}} = \Sigma X \div (\text{total no. of measurements})$

$$\text{or } \bar{\bar{X}} = \underline{\quad} \div \underline{\quad} = \underline{\quad}$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{R}$$

$$UCL_{\bar{X}} = \underline{\quad} + \underline{\quad} \times \underline{\quad} = \underline{\quad}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{R}$$

$$LCL_{\bar{X}} = \underline{\quad} - \underline{\quad} \times \underline{\quad} = \underline{\quad}$$

$$UWL_{\bar{X}} = \bar{\bar{X}} + \frac{2}{3} A_2 \bar{R}$$

$$UWL_{\bar{X}} = \underline{\quad} + \underline{\quad} = \underline{\quad}$$

$$LWL_{\bar{X}} = \bar{\bar{X}} - \frac{2}{3} A_2 \bar{R}$$

$$LWL_{\bar{X}} = \underline{\quad} - \underline{\quad} = \underline{\quad}$$

<sup>c/</sup> Use this second form when the numbers of measurements per subgroup differ.

For  $n = 2$ ,  $\bar{\bar{x}} \pm \frac{3}{\sqrt{2}} \frac{R}{1.13} = \bar{\bar{x}} \pm 1.88 \bar{R}$ ;  $A_2 = 1.88$  for  $n = 2$ .

For  $n = 4$ ,  $\bar{\bar{x}} \pm \frac{3}{\sqrt{4}} \frac{R}{2.06} = \bar{\bar{x}} \pm 0.73 \bar{R}$ ;  $A_2 = 0.73$  for  $n = 4$ .

### Control Charts Using the Standard Deviation (s)

For  $n = 2$ ,  $s = R/\sqrt{2}$  and hence there is no preference in using the range (R) or the standard deviation (s) from the statistical viewpoint. The range would be slightly easier to use since it would be necessary to divide the ranges by  $\sqrt{2}$ . Because of the ease of computing the standard deviation with the preprogrammed calculators, it is recommended that s be used for  $n > 2$  with the option of using R or s with  $n = 2$ . The procedure for using s to construct the control chart is given below.

1. Calculate  $\bar{X}$  for each sample (subgroup) and  $\bar{\bar{X}}$  for all samples.
2. Calculate s for each sample and  $\bar{s} = \Sigma s/k$  for all sets. This computation assumes equal sample sizes (n) for each of the k sets. See Reference 1 for unequal sample sizes.
3. Calculate the upper and lower control limits for s by using the equations

$$\begin{aligned}UCL_s &= B_4 \bar{s} \\LCL_s &= B_3 \bar{s}.\end{aligned}$$

The factors  $B_3$  and  $B_4$  are tabulated in Table H.4 and the control chart based on s is in Figure H.2.

### H.3.2 Control Charts for Averages ( $\bar{X}$ charts) - Mean or Nominal Value Basis

As previously stated, the control chart based on the range should only be used for  $n = 2$  and the chart based on the standard deviation s may be used for all n.

TABLE H.4.\* FACTORS\* FOR COMPUTING CONTROL CHART  
LINES USING  $\bar{s}$

Number of observations in subgroup $n$	Factor for $\bar{X}$ chart $A_1$	Factors for $\bar{s}$ chart	
		Lower control limit $B_3$	Upper control limit $B_4$
2	2.66	0	3.27
3	1.95	0	2.57
4	1.63	0	2.27
5	1.43	0	2.09
6	1.29	0.03	1.97
7	1.19	0.12	1.88
8	1.09	0.19	1.81
9	1.03	0.24	1.76
10	0.98	0.28	1.72
11	0.92	0.32	1.68
12	0.89	0.35	1.65
13	0.85	0.38	1.62
14	0.82	0.41	1.59
15	0.79	0.43	1.57
16	0.76	0.45	1.55
17	0.74	0.47	1.53
18	0.72	0.48	1.52
19	0.70	0.50	1.50
20	0.68	0.51	1.49

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_1 \bar{s}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_1 \bar{s}$$

$$UCL_s = B_4 \bar{s}$$

$$LCL_s = B_3 \bar{s}$$

$\bar{s}$  = average standard deviation for  $k$  groups of  $n$  observations each, i.e.,  $\bar{s} = \sum s/k$

\* All factors in Table H.4 are based on the normal distribution

Source: Grant, E. I. and Leavenworth, R. S., Statistical Quality Control, 4th Edition, McGraw-Hill Book Co., New York, p. 646, part of Table D.

$$A_1 = \sqrt{\frac{n-1}{n}} A_1 \text{ (as given in Table from Grant and Leavenworth)}$$

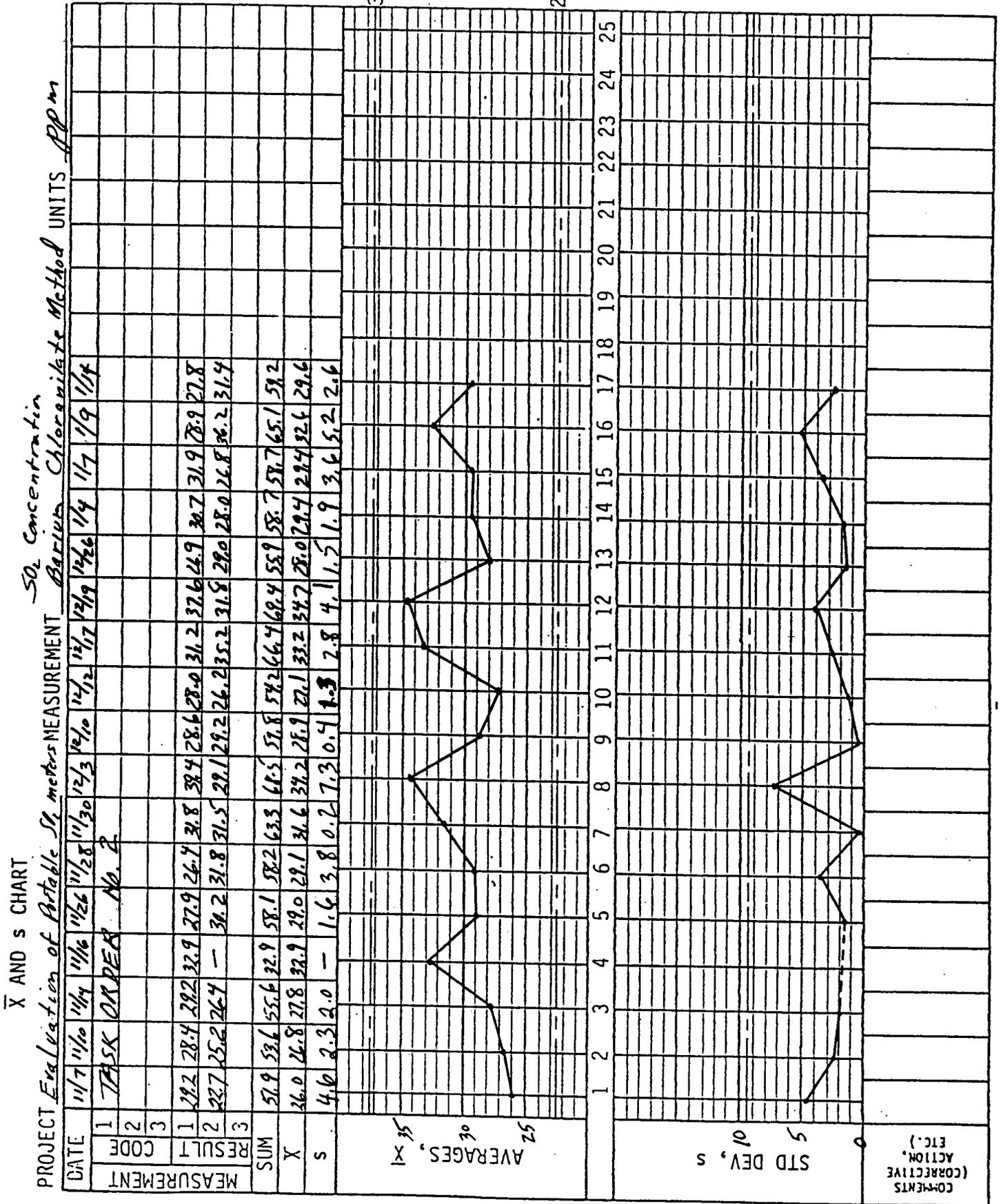


Figure H.2. X and s charts for data of Table H.2.

Control Charts Using the Range (R)

$\bar{X}$  charts simplify and render more exact the calculation of control limits since the distribution of data which conforms to the normal curve can be completely specified by  $\bar{X}$  ( $\cong \mu$ ) and  $\sigma$ . Step-by-step construction of an  $\bar{X}$  (and R) control chart based on duplicate sets of results obtained from consecutive analyses of a control sample serves as an example (bottom of Table H.3 with example calculations).

1. Calculate  $\bar{X}$  for each duplicate set and  $\bar{\bar{X}}$  for all sets.
2. Calculate R for each duplicate set and  $\bar{R}$  for all sets.
3. Calculate the upper and lower control limits by the equations:

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{R}, \quad LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{R}$$

Values of  $A_2$  vs. n are tabulated in Table H.3.

4. Calculate the upper and lower warning limits by the equations:

$$UWL_{\bar{X}} = \bar{\bar{X}} + (2/3)A_2 \bar{R}, \quad LWL_{\bar{X}} = \bar{\bar{X}} - (2/3) A_2 \bar{R}$$

5. Construct the lines corresponding to  $\bar{\bar{X}}$ ,  $UCL_{\bar{X}}$ ,  $LCL_{\bar{X}}$ , and if desired,  $UWL_{\bar{X}}$  and  $LWL_{\bar{X}}$  as shown in Figure H.1.
6. Plot  $\bar{X}$  for each sample and take appropriate action on points which fall outside of the warning limits.

Control Charts Using the Standard Deviation (s)

The procedure for constructing control charts for the mean based on the sample standard deviation is detailed below:

1. Calculate  $\bar{X}$  for each sample (subgroup) and  $\bar{\bar{X}}$  for all samples
2. Calculate s for each sample and the average  $\bar{s} = \Sigma s/k$  for all samples
3. Calculate the upper and lower control limits using the equations

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_1 \bar{s}, \quad LCL_{\bar{X}} = \bar{\bar{X}} - A_1 \bar{s}$$

where  $A_1$  is read from Table H.4.

4. Calculate the upper and lower warning limits (if desired) using the equations

$$UWL_{\bar{X}} = \bar{X} + \frac{2}{3} A_1 \bar{s}$$

$$LWL_{\bar{X}} = \bar{X} - \frac{2}{3} A_1 \bar{s}$$

See Figure H.2 for an illustration of this procedure.

### H.3.3 Control Charts for Percent of Defective Measurements

One type of control chart which is frequently used in industrial quality control is that for the fraction of defects in a production process. It is desired to maintain a low percentage of defects in order to minimize the expense of rework or waste. A comparable problem for air pollution data is the following: suppose that  $n$  hi-vol filters are visually checked for defects and the number of defective filters is  $d$ , then the fraction of defectives is  $p = d/n$ .

In the case of a measurement process, one may be concerned with, for example, the number of invalid measurements among those reported, or the number of missing values relative to the number of measurements to be taken. In these examples the true or average fraction of measurements which are missing or invalid will be denoted by  $\bar{p}$ . The observed fraction of invalid or missing data will be  $p$ ; then limits for  $p$  can be obtained as follows:

1. Assume that  $\bar{p}$  has been determined on the basis of recent history of a process

2. Calculate an upper control limit for  $p$  using

$$\bar{p} + 3 \left[ \frac{\bar{p}(1-\bar{p})}{n} \right]^{1/2}$$

3. Calculate a lower control limit for  $p$  using

$$\bar{p} - 3 \left[ \frac{\bar{p}(1-\bar{p})}{n} \right]^{1/2}$$

4. Draw lines,  $UCL_p$ ,  $LCL_p$  on a control chart with suitable scale for vertical axis to include expected range of variation of p.

5. Plot the fraction defective p for each sample along with the appropriate control limits. An example of such a control chart is given in Figure H.3 with the data taken from Table H.5. The sample size n is varied in order to indicate how such a control chart is constructed. A similar approach would be applicable to the  $\bar{X}$  and R charts using appropriate n.<sup>1</sup>

TABLE H.5. COMPUTATION OF CONTROL LIMITS FOR FRACTION DEFECTIVE

$$\bar{p} = 0.05, \sigma_{100}\{p\}^* = \left[ \frac{(0.05)(0.95)}{100} \right]^{1/2} \cong 0.022; \sigma_{45}\{p\} = 0.0325; \sigma_{50}\{p\} = 0.0308$$

Sample number	Number of measurements (n)	Number of defective measurements (d)	Fraction defective measurements (p=d/n)	UCL $\bar{p} + 3\sigma_n\{p\}$	LCL $\bar{p} - 3\sigma_n\{p\}$
1	50	1	0.02	0.142	0
2	50	3	0.06	0.142	0
3	45	2	0.044	0.148	0
4	100	6	0.06	0.116	0
5	45	4	0.088	0.148	0
6	50	3	0.06	0.142	0
7	45	0	0	0.148	0
8	45	3	0.066	0.148	0
9	50	2	0.04	0.142	0
10	100	7	0.07	0.116	0
11	45	2	0.044	0.148	0
12	45	1	0.022	0.148	0
13	45	3	0.066	0.148	0
14	50	4	0.08	0.142	0
15	45	1	0.022	0.148	0

$$*\sigma_n\{p\} = \left[ \frac{\bar{p}(1-\bar{p})}{n} \right]^{1/2}$$



#### H.3.4 Control Charts for Individual Results

In many instances a rational basis for subgrouping may not be available, or the analysis may be so infrequent as to require action on the basis of individual results. In such cases, charts of individual values  $X$  are employed. A control chart for individuals has the advantage of displaying each result with respect to its specification limits (Figure H.4). The disadvantages must be recognized when considering this approach.

1. Changes in dispersion are not detected unless an  $R$  (or  $s$ ) chart is included using the moving range (or standard deviation), see Section H.3.6.

2. The distribution of results must be approximately normal for the control limits to be valid using the standard normal table. Of course, other distributions can be used rather than the normal, (e.g., lognormal or Weibull).

#### H.3.5 Control Chart for Signed % Difference

One calculation which is performed frequently is the signed % difference (e.g., if  $Y_i$  is a routinely measured value and  $X_i$  is an audited value, then

$$d_i = 100 \frac{Y_i - X_i}{X_i}$$

is the signed % difference).

As an example consider the following TSP data obtained using one pair of collocated hi-vol samplers.

$\bar{X}$  AND R CHART

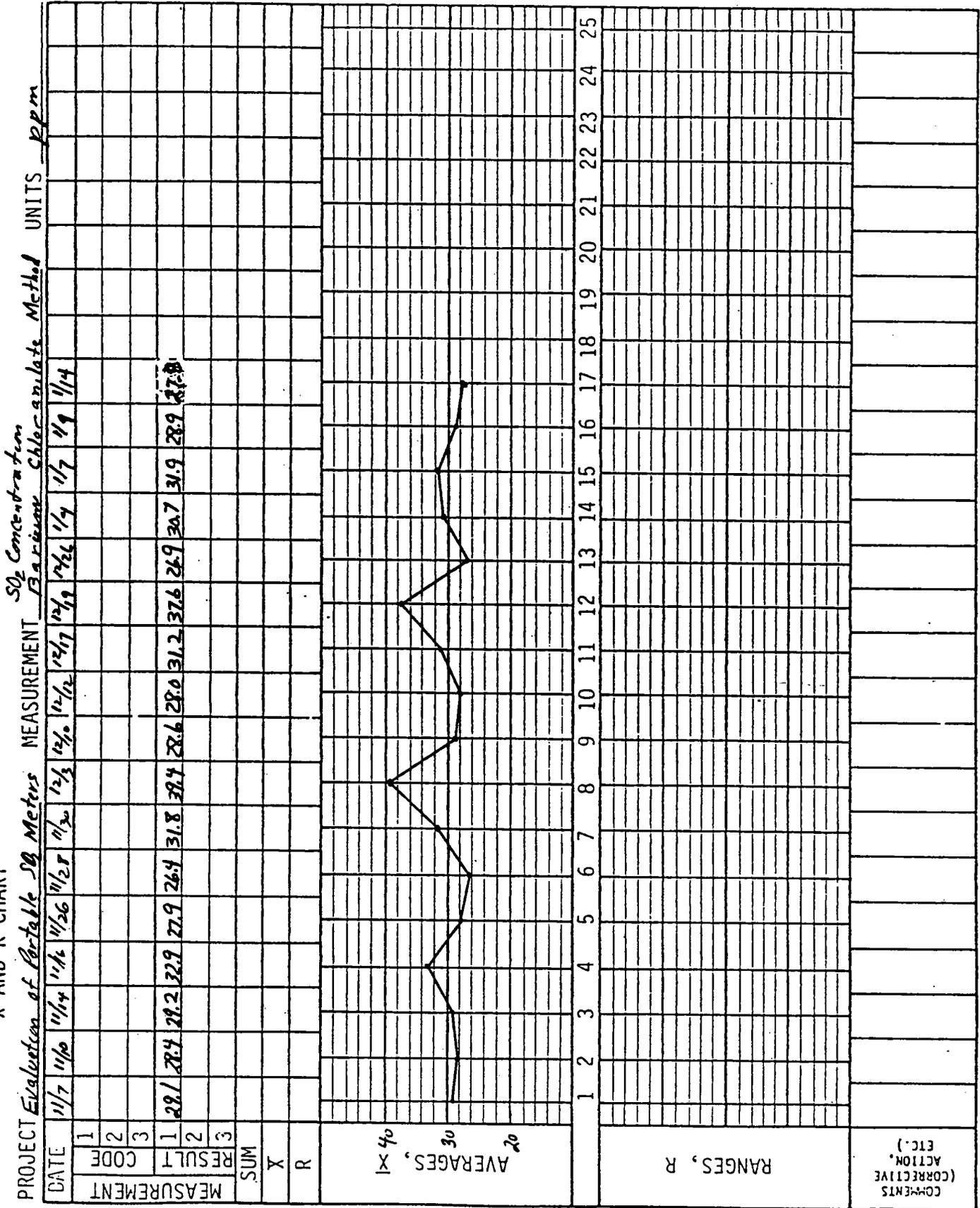


Figure H.4. Control chart for individual values in Table H.6.

Sampling period	Duplicate sampler ( $Y_i$ ), $\mu\text{g}/\text{m}^3$	Official sampler ( $X_i$ ), $\mu\text{g}/\text{m}^3$	Difference ( $d_i$ ), %
1	53.0	51.9	2.1
2	69.9	66.6	5.0
3	-	67.8	-
4	58.4	55.7	4.8
5	48.5	46.4	4.5
6	61.6	58.9	4.6
7	57.9	59.0	-1.9
8	67.5	64.2	5.1
9	58.0	55.4	4.7
10	55.0	59.0	-6.8
11	60.0	58.1	3.3
12	55.8	53.1	5.1
13	51.4	52.8	-2.7
			$\Sigma d_i = 27.8$

The average  $\bar{d}$  and the standard deviation of the  $d_i$  are:

$$\bar{d} = 27.8/12 = 2.3\%$$

$$s = \left[ \frac{236 - (27.8)^2/12}{11} \right]^{1/2} = 4.0\%$$

These values are used to obtain a control chart (Figure H.5) for individual values of  $d_i$  as follows:

$$\bar{d} \pm 3s = 2.3 \pm 3(4) = (-9.7, 14.3).$$

### H.3.6 Moving Averages and Ranges

The  $\bar{X}$  control chart is more efficient than an  $X$  chart for individual values for moderate changes in the mean as the subgroup size increases. A logical compromise between the  $X$  and  $\bar{X}$  approach would be application of the moving average. The moving averages are obtained for the data given in the first column of Table H.2, and presented in Table H.6.

The moving average and range of two observations are obtained for each successive pair; that is, observations 1 and 2, 2 and 3, etc.

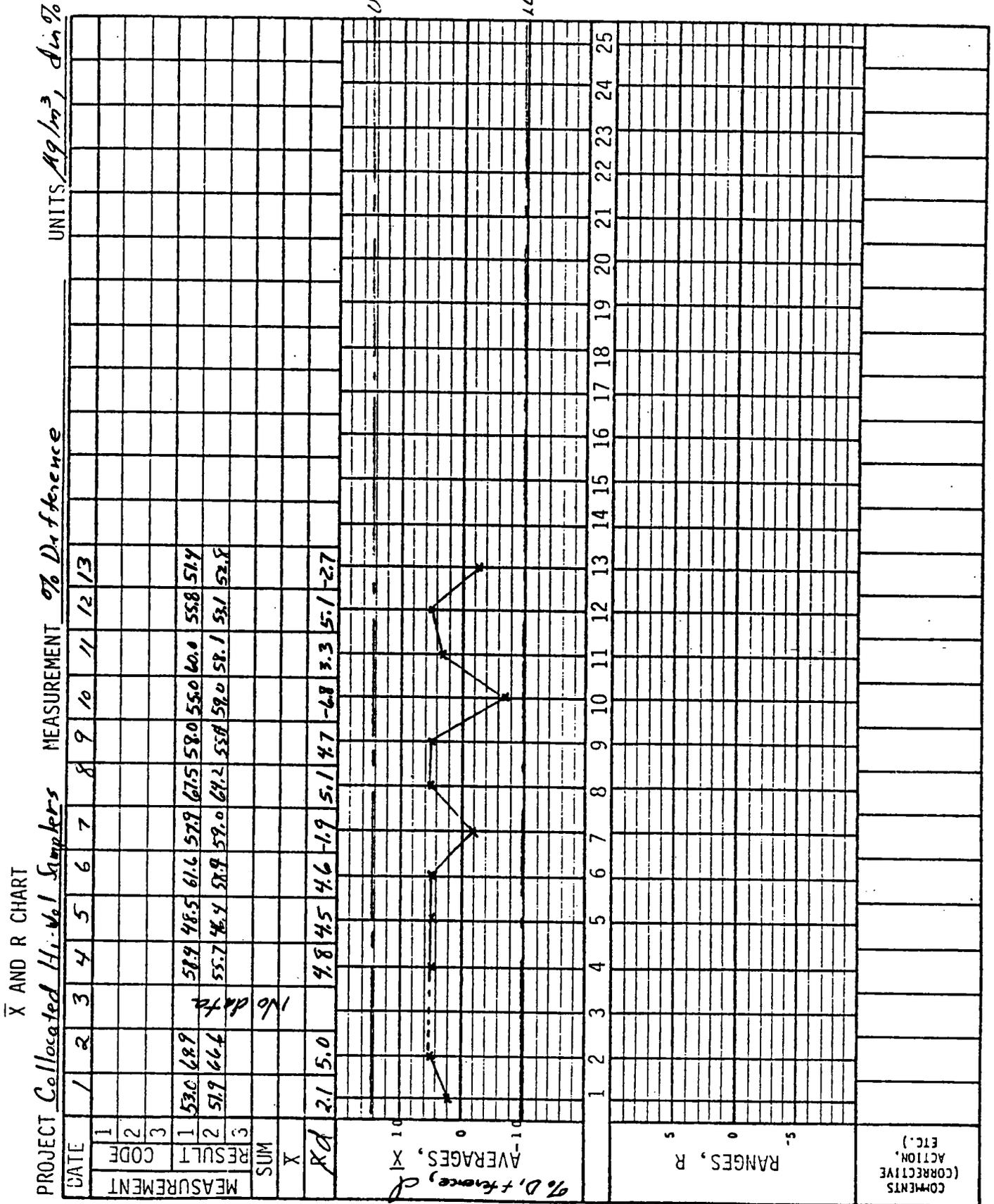


Figure H.5. Control chart for signed % difference.

The moving range serves well as a measure of variation when no rational basis for subgrouping is available or when results are infrequent or expensive to gather.

Control charts can be constructed with the use of moving averages and ranges in the same manner as for ordinary charts. That is, one computes the grand average  $\bar{\bar{X}}$  and  $\bar{\bar{R}}$  and then computes the limits. These control charts would only be approximate ones for they do not take into consideration the correlation between the successive observations. If these values are taken close in time they may be highly correlated, (e.g., the autocorrelation with lag 1 [successive values] may be 0.5 to 0.7).

$\bar{X}$  Chart (moving average of two consecutive values)

---

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{\bar{R}}, \quad A_2 = 1.88 \text{ for } n = 2$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{\bar{R}}$$

Range (R) Chart (moving range of two consecutive values)

---

$$LCL_R = 0$$

$$UCL_R = D_4 \bar{\bar{R}} \quad D_4 = 3.27 \text{ for } n = 2$$

---

The interpretation of the first point "out of control" is the same as for an ordinary chart. However, because successive plotted points use a common value and are thus correlated, two consecutive points out of control on a moving average or range chart can be due to a single value out of control on an X chart where each measurement corresponds to only one point.

For further details on the moving average charts one is referred to Grant and Leavenworth,<sup>1</sup> (pages 177 to 182). Because of their potential importance to air pollution measurements, an example application of the construction of such charts is given in Figure H.6 for moving averages and ranges of two consecutive observations.

X AND R CHART

PROJECT Evaluation of Portable SO<sub>2</sub> meters MEASUREMENT Berium Chloranilate Method UNITS PPM

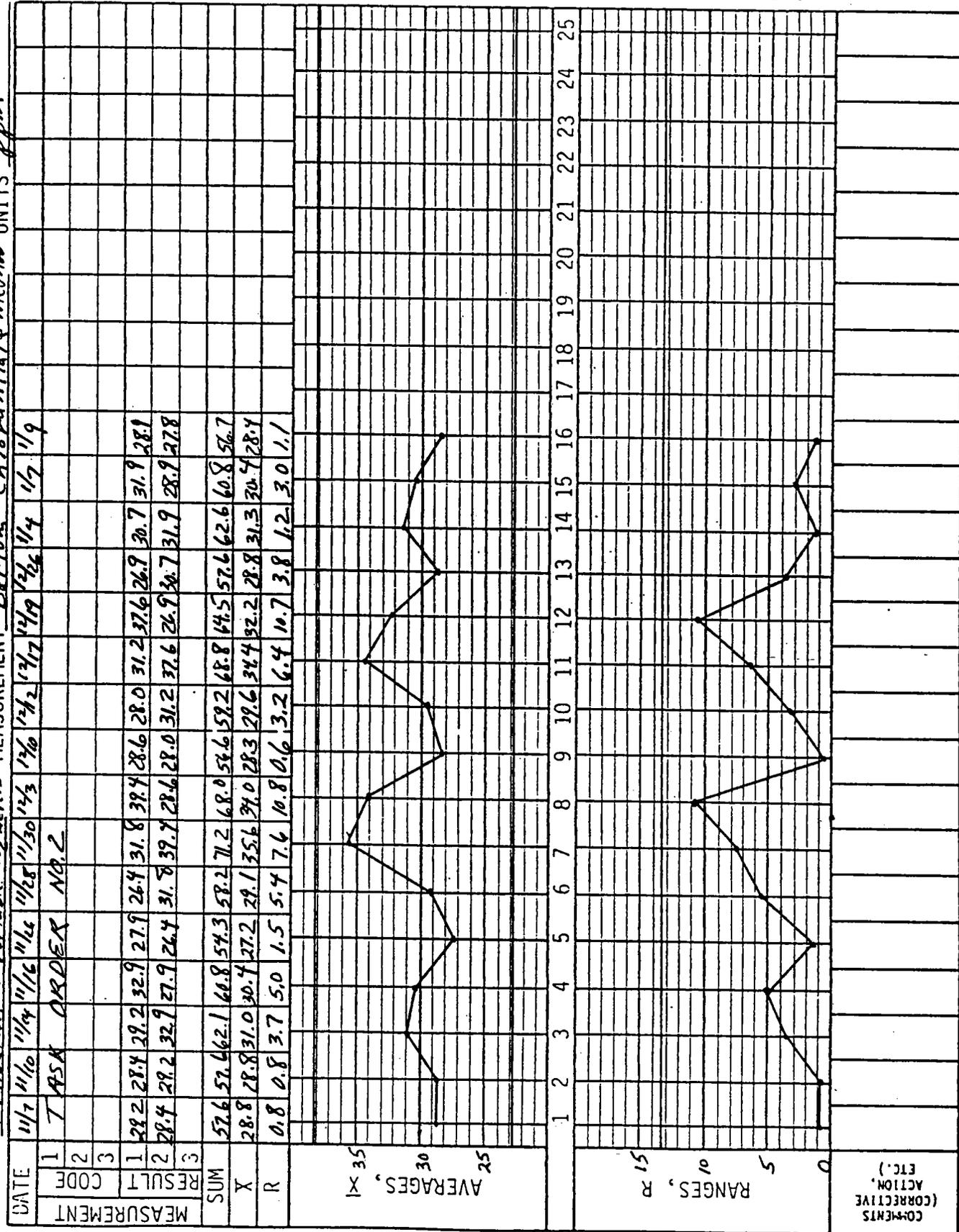


Figure H.6. X and r chart for moving averages and ranges data from Table H.6.

TABLE H.6. MOVING AVERAGE AND RANGE TABLE (n=2)

Sample number	Value	Moving averages of 2	Moving range
1	29.2	---	---
2	28.4	28.80	0.8
3	29.2	28.80	0.8
4	32.9	31.05	3.7
5	27.9	30.40	5.0
6	26.4	27.15	1.5
7	31.8	29.10	5.4
8	39.4	35.60	7.6
9	28.6	34.00	10.8
10	28.0	28.30	0.6
11	31.2	29.60	3.2
12	37.6	34.40	6.4
13	26.9	32.25	10.7
14	30.7	28.80	3.8
15	31.9	31.30	1.2
16	28.9	30.40	3.0
17	27.8	28.35	1.1
Totals:		488.3	65.6
Averages:		$\bar{X} = 30.52$	$\bar{R} = 4.1$

Refer to Table H.3

$$UCL_R = D_4 \bar{R} = 3.27 \times 4.1 = 13.4$$

$$LCL_R = D_3 \bar{R} = 0$$

$$UCL_{\bar{X}} = \bar{X} + A_2 \bar{R} = 30.5 + 1.88 \times 4.1 = 38.2$$

$$LCL_{\bar{X}} = \bar{X} - A_2 \bar{R} = 30.5 - 1.88 \times 4.1 = 22.8$$

These limits are plotted on Figure H.6.  
Warning limits could be computed in a similar manner.

### H.3.7 Other Control Charts

Although the standard  $\bar{X}$  and R control charts are the most common, they do not always do the best job. Several examples follow where other charts are more applicable.

H.3.7.1 Variable Subgroup Size - The standard  $\bar{X}$  and R charts are applicable for a constant size subgroup on n. If n varies control limit values must be calculated for each sample size. Plotting is done in the usual manner with the control limits drawn in for each subgroup depending on its size. Plotting may not be practical if the size of the subgroup varies a great deal; in this case a tabular calculation would be appropriate. An example with varying control limits was described under control charts for fraction-defective measurements.

H.3.7.2  $\sigma$  as Function of the Mean - When the standard deviation is a function of concentration, control limits can be expressed in terms of a percentage of the mean. In practice such control limits would be given as in the example below.

±5 units/liter for 0-100 units/liter concentration  
±5% for > 100 units/liter concentration

An alternative procedure involves transformation of the data. For example, logarithms would be the appropriate transformation when the standard deviation is proportional to the mean.

The frequent use of the relative standard deviation (RSD) and the percent difference (d) in air pollution measurements suggests that it may be desirable to construct control charts for the sample  $RSD = s/\bar{X}$  rather than transform the data. Using the results of Iglewicz and Myers<sup>2</sup> in comparing several approximations, the percentiles of the distribution of the sample RSD were obtained as a function of RSD of the population,  $RSD'$ .

To simplify the use of the limits for the reader, Figure H.7 contains the appropriate limits for sample sizes  $n = 2, 4$  and  $7$ , and for 95 and 99.5 probability limits (corresponding to approximately  $2\sigma$  and  $3\sigma$  limits).

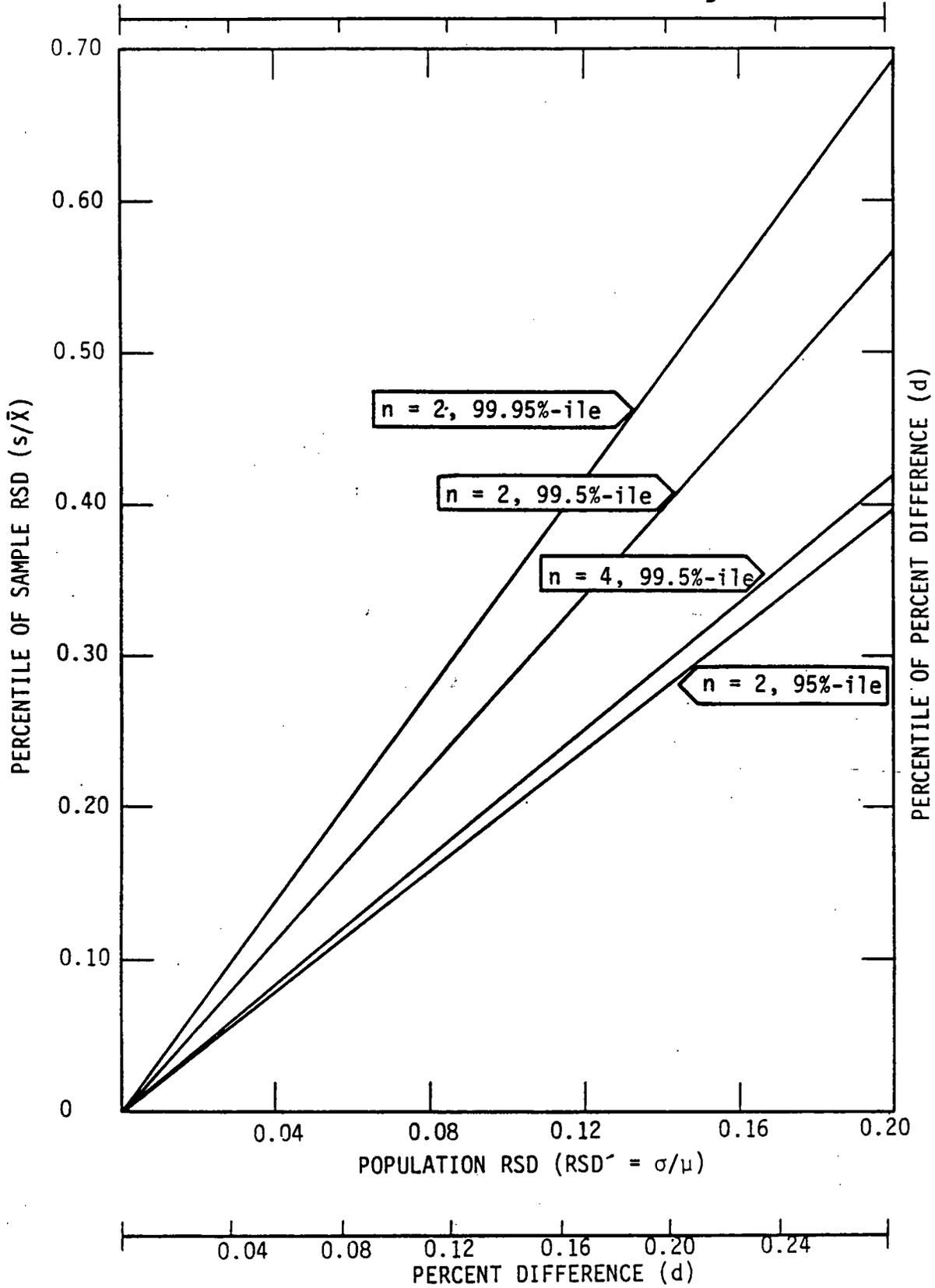


Figure H.7. Limits for use in constructing control charts for RSD and d.

As an example of how the limits may be calculated for a control chart for the sample RSD =  $s/\bar{X}$ , suppose that the RSD is known to be 0.04 when the method is in a state of control. For a sample of size two ( $n = 2$ ), the 99.5th percentile for the sample RSD ( $= s/\bar{X}$ ) is read from Figure H.7 to be about 0.113 or 11.3 percent. Hence, a control chart can be constructed with mean line at 0.04 or 4 percent and the upper control limit at 0.113 or 11.3 percent. If an observed RSD exceeds 0.113 it is indicative of a possible lack of control of the RSD of the measurement.

If one is using the percent error measurement,

$$d = \frac{|X_1 - X_2| \times 100}{(X_1 + X_2)/2}$$

for  $n = 2$ , then  $d = \sqrt{2}$  RSD. Hence the limits are multiplied by  $\sqrt{2}$  or the upper 99.5 percentile for  $d$  for  $n = 2$  and RSD = 0.04, becomes 0.160 or 16 percent and the mean value would be  $0.04 \times \sqrt{2} = 0.056$  or 5.6 percent. The 99.5 percentile is used as an approximation of the upper  $3\sigma$  control limit which corresponds to the 99.86th percentile.

H.3.7.3 Cusum Charts - The cumulative sum (quality control) chart has the advantage of identifying small persistent changes in the sampling/analytical process faster than the standard quality control chart using the  $3\sigma$  limits.<sup>3,4,5</sup> This is an advantage for processes requiring tight control but is a disadvantage otherwise. If the standard ( $3\sigma$ ) control chart is augmented by either warning lines, or a run test, for example, the chart efficiency approaches that of a cusum chart. Some disadvantages of the cusum chart are: (1) complicated calculations, (2) not as efficient as a  $3\sigma$  control chart in identifying a single change in the process (however, it is possible to augment the cusum chart to remove this disadvantage), (3) most charts use a cumbersome, movable V-mask to determine control.

The cusum charts would be useful in the application of quality control techniques to zero and span checks. An alternative would be to apply the usual  $3\sigma$  chart with the added features of warning  $2\sigma$  limits and/or a run test (e.g., a run of seven values above or below the mean line or an upward or downward trend of seven consecutive values).

#### H.4 INTERPRETATION OF CONTROL CHARTS FOR OUT-OF-CONTROL

Various criteria have been used<sup>6</sup> to determine when the measurement system is out-of-control. The more important criteria for out-of-control are as follows:

1. One or more points outside the control limits ( $3\sigma$ ).
2. A run of 2 or more points outside the warning limits ( $2\sigma$ ).
3. A run of 7 or more points (i.e., seven consecutive points with a common property--e.g., larger than a given value such as the mean). This might be a run up or run down or simply a run above or below the central line ( $\bar{X}$ ) on the control chart.
4. Cycles or non-random patterns in the data. Such patterns may be of great help to the experienced operator.<sup>7</sup> For example, the measurement may be subject to diurnal variation due to sensitivity of the measurement method to temperature variations.

#### H.5 STEPS IN DEVELOPING AND USING A CONTROL CHART SYSTEM

The following is the logic sequence which could be followed in developing and using a control chart system.

1. Determine which key data to chart. Obviously, all available data cannot be plotted. Only the more important data should be plotted, such as results of calibrations, checks of standards or blinds, or duplicate checks.
2. Decide what statistic to plot.
  - (a) Control of mean, plot  $\bar{X}$  or  $X$ .
  - (b) Control of variability, plot  $R$  or  $s$ .

3. Evaluate the form of the distribution. Determine the form of the distribution (i.e., whether the distribution is normal, lognormal, or as otherwise assumed).
4. Eliminate any outliers from the sample of past data. Obviously, the existence of out-of-control points should not be used in the establishment of control limits.
5. Determine the warning limits (2-sigma), if appropriate, and the control limits (3-sigma).
6. Use control chart form on page 2 or 3 and prepare any additional instructions for recording and plotting data and for taking action when out-of-control conditions are indicated.
7. Draw in central line and control limits with bold lines. Where specification limits exist, these may also be shown.
8. Maintain charts in the working area, if possible. Where possible and practicable, the operator should record the information and data, perform the necessary computations and plot the charts. (This would not be feasible when the results are for 'blinds' or unknown to the operator.) In many cases the charts should be posted on a wall or otherwise kept in view in the working area. Some charts may be kept in a loose-leaf binder by the supervisor.
9. Plot the points in bold fashion and join adjacent points by a straight line. It is very important that data be plotted as soon as they become available in order that out-of-control conditions can be detected as early as possible and that timely corrective action can be taken.
10. Circle or otherwise highlight out-of-control points or conditions. It is also desirable to indicate visible trends on the charts.
11. Take appropriate corrective actions when out-of-control conditions are indicated. Record on the chart the nature (and time) of the corrective action.

12. Revise control limits periodically. When recent past history indicates an improvement in control, the revised limits will be tighter than previously. Good justification should be made before relaxing limits.

13. Maintain an historical file:

a. Data and computations used for determining control limits.

b. Plotted control charts.

#### H.6 REFERENCES

1. Grant, E. I., and Leavenworth, R. S., Statistical Quality Control, Fourth Edition, McGraw-Hill Book Co. New York, 1972.
2. Iglewicz, B., and Myers, R. H., Comparison of Approximations to the Percentage Points of the Sample Coefficients of Variation, Technometrics, 12 1:166-170. February 1970.
3. Quality Assurance Practices for Health Laboratories, pp. 838-843. Stanley L. Inham, MD Editor, American Public Health Association, Washington, D.C. 1978.
4. Roberts, S. W. A Comparison of Some Control Chart Procedures. Technometrics, Vol. 8, No. 3. August 1966.
5. Lucas, J. M. A Modified "V" Mask Control Scheme. Technometrics, Vol. 15, No. 4. November 1973.
6. Duncan, A. J. Quality Control and Industrial Statistics, Third Edition, Richard D. Irwin, Inc., Homewood, Illinois, 1965.
7. Industrial Hygiene Laboratory Accreditations 589, NIOSH Quality Assurance Training Manual.

## APPENDIX I

### STATISTICAL SAMPLING

#### I.1 CONCEPTS

Suppose that filters to be used in hi-vol samplers are received in lots of size  $N = 100$  and that it is desired for a special project to use only filters with pH between two specified values, for example,

$$6.5 \leq \text{pH} \leq 7.5.$$

If the pH test destroys a filter then it is necessary to employ some sampling procedure to determine whether the lot of  $N = 100$  filters should be accepted for use consistent with prescribed specification limits such as given above. A sample of  $n$  filters is selected at random (this will be discussed in the next section), and if the number of defective filters,  $d$ , (i.e., with pH below 6.5 or above 7.5) exceeds a preselected value,  $c$ , the lot of filters is rejected. That is, it is presumed that the quality of the lot is not consistent with the desired specifications. This procedure is referred to in statistical and quality control texts as acceptance sampling by attributes. The approach can be applied to the acceptance of any product ordered in lots, and subject to desired specification levels. In some cases, 100% inspection is obviously impossible; while in other cases of nondestructive testing, 100% inspection is possible but may not be practicable because of cost or inspection fatigue and consequently an increased risk of the inspector misclassifying the item.

#### I.2 HOW DOES ONE SELECT A RANDOM SAMPLE, STRATIFIED RANDOM SAMPLE?

##### I.2.1 Random Sample

Assume that a lot of  $N = 100$  items is received, and a random sample of size  $n = 10$  is to be drawn. First of all, what

does one mean by a random sample of  $n = 10$ ? This can be defined as a sample of 10 items selected in such a manner that every possible sample of size 10 has an equal chance of being selected. The sample can be selected with replacement (putting an item back after selection and inspection) or without replacement. Both cases will be considered. A strictly random procedure is as follows:

1. Number the items in the lot from 1 to 100 (this does not have to be done by marking, but merely assign an ordering of the items to be sampled. For example, if they are in a single stack the top item can be taken as 1, and the bottom item as 100 (corresponding to 00 in the random number table).

2. Use a table of random numbers or a random number generator such as a deck of cards numbered 0, 1, 2, ... 9.

3. From the table of random numbers select a two digit number at random, often done by closing one's eyes and placing the finger on the page and identifying the closest number pair just above the finger. See Table I.1 as an example.

4. Record the number obtained in 3 above and the following nine two-digit numbers, for example, 59, 11, 13, 99, 93, 19, 78, 83, 72, 62. (See Table I.1 for 59 in a rectangular box.)

5. If there are any repeats, select another number in case of sampling without replacement; otherwise the ten numbers identify the items to be selected from the lot and checked against the specifications.

6. If one wishes to use a deck of ten cards numbered 0 to 9 instead of a random number table, one card can be selected at random, after shuffle say, replaced and followed by a second drawing after a reshuffle. The two digits yield a two-digit number between 00 and 99 and hence determine the item to be selected from the lot.

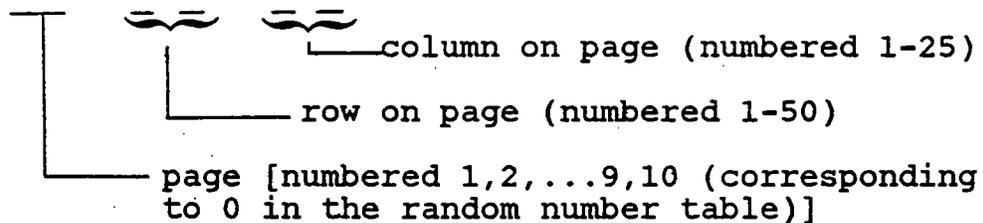
TABLE I-1. SHORT TABLE OF RANDOM NUMBERS

57	23	06	26	23	08	66	16	11	75	28	81	56	14	62	82	45	65	80	36	02	76	55	01	
87	78	16	06	57	12	46	22	90	97	73	67	39	06	63	60	51	02	07	16	75	12	89	41	17
25	71	15	08	82	64	87	29	01	20	46	72	05	80	19	27	47	15	76	51	58	67	06	50	74
42	67	93	41	67	44	28	71	45	08	19	47	76	30	26	72	33	69	92	51	95	23	20	80	76
05	83	93	84	32	62	83	27	48	83	09	19	84	90	20	20	50	87	74	93	51	62	10	23	30
60	46	18	41	23	74	73	51	72	90	40	52	95	41	20	89	48	98	27	38	81	33	83	82	94
32	80	64	75	91	98	09	40	64	89	29	99	46	35	69	91	50	73	75	92	90	56	82	93	24
79	86	53	77	78	06	62	37	48	82	71	00	78	21	65	65	88	45	82	44	78	93	22	78	09
45	13	23	32	01	09	46	36	43	66	37	15	35	04	88	79	83	53	19	13	91	69	81	81	87
20	60	97	48	21	41	84	22	72	77	99	81	83	30	46	15	90	26	51	73	66	34	99	40	60
67	91	44	83	43	25	56	33	28	80	99	53	27	56	19	80	76	32	53	95	07	53	09	61	98
86	50	76	93	86	35	68	45	37	83	47	44	92	57	66	59	64	16	48	39	26	94	54	66	40
66	73	38	38	23	36	10	95	16	01	10	01	59	71	55	99	24	88	31	41	00	73	13	80	62
55	11	50	29	17	73	97	04	20	39	20	22	71	11	43	00	15	10	12	35	09	11	00	89	05
23	54	33	87	92	92	04	49	73	96	57	53	57	08	93	09	69	87	83	07	46	39	50	37	85
41	48	67	79	44	57	40	29	10	34	58	63	51	18	07	41	02	39	79	14	40	68	10	01	61
03	97	71	72	43	27	36	24	59	88	82	87	26	31	11	44	28	58	99	47	83	21	35	22	88
90	24	83	48	07	41	56	68	11	14	77	75	48	68	08	90	89	63	87	00	06	18	63	21	91
98	96	97	42	27	11	80	51	13	13	03	42	91	14	51	22	15	48	67	52	09	40	34	60	85
74	20	94	21	49	96	51	69	99	85	43	76	55	81	36	11	88	68	32	43	08	14	78	05	84
94	67	48	87	11	84	00	85	93	56	43	99	21	74	84	13	56	41	90	96	30	04	19	68	73
58	18	84	82	71	23	66	33	19	25	65	17	90	84	24	91	75	36	14	83	86	22	70	86	89
31	47	28	24	88	49	28	69	78	62	23	45	53	38	78	65	87	44	91	93	91	62	76	09	20
45	62	31	06	70	92	73	27	83	57	15	64	40	57	56	54	42	35	40	93	55	82	08	78	87
31	49	87	12	27	41	07	91	72	64	63	42	06	66	82	71	28	36	45	31	99	01	03	35	76
69	17	22	23	46	10	75	83	62	94	44	65	46	23	65	71	69	20	89	12	16	56	61	70	41
93	57	21	56	98	42	52	53	14	86	24	70	25	18	23	23	56	24	03	86	11	06	46	10	23
77	56	18	37	01	32	20	18	70	79	20	85	77	89	28	17	77	15	52	47	15	30	35	12	75
37	07	47	79	60	75	24	15	31	63	25	93	27	66	19	53	52	49	98	45	12	12	06	00	32
72	08	71	01	73	46	39	60	37	58	22	25	20	84	30	02	03	62	68	58	38	04	06	89	94
55	22	48	46	72	50	14	24	47	67	84	37	32	84	82	64	97	13	69	86	20	09	80	46	75
69	24	98	90	70	29	34	25	33	23	12	69	90	50	38	93	84	82	28	96	03	65	70	90	12
01	86	77	18	21	91	66	11	84	65	48	75	26	94	51	40	51	53	36	39	77	69	06	25	07
51	40	94	06	80	61	34	28	46	28	11	48	48	94	60	65	06	63	71	06	19	35	05	32	58
58	78	02	85	80	29	67	27	44	07	67	23	20	28	22	62	97	59	62	13	41	72	70	71	07
33	75	88	51	00	33	56	15	84	34	28	50	16	65	12	81	56	43	54	14	63	37	74	97	59
53	69	87	45	62	09	95	93	16	59	35	22	91	78	04	97	98	80	20	04	38	93	13	92	30
72	13	12	95	32	87	99	32	83	65	40	17	92	57	22	68	98	79	16	23	53	56	56	07	47
22	21	13	16	10	52	57	71	40	49	95	25	55	36	95	57	25	75	05	38	05	62	57	77	
97	94	53	67	90	68	74	88	17	22	38	01	04	33	49	38	47	57	61	87	15	39	43	87	00
09	03	08	53	63	29	27	31	66	53	39	34	88	87	04	35	80	69	52	74	99	16	52	01	55
29	95	61	42	65	05	72	27	28	18	09	85	24	59	46	03	91	55	38	62	51	71	47	37	33
81	96	78	90	47	41	38	36	33	95	05	90	26	72	85	23	23	30	70	51	56	93	23	84	30
44	62	20	81	21	57	57	85	00	47	26	10	87	22	45	72	03	51	75	23	38	35	56	77	97
68	91	12	15	08	02	18	74	66	79	21	53	63	41	77	15	07	39	87	11	19	25	62	19	30
29	33	77	60	29	09	25	09	42	28	07	15	40	67	56	29	58	75	84	06	19	54	31	15	53
54	13	39	19	29	64	97	73	71	61	78	03	24	02	93	86	69	76	74	28	08	98	84	08	23
75	16	65	64	64	93	85	68	08	84	15	41	57	84	45	11	70	13	17	60	47	80	16	13	50
36	47	17	08	79	03	92	85	18	42	95	48	27	37	99	98	81	94	44	72	06	95	42	31	17
29	61	08	21	91	28	76	72	84	98	26	23	66	54	86	88	96	14	82	57	17	99	16	23	99

Reproduced with permission from "A Million Random  
Diget", Rand Corp., Copyright 1955, the Free Press.

7. Repeat procedure in (6) to obtain nine additional two-digit numbers for the total of the items to be selected. If there are any repeats draw another number, etc.

There are more elaborate procedures for finding a starting point which are applicable when there are several pages in the table. For example, the pages and the rows and columns on a page may be numbered, then a random number is selected to identify which page, row, and column is used as a starting point. In Table I.1 there are 25 two digit columns, 50 rows, and suppose there were ten such pages similar to the page given. A five digit number could be drawn to select the page, row and column as suggested below.



After using a set of random numbers, the stopping place can be noted, and the next set drawn with the next number in sequence. If the bottom of the table is reached, assuming the numbers are taken in vertical sequence, the next number can be taken from the top of the following column of numbers of the same number of digits. Theoretically, one can read the numbers horizontally if one wishes.

The example previously described involved the selection of a sample from 100 items and thus the numbering of the items can be put in one-to-one correspondence with the set of two digit numbers as follows:

01	02	...	...	99	00
1	2	...	...	99	100

If the lot or population from which the sample is drawn does not consist of 10, 100, 1000, etc., items, then the correspondence

can often be altered to simplify the drawing of the sample rather than throwing out all of the numbers above N the size of the population. Two examples are given below, one in which N divides evenly into 100, and one for which this is not the case.

EXAMPLE I.1 Draw a random sample of seven days from 25 days using the random number Table I.1.

A random position in the table is first selected, for example, the number 54 in the circle. Starting at that point the following eight numbers are recorded:

No.	Random No.	Remainder
1	54	4
2	71	21
3	71	21
4	23	23
5	17	17
6	53	3
7	2	2
8	64	14

These numbers exceed 25 in five cases. Thus one can either (1) continue to select numbers until all numbers fall between 1 and 25 or (2) divide each number which is larger than or equal to 25 by 25 and use the remainder as the random number. If the remainder is 0, the random number is taken to be 25. In this example one additional number had to be drawn to avoid repeats. The final sample consists of days numbered 2, 3, 4, 14, 17, 21, and 23.

In the example, the correspondence is established through the following:

Items in population:	1,	2,	3,	...	24,	25
Nos. in table	01,	02,	03		24,	25
	26,	27,	28		49,	50
	51,	52,	53		74,	75
	76,	77,	78		99,	00 (100)

Example I.2 Draw a sample of n = five days from n = 15 days using the random number Table I.1.

In the previous example, 100 is exactly divisible by 25 and thus each number 1 to 25 has an equal chance of being drawn under the system given, that is, dividing by 25 and taking the remainder. If this same system is used for  $N = 15$ ,  $n = 5$ , numbers 01 through 10 would have a greater chance of being selected than the numbers 11 to 15 because of the numbers 90 to 00 (100) being in the table. Thus disregarding these numbers the same approach can be employed as above. Using the same set of numbers for illustration the sample would be obtained as follows:

No.	Random No. (RN)	RN ÷ by 15 (Remainder)
1	54	9
2	71	11
3	71	11 (Repeat)
4	23	8
5	17	2
6	53	8 (Repeat)
7	02	2 (Repeat)
8	64	4

Thus the days numbered 2, 4, 8, 9, and 11 would be selected.

In this example, the correspondence is established as follows:

Items No. in Population:	1, 2, 3, ..., 14, 15
No. in Random Number Table:	01, 02, ..., 14, 15
	16, 17, ..., 29, 30
	31, 32, ..., 44, 45
	46, 47, ..., 59, 60
	61, 62, ..., 74, 75
	76, 77, ..., 89, 90
discarded	91, 92, 93, ..., 99, 00(100)

### I.2.2 Stratified Random Sample

Suppose the items are in stacks of 10 each. One could select one item at random from each stack of 10. The one item could be selected using either a random number table (one digit-0-9) or the deck of cards numbered 0, 1, 2, ..., 9. This is a

stratified random sample. That is, the lot of items to be sampled is first stratified or subdivided into sublots and a random sample is selected from each subplot in accordance to its size relative to the entire lot.

In many applications in sampling, the strata are defined to coincide with some characteristics of the population to be sampled. For example, the strata may be items produced or manufactured on one day, shift, or from one lot of material. In the case of sampling days from a year the strata might be weeks, months, or seasons or coincide with known production schedules of a particular manufacturing plant. In a stratified random sample the strata are sampled proportionally, thus providing in many applications a more representative sampling of the population. This is particularly true when there is considerable variation among strata and little variation within strata.

### I.2.3 Systematic Sample - (Pseudo-Random Sample)

Another procedure is to select a systematic sample of items by selecting the first item at random and every tenth item thereafter, depending on the sample size. In the case of selecting  $n = 10$  from  $N = 100$  items, the first item is selected at random from the first 10 items, and suppose it is a three (3), then items 13, 23, 33, etc.,...93 are selected. One needs to assess the possible consequences of the alternate sampling schemes, for if there is any possible relationship between the defective items and the systematic selection, it is obvious that some biased results could occur. For example, if the sample collection of a pollutant coincided with meteorological cycles or with weekly industrial activity patterns (i.e., say one of every seven days), the concentrations would tend to be less variable and may be relatively high (or low).

The reader is referred to standard statistical texts for a more complete discussion of sampling procedures.<sup>1,2,4,5</sup>

### I.3 ACCEPTANCE SAMPLING BY ATTRIBUTES

Now consider the problem of determining whether a lot of items should be accepted on the basis of given specifications. Suppose that a defective item is one with a physical defect which can be identified by a visual test or one for which a physical measurement falls outside a prescribed value(s) as in the example of the pH of the filters. Thus the sampling is by attributes, that is, an item is identified as either a defect or a good item (non-defect), often referred to as go/no-go inspection where this refers to whether the item meets the specification when checked by a gauge, calipers, sieve, etc.

A tabulation of sampling plans is given in MIL-STD-105.1 A complete discussion of the plans is given therein and not repeated here.

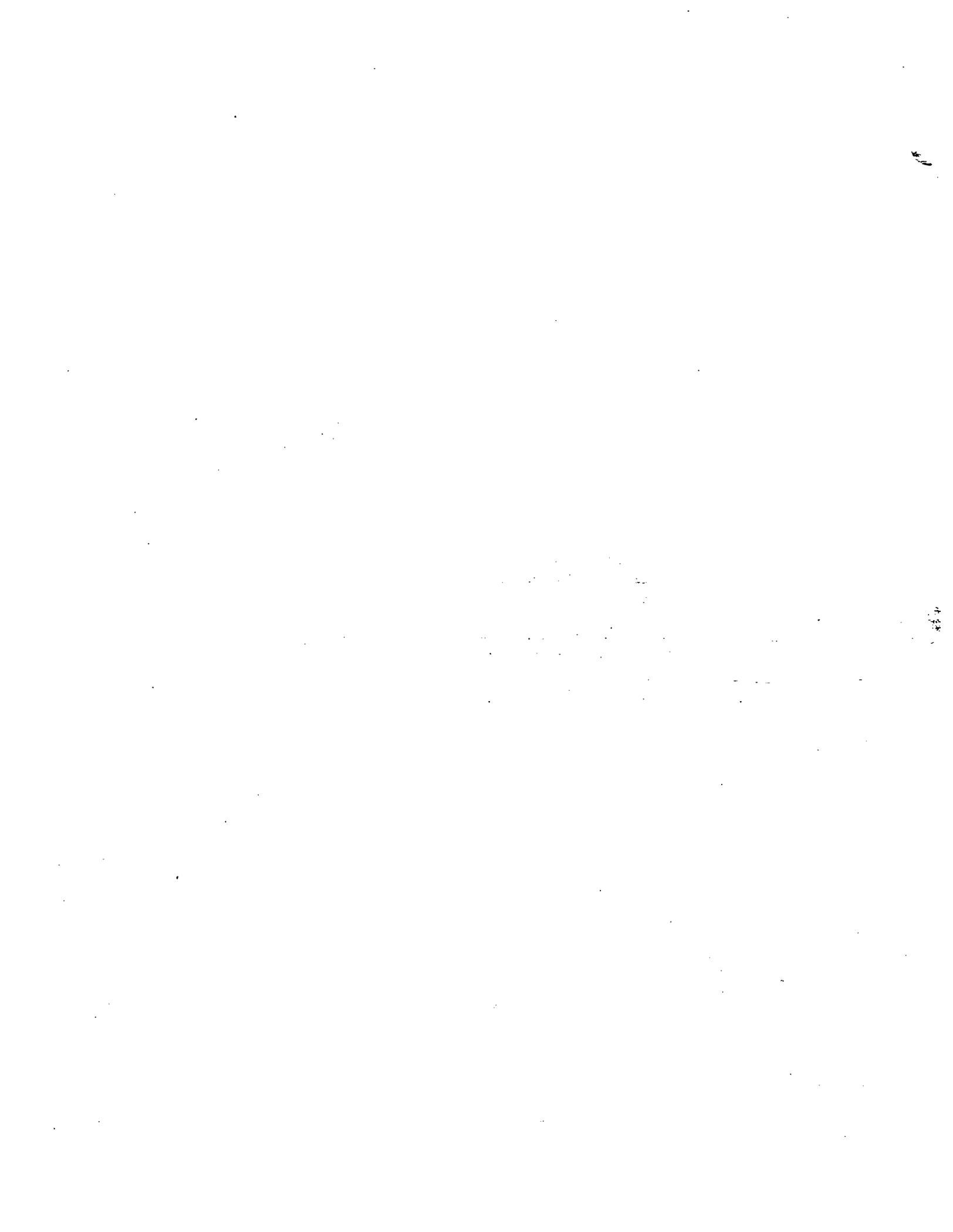
### I.4 ACCEPTANCE SAMPLING BY VARIABLES

A considerable savings in sampling may be achieved if the decisions concerning the acceptance of a lot of data (measurements) can be made on the basis of the actual measurement (a continuous value) rather than whether the measurements are outside specific limits. For example, in checking or auditing the quality of certain measurements the decision can be made to classify a routinely obtained field measurement as a defect if it deviates from the audit value by more than say 10% of the audited value. This results in sampling by attributes because one is classifying the items only as defects or nondefects (good measurements) rather than using the measured difference in the two values. If the latter is used, the decision on acceptance of a lot is made by variables, that is, on the basis of the mean and standard deviation of the sample of  $n$  measurements and given constants, very much like that used in control charts.

Variable sampling plans are described and tabulated in MIL-STD-414 (normal distribution only).<sup>4</sup> The reader is referred to this standard for further details and to Reference 3 for some specific variable sampling plans.

I.5 REFERENCES

1. MIL-STD-105, Sampling Procedures and Tables for Inspection by Attributes, Government Printing Office, Washington, D. C.
2. Juran, J. M., Quality Control Handbook, Third Edition, McGraw Hill Book Co., New York, 1974.
3. Owen, D. B., Variables Sampling Plans Based on the Normal Distribution, Technometrics, Vol. 9, No. 3, August 1967, pp. 417-424.
4. MIL-STD-414, Sampling Procedures and Tables for Inspection by Variables for Percent Defective, U.S. Department of Defense, Military Standard, Government Printing Office, Washington, D. C.
5. Bowker, A. H. and H. P. Goode, Sampling Inspection by Variables, McGraw Hill Book Co., New York, 1952.
6. Smith, F., Measuring Pollutants for Which Ambient Air Quality Standards have been Promulgated - Final Report, EPA-R4-73-028e.
7. Smith, F., Determination of Beryllium Emissions from Stationary Sources, EPA-650/4-74-005k.



APPENDIX J  
CALIBRATION

J.1 CONCEPTS

One of the most important steps in the measurement process is the calibration of the instruments for use in measuring the concentration of pollutants in ambient air. In addition to a multipoint calibration performed periodically, there need to be frequent checks, such as zero-span checks, to determine if there is a significant change in the calibration curve obtained at the most recent multipoint calibration. These checks should be performed at least biweekly, in accordance with Reference 1. The techniques for making some of the decisions concerning calibrations are not discussed in the general literature and some specific guidelines are presented in this section for use by the analyst/operator performing the calibrations. These guidelines emphasize that care should be taken to note trends in the results over time due to degradation of an instrument, calibration gas or other standards used in the frequent checks based on actual data. However, the operator should always consider his subjective feeling concerning a change in either the instrument/standard and/or environment which may alter the calibration. He should make the checks that he considers necessary to assure himself that valid results will be obtained with the measurement system.

The discussion of calibration procedures requires some consideration of statistical techniques for fitting a linear or nonlinear function to a set of calibration data. The technique most frequently used is the least-squares method. Although calibration data can usually be adequately fitted by "eyeball," that is, fitting a line or curve to the data based on a subjective fit by an analyst using a straight edge or a french curve,

this technique must be supplemented by some calculation of how well the data are "fitted" by the curve in order to predict the precision/accuracy of the reported data and to make objective checks on whether the calibration may have changed. The least-squares method and some associated statistical computations will be discussed briefly in Subsection J.2.

Some specific problems to be considered in this section are indicated below under the two headings, multipoint calibration and zero-span calibration.

#### Multipoint Calibration (MPC)

Some questions to consider with respect to MPC are:

1. How often to repeat the calibration?
2. How many points (levels of standards, e.g., concentrations of calibration gases) to use?
3. How to space the levels of standards?
4. Is the calibration curve linear or nonlinear?
5. Is the calibration equally precise for all concentration levels?
6. How to estimate the precision of the estimates read from the calibration curve or table?
7. How does the expected range of measured concentrations affect the selection of the levels to be used in the calibration?
8. How can a quality control chart be used to monitor changes in the MPC?

#### Zero-Span Calibration (OSC)

Some of the questions to be considered for the OSC are:

1. How often do we make an OSC?
2. At which levels of the standard should the checks be made?
3. What is the day-to-day variation among the checks?
4. How is a quality control chart used in determining if a significant drift or change in the calibration curve has occurred, based on the OSC?

5. How does the expected range of concentrations affect the selection of the OSC levels?

There are additional questions which may be asked relative to the calibration procedure, but the above represent some of the most important ones. The answers to all of the above questions are not easy to discuss briefly in this section. The approach will be to use examples to indicate the approach and refer to some appropriate statistical texts (References 2, 3) for a more detailed discussion of the basic mathematical structure of the problem.

### J.2 MULTIPOINT CALIBRATION (MPC)

Example J.1 provides multipoint calibration data for an NO<sub>2</sub> analyzer for five levels of concentration. These data are plotted in Figure J.1 to give Y as a function of X. A straight line was fitted to the five data points using the least-squares method. The necessary calculations are indicated below the example.

#### Example J.1

X (Concentration of NO <sub>2</sub> , ppm)	Y (Analyzer Reading, volts)
0.10	0.039
0.20	0.086
0.30	0.140
0.50	0.254
0.75	0.369
Totals 1.85	0.888

Assuming that the calculations are performed manually using a desk (or portable) calculator without a completely programmed feature, the following steps are performed in obtaining the equation of the "best fit" line by the least-squares method and the variance of the responses about the fitted line.

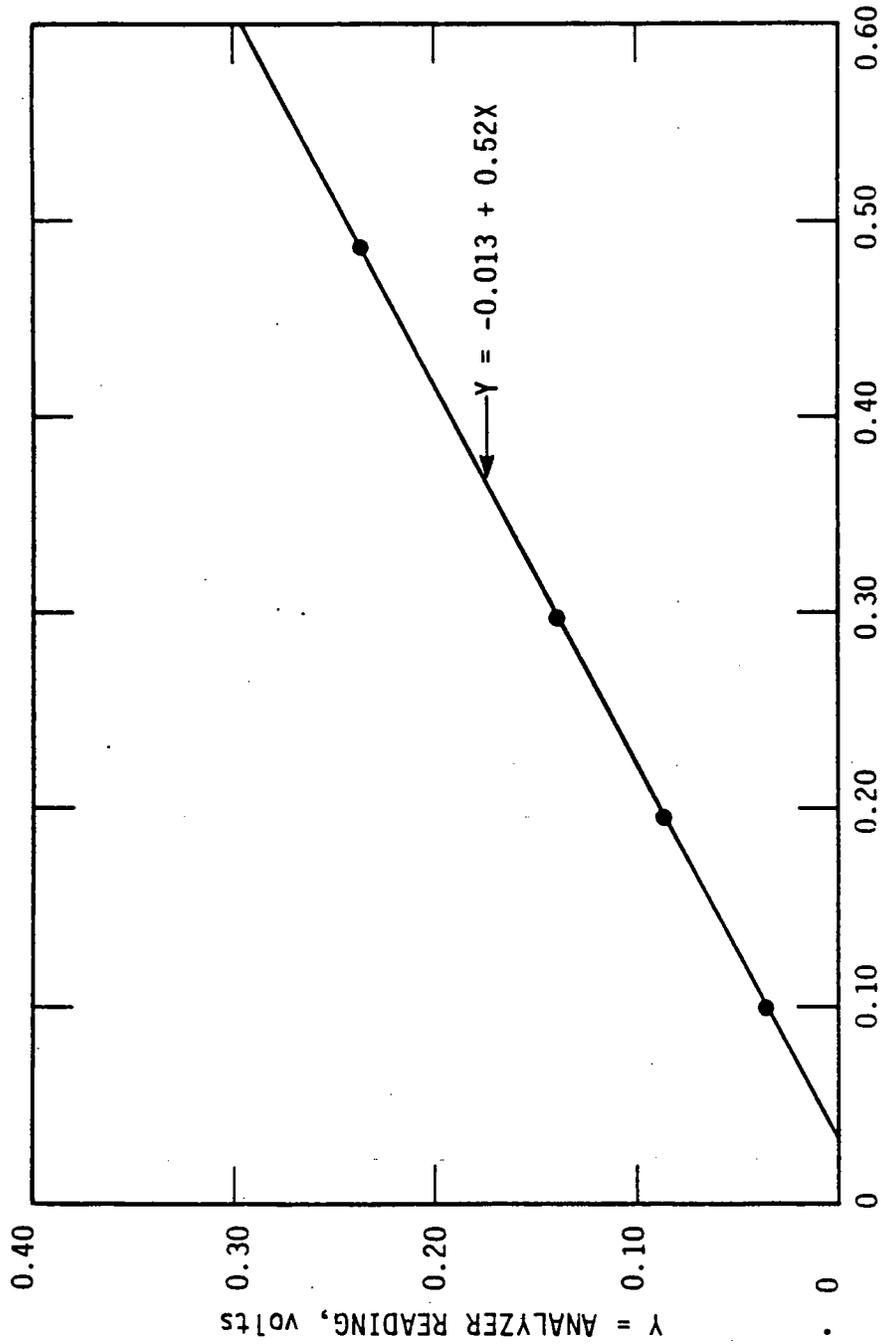


Figure J.1. Multipoint calibration of NO<sub>2</sub> analyzer.

Preliminary Computations

$$\begin{aligned}\Sigma X &= 1.85 & \Sigma Y &= 0.888 \\ \Sigma X^2 &= 0.9525 & \Sigma Y^2 &= 0.2292 \\ \Sigma XY &= 0.4668\end{aligned}$$

The notation  $\Sigma X$  is an abbreviation for  $\sum_{i=1}^n X_i$ , similarly for  $\Sigma Y$ ,  $\Sigma X^2$ ,  $\Sigma Y^2$ , and  $\Sigma XY$ . The subscripts are dropped for ease in typing and reading.

$$\bar{X} = \Sigma X/n = 1.85/5 = 0.37 \qquad \bar{Y} = \Sigma Y/n = 0.888/5 = 0.1776$$

$$b = \frac{\Sigma XY - (\Sigma X)(\Sigma Y)/n}{\Sigma X^2 - (\Sigma X)^2/n}$$

$$b = \frac{0.4668 - (1.85)(0.888)/5}{0.9525 - (1.85)^2/5} = \frac{0.1383}{0.2680} = 0.516.$$

Record the numerator and denominator of  $b$  for future use. The equation of the line fitted to the data is written as

$$\begin{aligned}Y_p &= \bar{Y} + b(X - \bar{X}) = a + bX \\ &= 0.1776 + 0.516(X - 0.37) \\ &= -0.013 + 0.52X\end{aligned}$$

where  $Y_p$  is the predicted mean response for the corresponding  $X$ .

The next step in the calculation is to determine how well the data are fitted by the line. The measure used is the sum of squares of deviations between the data points  $Y$  and the corresponding prediction  $Y_p$  given by the equation above, divided by the sample size less 2 (i.e.,  $n-2$ ). The reason for dividing by  $n-2$  instead of  $n-1$ , as used in Appendix C, is that two parameters of the line have been estimated from the data (the intercept and the slope), thus the remaining degrees of freedom in fitting the five data points is  $n-2 = 3$ .

The calculation of the variance and standard deviation of the responses is performed as follows:

$$s_{Y|X}^2 = \frac{\sum Y^2 - (\sum Y)^2/n - b\{\sum XY - (\sum X)(\sum Y)/n\}}{n-2}$$

$$= \frac{0.2292 - (0.1776)(0.888) - (0.516)(0.1383)}{3}$$

$$s_{Y|X}^2 = 0.000127/3 = 0.000042$$

and  $s_{Y|X} = 0.0065$  volts.

The notation  $s_{Y|X}$  (read "the standard deviation of Y, given X) is not used when it is clearly understood from the context of the discussion that the standard deviation is that of the response or Y value about the regression curve (or line).

In the formula for  $s_{Y|X}^2$ , all of the information has been previously calculated,  $\sum Y^2$ ,  $\bar{Y}$ ,  $\sum Y$ ,  $b$ , and the quantity in brackets {} is the numerator of the expression for  $b$ . The last term in the computation of the numerator of  $S_{Y|X}^2$  can also be written as

$$b \times b\{\sum X^2 - (\sum X)^2/n\} \text{ or } b^2\{\sum X^2 - (\sum X)^2/n\};$$

that is,  $b^2$  multiplied by the denominator used in calculating  $b$ . A calculation form (Figure J.2) is provided for manual computation.

Line Through the Origin - If the line must pass through the origin ( $X = 0, Y = 0$ ) then  $Y_p = b'X$  where

$$b' = \frac{\sum XY}{\sum X^2} = \frac{0.4668}{0.9525} = 0.49,$$

compared to 0.516 when not forced through the origin. The estimated variance  $s_{Y|X}^2$  for the case in which the line must pass through the origin is given by

$$s_{Y|X}^2 = \frac{\sum Y^2 - b' \sum XY}{n-1} = \frac{0.2292 - (0.49)(0.4668)}{4} = 0.000111,$$

$$s_{Y|X} = 0.0105 \text{ volts.}$$

Certain assumptions are implied by the least-squares analysis as described:

Calibration Form  
Linear Regression (Least Squares)  
(Straight line--not through the origin)

Sample number	<u>X</u>	<u>Y</u>	<u>Y<sub>p</sub> (predicted mean)</u>
1	_____	_____	_____
2	_____	_____	_____
3	_____	_____	_____
4	_____	_____	_____
5	_____	_____	_____
6	_____	_____	_____
7	_____	_____	_____
8	_____	_____	_____
9	_____	_____	_____
10	_____	_____	_____
	ΣX = _____	ΣY = _____	
	$\bar{X}$ = _____	$\bar{Y}$ = _____	
	ΣX <sup>2</sup> = _____	ΣY <sup>2</sup> = _____	ΣXY = _____
	(ΣX) <sup>2</sup> /n = _____	(ΣY) <sup>2</sup> /n = _____	(ΣX·ΣY)/n = _____
Subtract	Σ(X- $\bar{X}$ ) <sup>2</sup> = _____	Σ(Y- $\bar{Y}$ ) <sup>2</sup> = _____	Σ(X- $\bar{X}$ )(Y- $\bar{Y}$ ) = _____
Divide by	n-1 = _____	n-1 = _____	n-1 = _____
	s <sub>X</sub> <sup>2</sup> = _____	s <sub>Y</sub> <sup>2</sup> = _____	s <sub>XY</sub> = _____
Slope of fitted line:	b = s <sub>XY</sub> /s <sub>X</sub> <sup>2</sup> = _____		
Equation of fitted line:	Y <sub>p</sub> = $\bar{Y} + b(X-\bar{X}) = (\bar{Y}-b\bar{X}) + bX = a + bX = \underline{\hspace{2cm}} + \underline{\hspace{2cm}} X$		
Calculate predicted mean of Y for each X, record above. Variance of observed values (Y) about fitted values (Y <sub>p</sub> ) is calculated below.			
	ΣY <sup>2</sup> = _____	(ΣY) <sup>2</sup> /n = _____	bΣ(X- $\bar{X}$ )(Y- $\bar{Y}$ ) = _____
	s <sub>Y X</sub> <sup>2</sup> = [ΣY <sup>2</sup> - (ΣY) <sup>2</sup> /n - b Σ(X- $\bar{X}$ )(Y- $\bar{Y}$ )]/(n-2) = _____		
	s <sub>Y X} = √s<sup>2</sup> = _____</sub>	s <sub>b</sub> = √ $\frac{s^2}{\Sigma(X-\bar{X})^2}$ = _____	

Figure J.2. Calculation form for least squares fit of a straight line to calibration data.

1. The function is linear. This seems reasonable from the plot in Figure J.1. However, the appropriateness of the linear function should be checked by techniques given in the literature.<sup>2,3</sup>

2. The variation of the analyzer readings for a given level of concentration of NO<sub>2</sub> is the same for all levels. This is not usually a valid assumption. Although there are insufficient data to check this assumption in this example, the assumption can be checked in practice by taking several observations at each concentration such as one would do in repeated zero-span checks. There are techniques<sup>3</sup> for fitting a weighted linear function and these can be employed when it is determined that the variations depend on the concentration levels.

3. The levels of concentration of the standards are assumed to be precisely known. This is not exactly true as the standards are probably known to within 1 or 2 percent (absolute). The important point here is that the concentrations should be known with much greater precision than the precision of the response. In addition the error in these concentrations should be small relative to the variation of the concentration (e.g., a ratio less than 0.1). The effect of violations in the assumption on the above procedure is described in Reference 2 under the topic "Error in Both Variables."

In order to simplify the calculations and the discussion herein, it is assumed that the above assumptions apply and that no modified approach need be taken. These assumptions seem appropriate to illustrate the techniques to be described in this section and are adequate in most practical applications.

After the multipoint calibration curve has been determined, the curve is used to estimate the concentration of NO<sub>2</sub> from a given analyzer reading in volts for each sample until a recalibration is performed according to schedule or indicated need. The procedure to be used to make this decision will be discussed

in Subsection J.3. The estimated precision of a reported concentration in ppm from a given analyzer reading involves an inverse prediction using the regression line of Figure J.3. That is, given an ordinate value Y, what is the corresponding X?<sup>4</sup> Note that the fitted line was derived on the assumption that given X, what is the predicted mean value of Y? The analyzer readings vary from replication-to-replication and day-to-day and thus there is an inherent error in these readings. This in turn is combined with the variation in the predicted values of Y for given X as determined by the least-squares fitted line, to yield an estimate of the precision in the predicted concentration. This prediction is obtained by inverting the equation  $Y = -0.013 + 0.52X$  to obtain  $X_p$  the value of X predicted from the calibration curve,

$$X_p = \frac{Y + 0.013}{0.52}$$

or

$$X_p = 0.0258 + 1.938Y \text{ ppm,}$$

thus for  $Y = 0.10$ ,  $X_p = 0.22$  ppm. The question is, can an interval be given in the form that the true concentration falls between the limits  $(0.22 \pm w)$  ppm? The answer is of course yes, but such an interval is rather tedious to calculate. Using the appropriate procedure,<sup>3</sup> the interval for the concentration of  $\text{NO}_2$  for  $Y = 0.10$  volts, is about  $0.22 \pm 0.07$  ppm. Thus even though the calibration curve (line), appears to be reasonably precise, the resulting error in the predicted concentration can be quite large. This points out one of the important aspects in the calibration procedure, that is, relatively small errors in the deviations between the values observed and the calibrated values can yield relatively large errors in predicted concentrations and thus this becomes one important source of uncertainty in the measurement process. This result also emphasizes the need for a precise calibration technique.

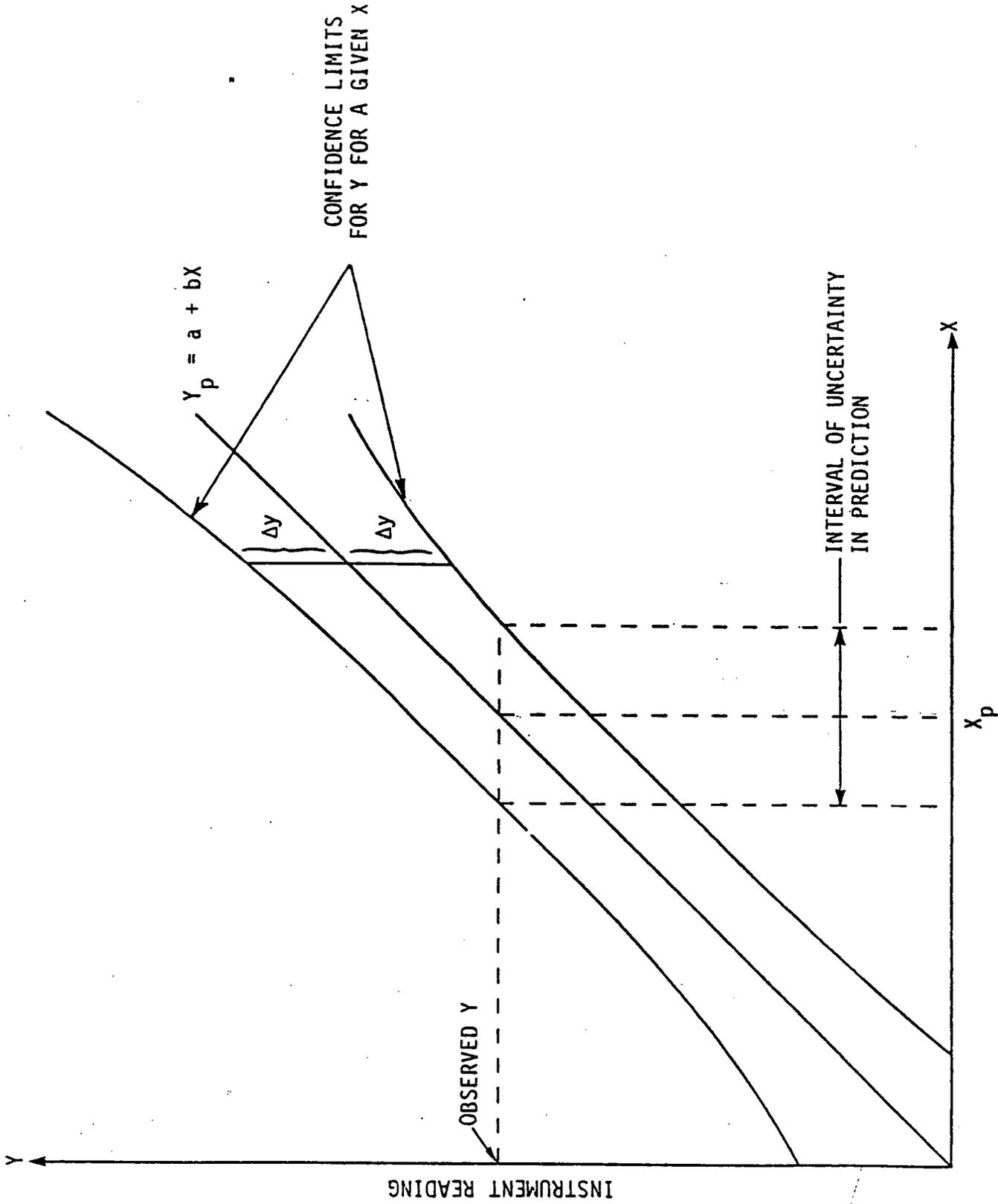


Figure J.3. Inverse prediction using regression line.

The prediction procedure described herein is known as the classical method<sup>4</sup> in the statistical literature on the subject of calibration. Reference 4 compares several prediction procedures and recommends a procedure different from the classical method. Although the application of this procedure will not alter practically the results for the example given herein (e.g., the prediction equation for  $X_p$  is

$$X_p = 0.0261 + 1.9365Y \text{ ppm}$$

using the proposed estimator),<sup>4</sup> it would be advisable to consider the approach in this paper for general usage.

The questions given in Subsection J.1, under MPC now become important in terms of how these predictive errors can be reduced to an acceptable order of magnitude, if this has not already been achieved. To reduce these errors, one or more of the following steps can be taken:

1. Improve calibration standards (i.e., be sure that their levels are known as precisely as possible)
2. Repeat the MPC entirely
3. Adjust the spacing of the levels of the standards to yield improved precision within this expected range of the Y's to be observed
4. Repeat some levels of the standards used in the MPC.

Consider here generally the effect of each of the above actions. First of all, action (1) will increase the cost of the acquisition and testing of the standards, and in maintaining and replacing of same. The tradeoff between these increased costs and the improvement which can be achieved would need to be considered by employing the techniques given under the subject "Error in Both Variables."<sup>2</sup> It is intuitively felt that if the stated concentrations of the standards are within 2 percent of their true values, then their overall impact on the precision of the predicted values will be sufficiently small.

Now consider what will happen if the MPC is repeated, action (2), (e.g., it may be redone on 5 consecutive days and

the results put together to obtain on MPC based on 25 measurements, 5 at each level obtained on each of 5 days.) If the day-to-day variations are negligible, this procedure reduces the interval to about one-half the width for one MPC, (i.e., the result could be reported as  $0.22 \pm 0.035$  ppm instead of  $0.22 \pm 0.07$  ppm as indicated for the previous example). In combining the results over 5 days, checks must be made of the consistency over the 5 days; if the analyzer behaved erratically on a specific day, this should be considered in the analysis, with the calibration data for that day possibly being discarded (See Appendix F). With additional repeated calibrations the interval width could be further reduced, but this approach is not suggested as a practical one for reasons to be given later.

The next consideration, action (3), is that of adjusting the spacing levels of the calibration gases. For example, if it is known (without any doubt) that the calibration curve is linear, then the best allocation of n levels of concentration would be as follows:

- n even: make 50 percent of the observations at each end of the range of concentrations over which the linear relationship holds.
- n odd: make 1 observation at the center as a confirmation of continued linearity and 50 percent of the remaining observations at each end of the range.

For additional information see a discussion of the optimal spacing of data for such problems.<sup>5</sup> If there is some doubt concerning the linearity assumption, the above approach would be poor and one should then take more center points or add two intermediate points to be able to check for nonlinearity. Some experience in the type of nonlinearity would be valuable in the selected spacing of the concentrations of the calibration gases. This approach to improving the MPC and the resulting precision of the predicted concentrations is appropriate as considerable

improvement in the precision can be obtained without repeating an entire MPC. This leads to the last and recommended procedure for improving almost all calibration procedures.

The fourth action suggests that certain levels of the MPC be repeated to improve on the precision of the results. It is beneficial to consider the combination of zero-span calibration or checks with the MPC in this approach. Hence before discussing this further, consider first the following discussion on OSC.

### J.3 ZERO-SPAN CALIBRATION CHECKS (OSC)

The zero-span calibration checks can be used to significantly improve on the MPC and to detect when a change may have occurred in the calibration (e.g., the instrument may have degraded because of a particular component wear-out, or the calibration gas may have degraded, changing concentration with time). These changes can be sudden (or catastrophic) or a slow degradation. The desire is to detect any changes that would significantly alter the calibration. The approach might be suggested that the daily OSC be used for that day and the equipment adjusted each day accordingly. However, this approach may result in predictive results with relatively poor precision. If the day-to-day variations are random variations in analysis which are consistent with the capabilities of the measurement process, then the most recent MPC relationship should be used; if not consistent, a new MPC should be obtained. Before suggesting a calibration procedure, consider the questions posed in Subsection J.1, under OSC.

1. How often do we make an OSC? This will be determined from the experience with day-to-day and within day variations in the OSC just as one needs to determine the causes of variability in a production process prior to determining the best sampling procedure to use in quality control checks. It would be suggested that initially they be made very frequently (e.g., at

least once a day) until sufficient experience is gained, such as 25 to 30 OSC's should be required to test for trends or shifts in the levels.

2. At which levels of the standards should the OSC be made? The answer to this question is similar to that point discussed under MPC, that is, if the calibration is truly linear, the precision can be increased by spreading out the levels of concentration. If however, the calibration is not linear the levels should be closer to the expected range of observed values of concentration for which the calibration curve will be used. Suggested levels are specified in Reference 1.

3. What is the day-to-day and within day variation among the OSC checks? These variations are determined by using the same concentration levels on each day (or time of calibration check) and measuring the variation in the observed analyzer readings. These variations may depend on the level of the standard. For example, in the MPC example of Subsection J.2, the levels might be taken at 0.05 ppm and at 0.85 ppm. Calculate the value of  $s^2$  for each set of measurements, (i.e., analyzer readings at 0.05 ppm and at 0.85 ppm).

4. How is a quality control chart used in determining if a significant drift or change in the calibration curve has occurred? The results obtained with each OSC can be plotted on a quality control chart for individuals (see Appendix H). An example chart will be described in the following subsection.

5. How does the expected range of concentration of the pollutant in the particular region (site being monitored) affect the selection of the OSC levels? The answer to this question would follow the recommendation under question (2) above.

#### J.4 SUGGESTED CALIBRATION PROCEDURE

It is not expected that a single proposed procedure could be appropriate for all calibrations used in ambient air pollutant measurement processes. However, it is felt that these procedures should follow the general set of guidelines set forth

in this section. These guidelines provide some flexibility in the approach depending on the situation. Basically, it is assumed that an MPC can serve as a basis for predicting the concentration of a pollutant in an unknown sample over a period of several days or weeks and that the OSC's can be used to check on this prediction (i.e., its precision) until a significant change is detected in the measured values (by an NO<sub>2</sub> analyzer, for example) in the zero-span gases. (The discussion herein can apply to calibration of a rotameter, pitot tube, etc.; however, the terminology for calibration of gas analyzers is used throughout this section to simplify the presentation and hopefully make it more readily understood). The significant change will be determined by an appropriate quality control chart or limits within which certain measurements of results should fall as previously suggested. If a point falls outside the prescribed control limits, the MPC should be checked and, if necessary, a new calibration gas may need to be acquired.

Example J.2

Consider the following zero and span drift data.

<u>Day</u>	<u>Slope</u>	<u>Zero drift, %</u>	<u>Span drift, %</u>
1	1.99	0.88	6.47
2	2.12	-0.18	5.23
3	2.23	0.06	-1.70
4	2.20	-0.34	-3.64
5	2.12	0.38	-2.36
6	2.06	0.12	0.73
7	2.08	-0.04	0.48
8	2.09	-0.26	-8.56
9	1.91	0.32	1.52
10	1.94	0.02	2.89
11	2.00	-0.30	-3.46
Avg ( $\bar{X}$ )		0.06	-0.22
Std dev(s)		0.36	4.33
$\bar{X} + 3s$		1.08	13.00
$\bar{X} - 3s$		-1.08	13.00

The zero and span drifts are plotted on control charts (Figures J.4 and J.5). These limits are centered at 0, the expected values of zero and span drift.

Because of the possible tendency of the zero and span values to change together or be correlated, a quality control chart may be maintained on the slope of the linear calibration (or the difference in the zero and span instrument readings, which is a constant multiple of the slope) and one point, say the span value. This quality control chart for the slope (or difference in zero and span values) may be more sensitive to changes in the instrument depending on the degree of correlation in the two readings.

#### J.5 OTHER REGRESSION PROBLEMS

In the previous sections it was assumed that  $Y$  is predicted by a simple linear relationship with  $X$ , that is,  $Y_p = a + bX$ . This assumption is adequate for a large number of regression (least squares) problems. However, there are also many problems for which either (1) a transformation can be performed on the  $X$ 's and/or the  $Y$ 's in order to obtain a linear relationship or (2) the prediction relationship cannot be so transformed and nonlinear least squares techniques must be applied if it is desired to obtain a fitted curve using statistical techniques. Of course, it is always possible to draw a smooth curve through the data points, particularly when they appear to follow such a curve. There are, however, advantages to fitting a curve by the method of least squares and obtaining the associated information on goodness of fit. This is particularly true when the theoretical form of the curve is known and the estimated fitted parameters of the prediction relationship can be used to interpret the results. Nonlinear relationships for  $Y$  in terms of  $X$  can often be approximated by using polynomial functions. However, the use of such functions should be limited to the range of the data (interpolation) because the approximation usually is very poor outside the range of observation (extrapolation). Some examples are given to illustrate these points.



UCL: 13

LCL: -13

$\bar{X}$  AND R CHART

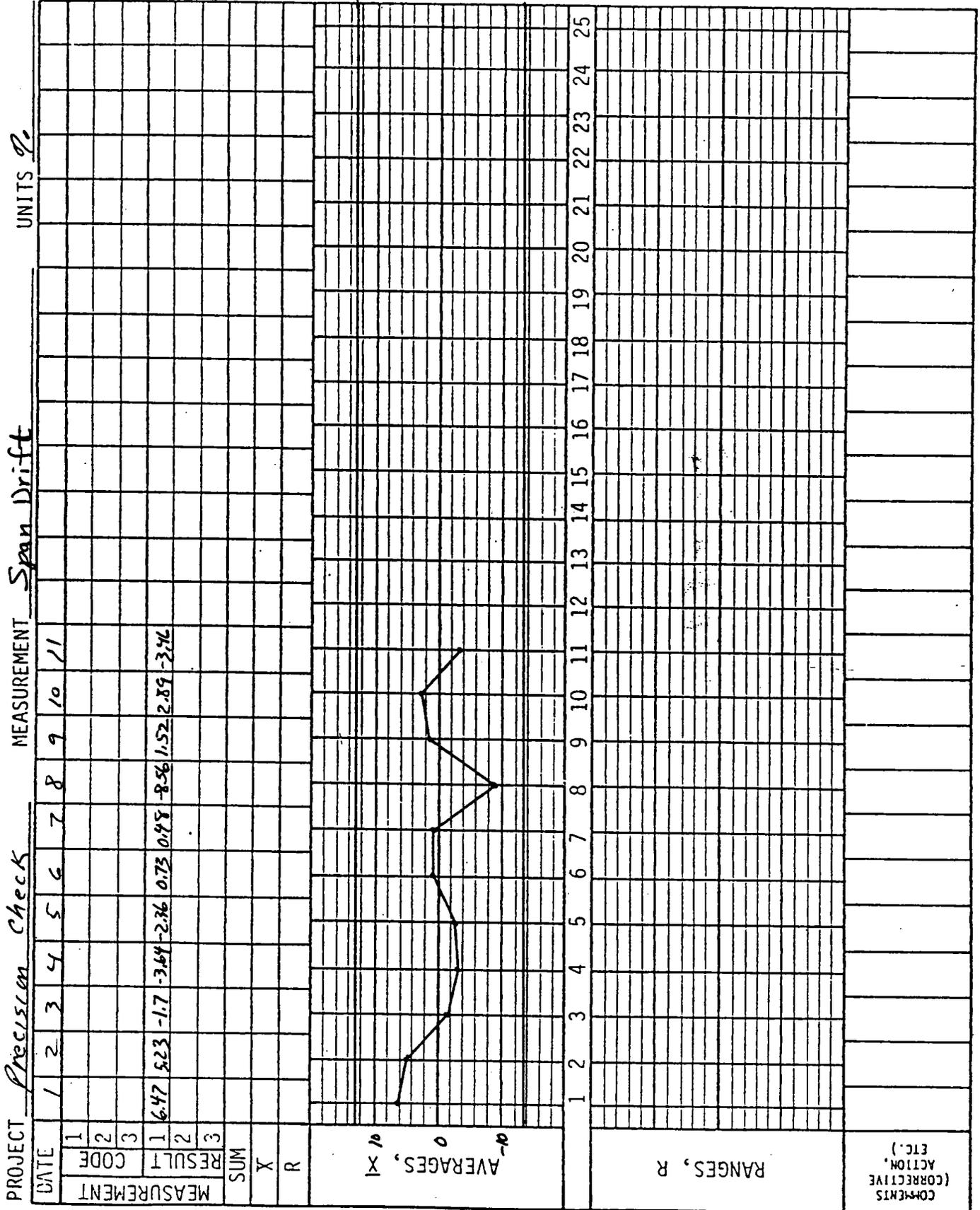


Figure J.5. Control chart for span drift, Example J.2.

Example J.3

A plot of transmittance (T) versus the concentration of the SO<sub>2</sub> standards in µg/ml yields a nonlinear relationship for the following data.

<u>SO<sub>2</sub> concentration, µg/ml</u>	<u>Transmittance (T)</u>	<u>Absorbance (A)</u>
0	0.863	0.064
0.032	0.815	0.089
0.081	0.752	0.124
0.162	0.650	0.187
0.326	0.484	0.315
0.663	0.279	0.555
0.952	0.165	0.782

However, transform the transmittance values to absorbance values using the relationship,

$$A = \log (1/T) = -\log T.$$

See Figure J.6 for a plot of these data before and after transformation. A straight line is then fitted to the transformed data, A versus the SO<sub>2</sub> concentration C, µg/ml, by the method of least squares.

The prediction equation is

$$A_p = 0.065 + 0.750 C$$

where

C = concentration of SO<sub>2</sub>, µg/ml,

A<sub>p</sub> = predicted mean absorbance level for given concentration C.

The standard deviation of the observed values of A about the fitted line is  $s_{A C} = 0.0044$  and the  $r^2$  (square of the correlation coefficient) is 0.9998.

Example J.4

Consider the following data for a calibration curve of absorbance versus concentration of SO<sub>4</sub><sup>=</sup> in µg/ml.

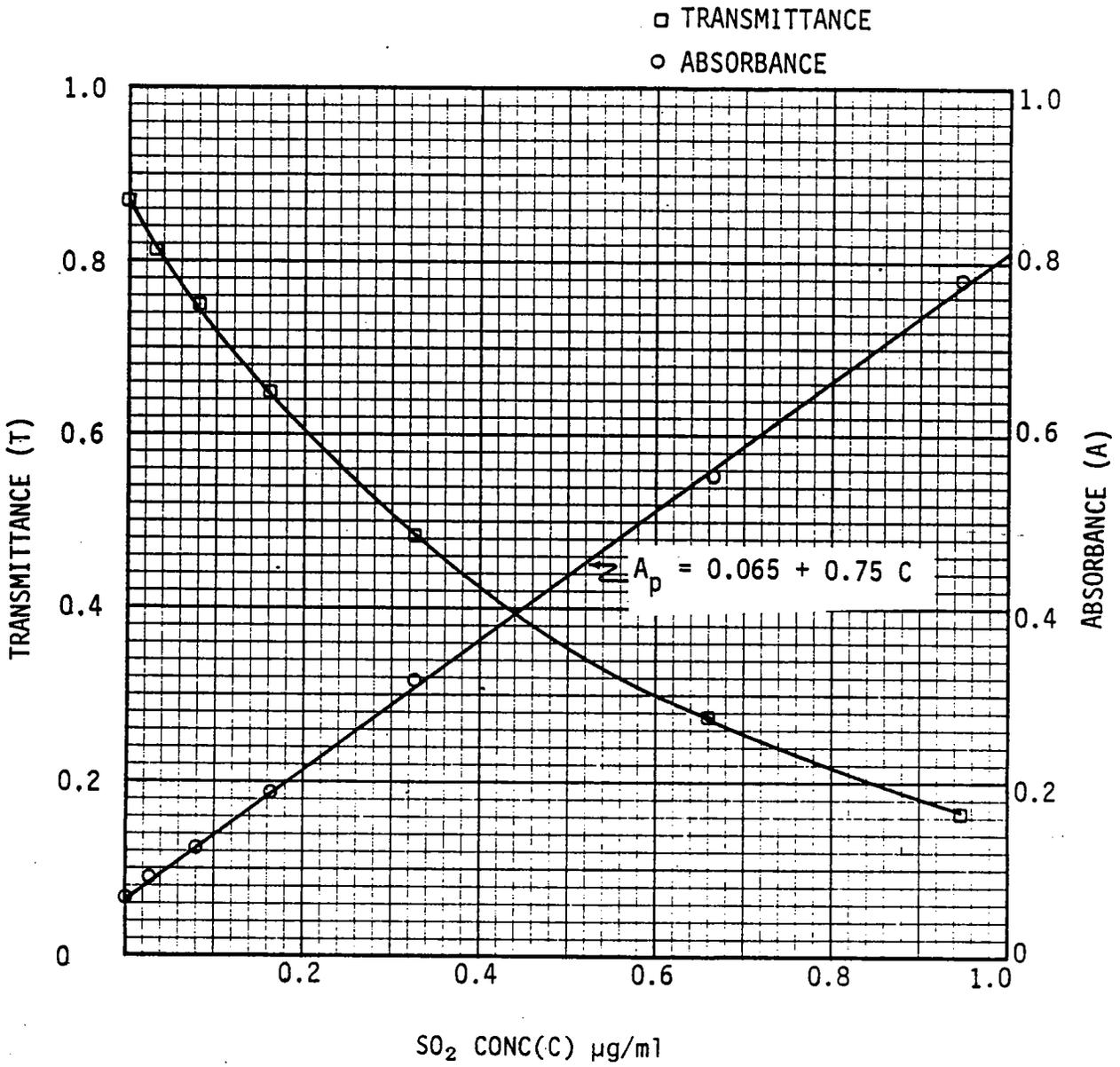


Figure J.6. Plot of transmittance and absorbance data versus SO<sub>2</sub> concentration.

$\text{SO}_4^-$ , $\mu\text{g/ml}$	Absorbance <u>(A)</u>
0	0
20	0.220
30	0.365
40	0.435
60	0.490
80	0.565

A plot of these data is in Figure J.7 and the data are fitted by a quadratic relationship

$$A_p = a + b_1C + b_2C^2,$$

where

$A_p$  is the predicted mean absorbance for given C,  
 $a, b_1, b_2$  are constants to be estimated by the method of least squares,

C is the concentration of the  $\text{SO}_4^-$  standard,  $\mu\text{g/ml}$ .

The calculation details are not given in this Appendix, however, any statistical text on least squares and/or regression will give the procedure.<sup>6</sup>

The prediction equation for absorbance is

$$A_p = -0.0010 + 0.0140 C - 0.000088 C^2.$$

Note that this relationship will not be appropriate for large C because the curve will turn downward and eventually attain negative values.

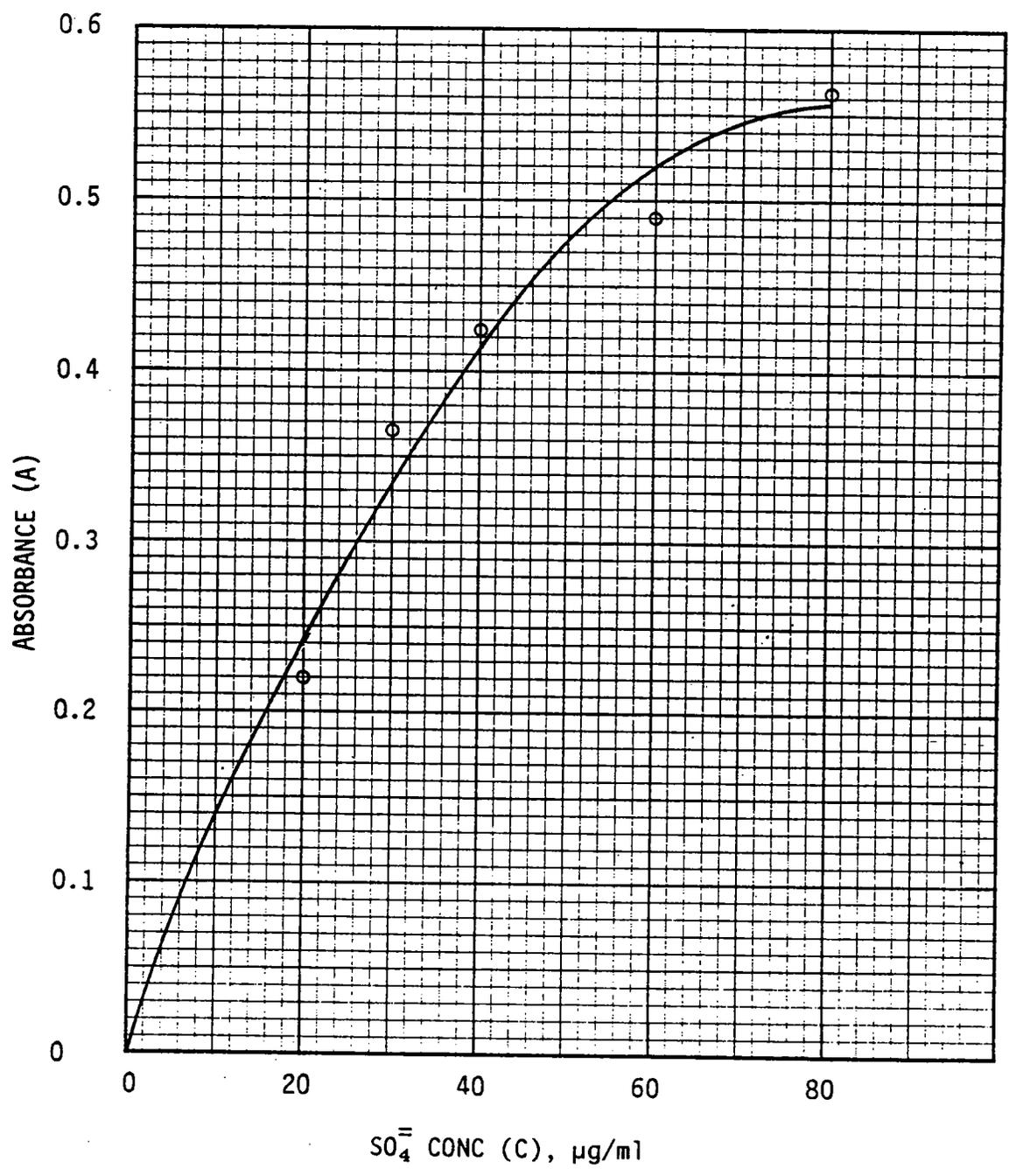


Figure J.7. Plot of absorbance versus SO<sub>4</sub><sup>2-</sup> concentration.

J.6 REFERENCES

1. Federal Register, Rules and Regulations, Vol. 44, No. 92, Thursday, May 10, 1979.
2. Mandel, John, The Statistical Analysis of Experimental Data, Interscience Publishers, John Wiley & Sons, Inc., New York, 1964.
3. Acton, F., Analysis of Straight Line Data, John Wiley & Sons, Inc., New York.
4. Tucker, W. T. The Calibration Problem Revisited. General Electric Company, Schenectady, NY. (Paper Presented at ASQC Chemical Division and Environmental Technical Committee, Conference in Cincinnati, Ohio, October 1980).
5. Ott, R. L., and Myers, R. H., Optimal Experimental Designs for Estimating the Independent Variable in Regression, Technometrics, Vol. 10, No. 4, November, 1968.
6. Hald, A., Statistical Theory with Engineering Applications. John Wiley and Sons, New York, 1952.



## APPENDIX K

### INTERLABORATORY TESTS

#### K.1 CONCEPTS

One of the most commonly employed procedures to assess the quality of data being reported by a large number of agencies/laboratories involved in a monitoring program is to conduct an interlaboratory test. Such a test typically involves preparing a set of samples, as nearly identical as possible, submitting them to a relatively large number of laboratories which are conducting the particular analysis, requesting that they be analyzed in the manner that a routine sample would be treated, and that the results be reported to the laboratory coordinating or conducting the test program.

One such interlaboratory procedure is the EPA performance audit program for analysis of simulated ambient air pollutants and source emission pollutants as described in two recent EPA reports.<sup>1,2</sup> The reported results are analyzed to determine the variation among the participating laboratories.

Another type of interlaboratory test program would be a collaborative test in which about ten to thirty selected labs might participate. The samples would be prepared in a similar manner to that in any interlab study, but in this case the analysis might be done on two or three days, duplicate or triplicate samples for each day. In this test, the data can be analyzed to determine not only the variation among laboratories, but also that within laboratories--within days and among days.<sup>3</sup>

#### K.2 INTERLABORATORY TEST

The test program for simulated CO samples is characteristic of almost any interlaboratory study; that is, samples were prepared at several levels of concentration (3 in this case) and

sent to each laboratory for analysis. Thus the results might appear as follows where L laboratories participate in the test program. Ideally each lab should report a value for each concentration level.

<u>Laboratory</u>	<u>Concentration level</u>		
	<u>1</u>	<u>2</u>	<u>3</u>
1	4.1	14.5	37.7
2	6.3	21.5	44.5
3	3.8	12.9	40.2
.	.	.	.
.	.	.	.
.	.	.	.
L	4.2	16.1	38.5

In parallel with the test data provided by the individual labs, the laboratory coordinating the study should analyze (or have analyzed) each simulated sample with a sufficiently large number of analyses to estimate the mean and standard deviation of the concentration of the simulated samples with the desired precision. This provides an independent measure of the target value (overall mean) and the variation in the simulated samples provided to the laboratories. These data can thus be used to determine limits within which the measurements obtained by the individual labs should fall with a given level of confidence.

K.2.1 Analysis of Interlaboratory Test Data - A brief discussion of the data analysis is given here in order that the participant/analyst in such programs can interpret the published results and the implication for his/her lab. There are several types of analyses given in the summary of results of such studies and this discussion will treat some of the more important aspects of these analyses, and refer the interested reader to some particular methods as described in the literature.<sup>1,2</sup> Clearly the sophistication of the analysis can vary considerably but the final results should be summarized so that they are useful to the laboratory analyst who wishes to determine if his analysis procedure is satisfactory and also to put some limits

on his reported data. Thus the analysis should have three purposes:

1. To summarize the overall test results.
2. To provide the necessary information to each participant for use in describing the quality of his data. This information need be given only to each lab as it should not be important for each lab to know where another particular lab stands, only how all labs are doing as given by (1) above, in order that one lab may compare its results against all lab results.
3. To provide each participant with a means of improving his/her analytical methods--for example, preference for a certain method of analysis due to more accurate and precise results, or identification of critical steps in the analysis by means of better written procedures. The analysis of the data only identifies the need for improving the method, not usually the means by which this can be done.

The ultimate objective in interlab studies is to improve data quality as reported by many laboratories so that comparative studies can be made with the assurance that the data quality is sufficient to yield valid results and decisions.

K.2.1.1 A summary analysis - As previously mentioned, one method of analysis of interlaboratory test data is described in an EPA report.<sup>2</sup> This provides an example of how the data from a large number of laboratories participating in an interlab test can be summarized. A summary analysis involves consideration of the following:

1. Identification and elimination of outliers. There are two aspects to this analysis: (a) determine if a laboratory is an outlier, that is, all of the observations from a laboratory indicate a significant bias and/or lack of precision in the measurement technique employed by a lab, and (b) determine if a single measurement by a lab is an outlier. Various test procedures are applied in the literature for identifying individual

outliers (e.g., see the references at the end of Appendix F). Similarly, there are alternative procedures for checking the consistency of one lab versus all other labs. Some of these procedures are described in the referenced report.<sup>1</sup> Some other procedures are given in appropriate statistical texts.<sup>4,5</sup>

2. Estimation of within and among laboratory variations. After the elimination of outliers (laboratories or individual measurements) by some appropriate method, it is then desired to summarize the remaining data by means of a measure of variation of the results within a laboratory and the variation among laboratories. The latter variation can be considered a measure of the variation of the inaccuracies or biases among the participating labs. Because the variation is typically dependent on the concentration level, it is usually desired to conduct the analysis in a manner to make the appropriate estimates of this dependence. For example, the analysis can be separately performed for each concentration level, or by making a transformation of the data.<sup>6</sup>

K.2.1.2 Analysis for a particular lab - A second purpose for the analysis is to provide the results in a manner as to be useful to a participant in the test program, that is, so that the lab may evaluate if its analysis procedure yields results consistent with those of all laboratories (excluding outliers) and assess the quality of its reported data. Suppose that the results of two particular labs are as indicated in Table K.1 and that the results of all labs (301 measurements) are summarized by their mean  $\bar{X}$  and standard deviation  $s$ .

If  $n$  is reasonably large, then  $\frac{X - \bar{X}^*}{s}$  is approximately normally distributed with mean zero and standard deviation one. Hence this value can be plotted on a quality control chart with

---

\*Actually,  $(X - \bar{X})/s\sqrt{1 - \frac{1}{n}}$  has a  $t$  distribution where  $X$  is one of the measurements of the sample from which the sample mean  $\bar{X}$  is determined.

limits at  $UCL = 3$ ,  $LCL = -3$ ,  $\bar{X} = 0$ . The value for each concentration, three in all, would be plotted at each sample number, or one can plot their range on a range chart and their average on an  $\bar{X}$  chart with limits given by:

Range chart:  $\bar{R} = 2.3$ ,  $UCL_R = D_4\bar{R} = 4.91$ ,  $LCL_R = 0$

Average chart:  $\bar{X} = 0$ ,  $UCL_{\bar{X}} = \frac{3}{\sqrt{3}} = 1.73$ ,

$LCL_{\bar{X}} = -\frac{3}{\sqrt{3}} = -1.73$

TABLE K.1. COMPARISON OF INDIVIDUAL LAB RESULTS WITH ALL LAB RESULTS

Data	Concentration level		
	1	2	3
<u>Lab X</u>			
	4.1	14.5	37.7
$(X-\bar{X})/s$	0.6	-0.2	0.4
<u>Lab Y</u>			
	6.3	21.5	44.5
$(Y-\bar{X})/s$	3.4	3.5	4.2
<u>All Labs</u>			
$n^a$	301	301	301
True value	3.8	14.6	36.4
$\bar{X}$	3.6	14.8	36.9
Median	3.6	14.7	36.9
Range	8.5	32.1	25.5
$s$	0.8	1.9	1.8
RSD	22.1	12.7	4.8
Accuracy <sup>b</sup>	-6.0	1.0	1.2

<sup>a</sup>  $n$  is the total number of participant measurements, no outliers removed.

<sup>b</sup> Accuracy =  $\frac{\text{Median} - \text{True Value}}{\text{True value}} \times 100$

The plot of the individuals can reveal some useful information; particularly, if the values are identified by concentration level. For example, a relationship between  $\frac{X-\bar{X}}{s}$  and concentration may indicate a poor calibration (error in slope).

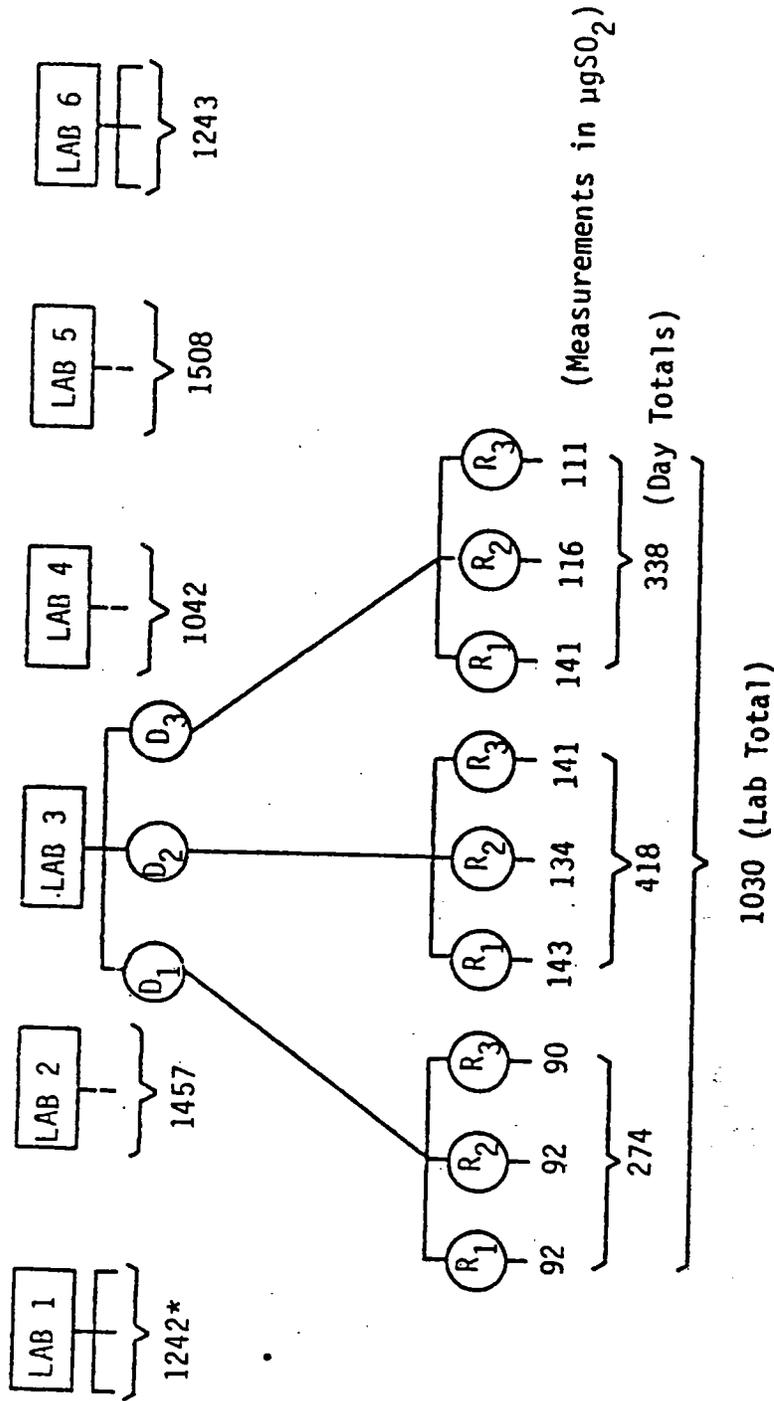
For the hypothetical laboratories given in Table K.1, one would infer that Lab X has obtained values consistent with the overall lab average, whereas Lab Y is biased high for all concentration levels with respect to the mean for all labs.

K.2.2 Feedback of information on methods - In order to attain the goal of overall improvement in data quality, there needs to be a feedback of information to the laboratories having some difficulties with the measurement process. This can be done, for example, through better method description, ruggedness tests to determine the critical steps in the method (this really should have been done prior to extensive use of a method), and/or consultation with experienced and well-qualified analysts.

### K.3 COLLABORATIVE TEST

The analysis of collaborative test data involves a more detailed breakdown of the total variation in measurements, that is, not only the variation among labs (reproducibility), but an estimate of the variation among days within labs (repeatability) and among replicates within days (replicability). A partial selection of data from a collaborative test study as given in Figure K.1 indicates the pattern of variation which is reasonably typical of such test data. For such data, the smallest component of variation is expected for within day or a measure of replicability, the next largest for among days or repeatability, the largest being among labs or reproducibility. These three components of variation are denoted by  $\sigma_r^2$ ,  $\sigma_d^2$ , and  $\sigma_l^2$  which are the variances for the replicates, days, and labs, respectively. These three components are then combined to obtain the three measures of variation.

<u>Measure</u>	<u>Variance</u>	<u>Standard Deviation</u>
Replicability	$\sigma_r^2$	$\sigma_r$
Repeatability	$\sigma_r^2 + \sigma_d^2$	$\{\sigma_r^2 + \sigma_d^2\}^{1/2}$
Reproducibility	$\sigma_r^2 + \sigma_d^2 + \sigma_l^2$	$\{\sigma_r^2 + \sigma_d^2 + \sigma_l^2\}^{1/2}$



\* Only the lab total is given for labs other than Lab 3.

Figure K.1. Data selected from a collaborative test study<sup>2</sup> ( $\mu\text{gSO}_2$ ).

Thus a measure of replicability is the variance or standard deviation among measurements made in the same day by the same analyst in the same laboratory. Repeatability is the sum of the replicate (within day) and the day-to-day variance for the same laboratory. Finally, reproducibility is the measure of total variation across days and laboratories. Typically the square roots of these variances (i.e., the standard deviations) are used as the corresponding measures. In some references, the measures are defined as the upper confidence limit for the difference between two repeat measurements which would be exceeded at most 10 percent (1%) of the time.

The specific method of analysis is described in most statistical texts on design and analysis of experiments under the subjects of nested experiments and components of variance.

### K.3 REFERENCES

1. Streib, E. W. and M. R. Midgett, A Summary of the 1982 EPA National Performance Audit Program on Source Measurements. EPA-600/4-83-049, December 1983.
2. Bennett, B. I., R. L. Lampe, L. F. Porter, A. P. Hines, and J. C. Puzak, Ambient Air Audits of Analytical Proficiency 1981, EPA-600/4-83-009, April 1983.
3. McKee, H. C., Childers, R. E., Saenz, Jr., O. S. Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method), Contract CPA 70-40, SWRI Project 21-2811, June 1971.
4. Mandel, J. The Statistical Analysis of Data, Interscience Publishers, a division of John Wiley and Sons, Inc., New York, 1964.
5. Youden, W. J. and E. H. Steiner. Statistical Manual of the Association of Official Analytical Chemists, The Association of Official Analytical Chemists, Box 540, Benjamin Franklin Station, Washington, DC 20044. 1975.
6. Interlaboratory Testing Techniques, Chemical Division Technical Supplement, American Society of Quality Control.

## APPENDIX L

### RELIABILITY AND MAINTAINABILITY

#### L.1 INTRODUCTION

Reliability and maintainability are those aspects of quality assurance which are concerned with the quality of performance of a component, instrument, or a process over time. One definition of reliability is--the probability that an item performs a specified function without failure under given conditions for a specified period of time or cycles of usage.

Maintainability is defined as--the probability that given a component (instrument or process) has failed, the component will be repaired or replaced and thus become operational within a specified time. An easily maintained system would be one for which troubleshooting for failures and the necessary repair or replacement of components can be performed in a short time. Diagnostic troubleshooting procedures provided by the manufacturer need to be included in the operator's manual.

Maintenance for automatic monitoring equipment can be divided into two types: routine (preventive) and nonroutine (corrective). Routine maintenance procedures are necessary to provide optimum operational performance and minimum instrument downtime. Nonroutine maintenance (corrective) is performed to rectify instrument failures or degraded performance of the instrument or to repair any part of the total system. Preventive maintenance should decrease the need for corrective maintenance.

An overall measure of system performance over time is availability. This can be simply defined as the ratio of the uptime of the system to the sum of uptime and downtime. Thus A is a measure of the ability of the system to perform its intended function when called upon to do so,

$$A = \frac{U}{U+D}$$

where            A = availability of the system,  
                  U = uptime,  
                  D = downtime.

Note that U+D is not necessarily the total calendar time, this would only be so if the system were used 24 hours per day every day. If the system is being used every third day, preventive maintenance can be performed on the off-days, not decreasing the system availability. Thus the availability is increased by having a reliable system (few failures) or a highly maintainable system (ease of maintenance) or a combination of these. In the pollutant monitoring systems it is desired that a system provide consistent and dependable data over long periods of time or continuously, thus a high availability is necessary. See Figure L.1 for a flow diagram of reliability and maintainability actions.

The causes of unreliability of an instrument are many. The complexity of measuring instruments consisting of many components may result in failure if any one component fails. Hence the likelihood of failure increases rapidly with the increasing number of components which need to perform a specified function in order that the system performs its required function. Records need to be kept on the system and on each component so that the time-to-failure (or degraded performance), cause of failure, and maintenance time can be determined for future use in procurement and maintenance practices. Clearly unreliable equipment or poorly maintainable equipment are to be eliminated from consideration in future procurement.

Field operational reliability depends upon three important variables in addition to the reliability of the components. Thus the reliability of the entire monitoring system is determined by the proper consideration of these variables:

1. Environment or operating conditions.

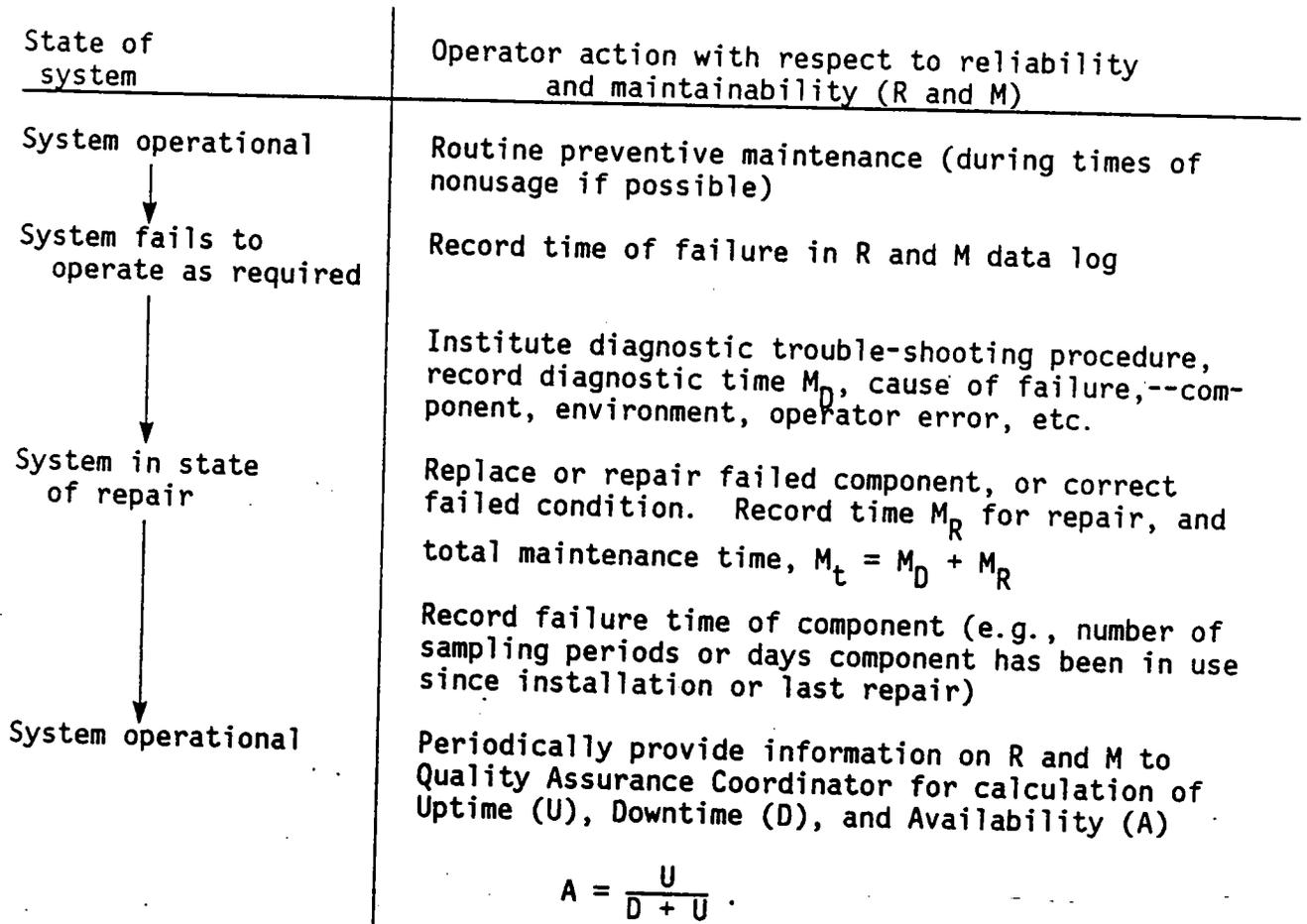


Figure L.1. Flow diagram of reliability and maintainability actions.

2. Functional interactions between various subsystems.
3. Conditions imposed by the operators.

An analytical calculation of reliability requires a great deal of knowledge concerning these factors. However, an empirical estimate can be obtained if the proper failure data are recorded. It is important to realize that some systems may be unreliable in some environments, (e.g., high humidity/high ambient temperature) whereas other systems may be unreliable if the operator is not properly trained.

Figure L.2 illustrates the typical life history of some types of equipments. The first phase is a debugging phase, the presence of marginal or short-life parts at original operation

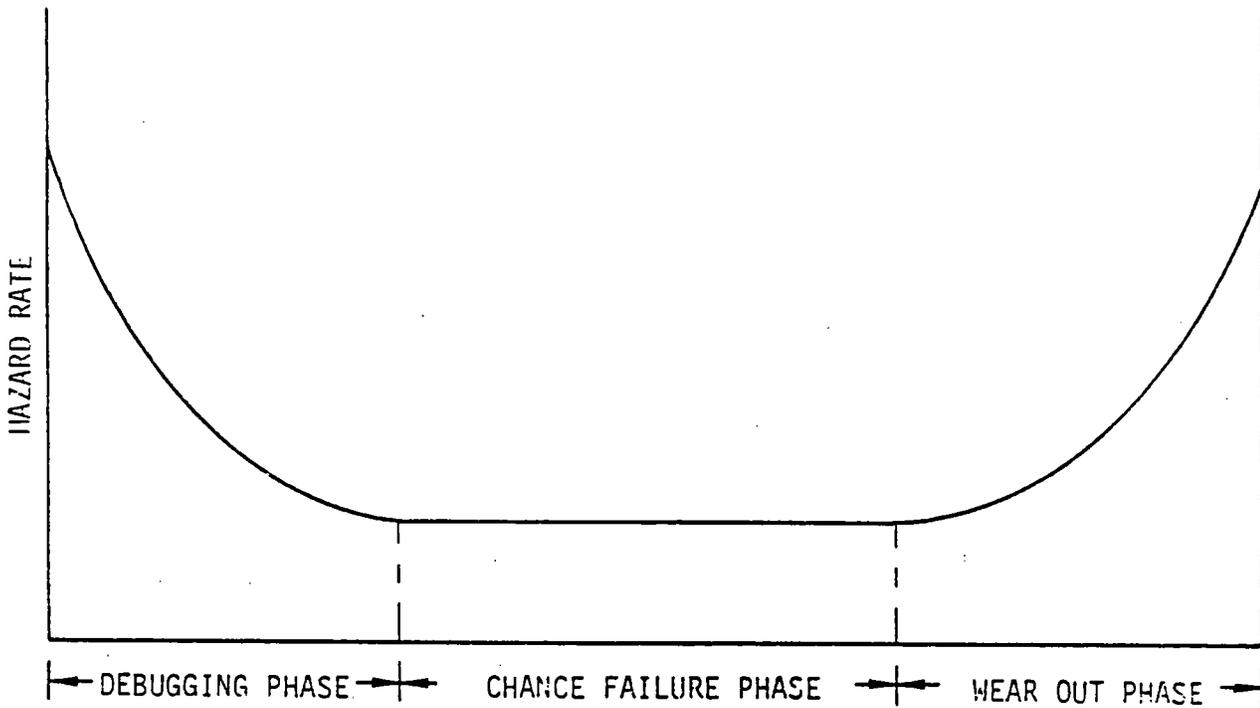


Figure L.2. Life history of some types of equipment (the bathtub curve).

(burn-in) is characterized by a decreasing hazard rate (instantaneous failure rate). The higher hazard rate period is commonly called the infant mortality period, and it results primarily from poor or marginal workmanship. The next phase is characterized by a relatively constant hazard rate which is the effective life of the equipment. This is followed by a period of increasing hazard rate, the beginning of wear-out failures in the equipment. This curve demonstrates the need for preventive maintenance in equipment exhibiting a history of behavior as illustrated in Figure L.3. If the debugging phase is short and if the hazard rate is relatively high, the equipment should be (1) tested through the burn-in period, or (2) the short-lived components should be identified and the appropriate equipment design modifications implemented. There are several techniques which can be employed to improve the reliability/availability of equipment. Some of these are listed here with no discussion. Reference to appropriate books on reliability would be necessary to select and apply the best technique.<sup>1,2,3</sup>

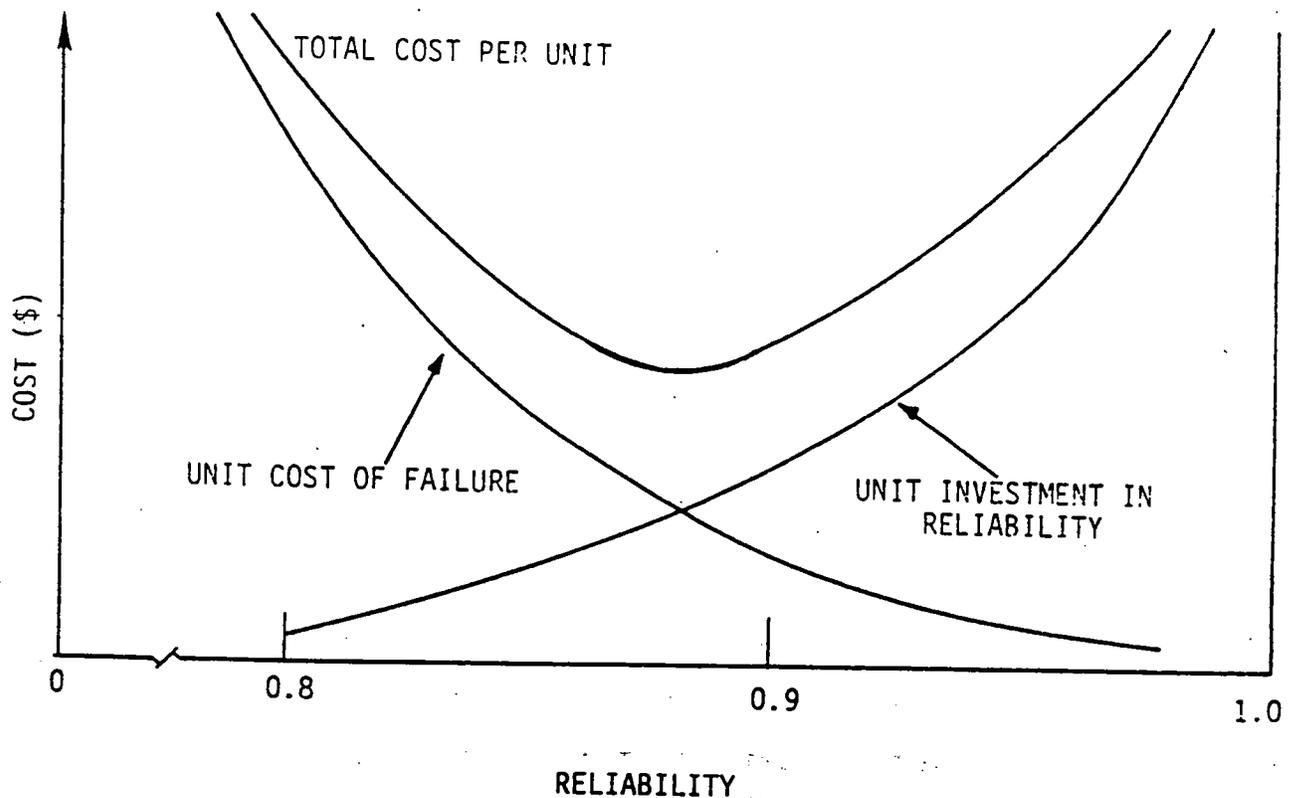


Figure L.3. Reliability of cost-trade-off considerations.

1. Design should be simple as possible.
2. Derating is used to achieve higher reliability (safety margin).
3. Burn-in if decreasing hazard rate in early life.
4. Redundancy (active, standby, and spares).
5. Protection from environmental or operational stresses--packaging and handling.
6. Ease of maintenance, service (actually an aid to increasing availability not reliability).
7. Use of appropriate screening tests for parts, components.
8. Specify replacement schedules (increases availability by preventive maintenance).

9. Conduct reliability qualification/demonstration tests.

Designing for reliability requires a cost-trade-off between unit cost of failure and the unit investment in reliability as illustrated by Figure L.3.

## L.2 LIFE TESTING

The life of an item is an important characteristic to the user and consequently life testing is often specified to demonstrate that a part, component or system has a mean life of a required number of hours. In order to assure with high confidence that a system or item is acceptable it may be required that  $n$  systems be tested, for example, until either

1.  $r$  systems have failed
2. The test time does not exceed a specified time, for example, 100 hours.

Thus, the test may be failure terminated, (1) above, or time terminated, (2) above. The number of failures  $r$  may be put equal to  $n$ , in which case the test is run until all items have failed. This is usually not practical due to the expense of testing and the consequent time delay in reaching a decision.

The definition of failure may be a sudden discontinuity of operation or a gradual degradation of the performance of the item as measured by pertinent performance parameters. Life of an item is thus defined as the time that the item continues to perform satisfactorily under specified conditions of operation.

The information required to plan a life test consists of

1. The sample size - number of items to be placed on test
2. The testing approach (whether failure-or time-terminated)
3. The decision of whether to replace failed items.
4. The termination criteria (failure criteria).

5. The degree of precision (e.g., the desired statistical confidence in the case of estimating the mean life of the item).

6. The model assumed for the distribution of failure times (e.g., exponential, Weibull, lognormal).

Test plans have been developed for life testing (acceptance sampling) when the failure time distribution is the negative exponential (MIL-STD-781)<sup>4</sup> and other distributions. The negative exponential implies a constant hazard rate (i.e., the time-to-failure of an item is independent of the time that it has been on test). On the other hand, the Weibull distribution can be applied when the item has a decreasing hazard rate (early life, debugging phase) or increasing hazard rate (wearout) phase at end of life), see the bathtub curve of Figure L.1. It is not possible to treat the many problems of testing in this short appendix, thus the reader is referred to one of the many texts on the subject.<sup>1,2,3</sup>

Example

Suppose that a record is maintained for the hours of operation of motor brushes for a hi-vol sampler motor. Assume the following data for 15 motor brushes.

<u>Motor brush</u>	<u>Hours before replacement</u>
1	600
2	696
3	782
4	576
5	648
6	576
7	806
8	720
9	624
10	504
11	408
12	384
13	648
14	744
15	432

Determine the mean life in hours before replacement and the likelihood that a failure (wear-out requiring replacement) will occur prior to 400 h of operation.

The mean life of the motor brushes is 610 h. The chances of failure prior to 400 h is  $1/15 = 0.067$  or 6.7% based on the one observation less than 400 h. This estimate does not assume a specific distribution form. If one assumes a normal distribution and uses the observed mean ( $\bar{X}$ ) and standard deviation (s), a second estimate can be obtained. For these data  $\bar{X} = 610$ ,  $s = 132$ , thus

$$z = \frac{400 - \bar{X}}{s} = \frac{400 - 610}{132} = -1.59$$

and the probability that a value less than 400 occurs is estimated by 0.056 or 5.6%.

This is a very simple example of the value of maintaining records of the life of an equipment or a component of an equipment. The mean life can be used to project spare and repair requirements, to compare different equipments/components, and to establish a replacement policy (time to check for wear out or possibly a preventive maintenance schedule).

### L.3 REFERENCES

1. Juran, J. M., Quality Control Handbook, Third Edition, McGraw Hill Book Co., New York, Section 8. (Also see Section 25, pp. 25-41, for a discussion of life testing and reliability.
2. Enrick, N. L., Quality Control and Reliability, Sixth Edition, Industrial Press, Inc., New York, 1972.
3. Myers, R. H., Wong, K. L., and Gordy, H. M., Reliability Engineering for Electronic Systems, John Wiley & Sons, Inc., New York, 1964.
4. MIL-STD-781, Reliability Tests, Exponential Distribution, 1967, Department of Defense, Washington, D. C.

## APPENDIX M

### INTERIM GUIDELINES AND SPECIFICATIONS FOR PREPARING QUALITY ASSURANCE PROJECT PLANS

#### M.1 INTRODUCTION

Environmental Protection Agency (EPA) policy requires participation by all EPA regional offices, program offices, EPA laboratories and States in a centrally-managed QA program as stated in the Administrator's Memorandum of May 30, 1979. This requirement applies to all environmental monitoring and measurement efforts mandated or supported by EPA through regulations, grants, contracts, or other formalized means not currently covered by regulation. The responsibility for developing, coordinating and directing the implementation of this program has been delegated to the Office of Research and Development (ORD), which has established the Quality Assurance Management Staff (QAMS) for this purpose.

Each office or laboratory generating data has the responsibility to implement minimum procedures which assure that precision, accuracy, completeness, and representativeness of its data are known and documented. In addition, an organization should specify the quality levels which data must meet in order to be acceptable. To ensure that this responsibility is met uniformly across the Agency, each EPA Office or Laboratory must have a written QA Project Plan covering each monitoring or measurement activity within its purview.

#### M.2 DEFINITION, PURPOSE, AND SCOPE

##### M.2.1 Definition

QA Project Plans are written documents, one for each specific project or continuing operation (or group of similar

projects or continuing operations), to be prepared by the responsible Program Office, Regional Office, Laboratory, Contractor, Grantee, or other organization. The QA Project Plan presents, in specific terms, the policies, organization, objectives, functional activities, and specific QA and QC activities designed to achieve the data quality goals of the specific project(s) or continuing operation(s). Other terms useful in understanding this guideline are defined in Appendix A of this volume.

#### M.2.2 Purpose

This document (1) presents guidelines and specifications that describe the 16 essential elements of a QA Project Plan, (2) recommends the format to be followed, and (3) specifies how plans will be reviewed and approved.

#### M.2.3 Scope

The mandatory QA program covers all environmentally-related measurements. Environmentally-related measurements are defined as all field and laboratory investigations that generate data. These include (1) the measurement of chemical, physical, or biological parameters in the environment, (2) the determination of the presence or absence of pollutants in waste streams, (3) assessment of health and ecological effect studies, (4) conduct of clinical and epidemiological investigations, (5) performance of engineering and process evaluations, (6) study of laboratory simulation of environmental events, and (7) study or measurement on pollutant transport and fate, including diffusion models. Each project within these activities must have a written and approved QA Project Plan.

### M.3 PLAN PREPARATION AND RESPONSIBILITIES

#### M.3.1 Document Control

All Quality Assurance Project Plans must be prepared using a document control format consisting of information placed in

the upper right-hand corner of each document page (Section 1.4.1 of this Volume):

1. Section number
2. Revision number
3. Date (of revision)
4. Page.

### M.3.2 Elements of QA Project Plan

Each of the sixteen items listed below must be considered for inclusion in each QA Project Plan:

1. Title page with provision for approval signatures
2. Table of contents
3. Project description
4. Project organization and responsibility
5. QA objectives for measurement data in terms of precision, accuracy, completeness, representativeness and comparability
6. Sampling procedures
7. Sample custody
8. Calibration procedures and frequency
9. Analytical procedures
10. Data reduction, validation, and reporting
11. Internal quality control checks and frequency
12. Performance and system audits and frequency
13. Preventive maintenance procedures and schedules
14. Specific routine procedures to be used to assess data precision, accuracy and completeness of specific measurement parameters involved
15. Corrective action
16. Quality assurance reports to management.

It is Agency policy that precision and accuracy of data shall be assessed on all monitoring and measurement projects. Therefore, Item 14 must be described in all Quality Assurance Project Plans.

### M.3.3 Responsibilities

M.3.3.1 Intramural Projects - Each Project Officer working in close coordination with the QA Officer is responsible for the preparation of a written QA Project Plan for each intramural project that involves environmental measurements. This written plan must be separate from any general plan normally prepared for the project (see caveat presented in Section M.6). The Project Officer and the QA Officer must ensure that each intramural project plan contains procedures to document and report precision, accuracy and completeness of all data generated.

M.3.3.2 Extramural Projects - Each Project Officer working in close coordination with the QA Officer has the responsibility to see that a written QA Project Plan is prepared by the extramural organization for each project involving environmental measurements. The elements of the QA Project Plan must be separately identified from any general plan normally prepared for the project (see caveat presented in Section M.6). The Project Officer and the QA Officer must ensure that each extramural project plan contains procedures to document and report precision, accuracy and completeness of all data generated.

### M.4 PLAN REVIEW, APPROVAL, AND DISTRIBUTION

#### M.4.1 Intramural Projects

Each QA Project Plan must be approved by the Project Officer's immediate supervisor and the QA Officer. Completion of reviews and approvals is shown by signatures on the title page of the plan. Environmental measurements may not be initiated until the QA Project Plan has received the necessary approvals. A copy of the approved QA Project Plan will be distributed by the Project Officer to each person who has a major responsibility for the quality of measurement data.

#### M.4.2 Extramural Projects

Each QA Project Plan must be approved by the funding organization's Project Officer and the QA Officer. In addition, the

extramural organization's Project Manager and responsible QA official must review and approve the QA Project Plan. Completion of reviews and approvals is shown by signatures on the title page of the plan. Environmental measurements may not be initiated until the QA Project Plan has received the necessary approvals. A copy of the approved QA Project Plan will be distributed by the extramural organization's Project Director to each person who has a major responsibility for the quality of the measurement data.

#### M.5 PLAN CONTENT REQUIREMENTS

The sixteen (16) essential elements described in this section must be considered and addressed in each QA Project Plan. If a particular element is not relevant to the project under consideration, a brief explanation of why the element is not relevant must be included. EPA-approved reference, equivalent or alternative methods must be used and their corresponding Agency-approved guidelines must be applied wherever they are available and applicable.

It is Agency policy that precision and accuracy of data shall be assessed routinely and reported on all environmental monitoring and measurement data. Therefore, specific procedures to assess precision and accuracy on a routine basis during the project must be described in each QA Project Plan. Procedures to assess data quality are being developed by QAMS and the Environmental Monitoring Systems/Support Laboratories. Additional guidance can be obtained from QA handbooks for air, water biological, and radiation measurements (References 1, 2, 3, 12, 17, and 18).

The following subsections provide specific guidance pertinent to each of the 16 components which must be considered for inclusion in every QA Project Plan.

##### M.5.1 Title Page

At the bottom of the title page, provisions must be made for the signatures of approving personnel. As a minimum, the QA Project Plan must be approved by the following:

1. For intramural projects
  - a. Project Officer's immediate supervisor
  - b. QA Officer (QAO)
2. For extramural projects
  - a. Organization's Project Manager
  - b. Organization's responsible QA Official
  - c. Funding organization's Project Officer
  - d. Funding organization's QA Officer.

#### M.5.2 Table of Contents

The QA Project Plan Table of Contents will address each of the following items:

1. Introduction
2. A serial listing of each of the 16 quality assurance project plan components
3. A listing of any appendices which are required to augment the Quality Assurance Project Plan as presented (i.e., standard operating procedures, etc.).

At the end of the Table of Contents, list the QA official and all other individuals receiving official copies of the QA Project Plan and any subsequent revisions.

#### M.5.3 Project Description

Provide a general description of the project. This description may be brief but must have sufficient detail to allow those individuals responsible for review and approval of the QA Project Plan to perform their task. Where appropriate, include the following:

1. Flow diagrams, tables, and charts
2. Dates anticipated for start and completion
3. Intended end use of acquired data.

M.5.4 Project Organization and Responsibility

Include a table or chart showing the project organization and line authority. List the key individuals, including the QAO, who are responsible for ensuring the collection of valid measurement data and the routine assessment of measurement systems for precision and accuracy.

M.5.5 QA Objectives for Measurement Data in Terms of Precision, Accuracy, Completeness, Representativeness, and Comparability

For each major measurement parameter, including all pollutant measurement systems, list the QA objectives for precision, accuracy and completeness. These QA objectives will be summarized in a Table M.1.

TABLE M.1. EXAMPLE OF FORMAT TO SUMMARIZE PRECISION, ACCURACY AND COMPLETENESS OBJECTIVES

Measurement parameter (Method)	Reference	Experimental conditions	Precision, std. dev.	Accuracy	Completeness
NO <sub>2</sub> (Chemiluminescent)	EPA 650/4-75-011 February 1975	Atmospheric samples spiked with NO <sub>2</sub> as needed	<±10%	± 5%	90%
SO <sub>2</sub> (24 h) (Pararosaniline)	EPA 650/4-74-027 December 1973	Synthetic atmosphere	<±20%	±15%	90%

All measurements must be made so that results are representative of the media (air, water, biota, etc.) and conditions being measured. Unless otherwise specified, all data must be calculated and reported in units consistent with other organizations reporting similar data to allow comparability of data bases among organizations. Definitions for precision, accuracy and completeness are provided in Subsection M.10 and Appendix A.

Data quality objectives for accuracy and precision established for each measurement parameter will be based on prior knowledge of the measurement system employed, method validation studies using, for example, replicates, spikes, standards, calibrations, and recovery studies and on the requirements of the specific project.

#### M.5.6 Sampling Procedures

For each major measurement parameter(s), including all pollutant measurement systems, provide a description of the sampling procedures to be used. Where applicable, include the following:

1. Description of techniques or guidelines used to select sampling sites
2. Inclusion of specific sampling procedures to be used (by reference in the case of standard procedures and by actual description of the entire procedure in the case of nonstandard procedures)
3. Charts, flow diagrams or tables delineating sampling program operations
4. A description of containers, procedures, reagents, et-cetera, used for sample collection, preservation, transport, and storage
5. Special conditions for the preparation of sampling equipment and containers to avoid sample contamination (e.g., containers for organics should be solvent-rinsed; containers for trace metals should be acid-rinsed)
6. Sample preservation methods and holding times
7. Time considerations for shipping samples promptly to the laboratory
8. Sample custody or chain-of-custody procedures
9. Forms, notebooks and procedures to be used to record sample history, sampling conditions and analyses to be performed.

#### M.5.7 Sample Custody

Sample custody is a part of any good laboratory or field operation. Where samples may be needed for legal purposes, "chain-of custody" procedures, as defined by the Office of Enforcement, will be used. However, as a minimum, the following sample custody procedures will be addressed in the QA Project Plans:

##### 1. Field Sampling Operations:

- a. Documentation of procedures for preparation of reagents or supplies which become an integral part of the sample (e.g., filters, and absorbing reagents)
- b. Procedures and forms for recording the exact location and specific considerations associated with sample acquisition
- c. Documentation of specific sample preservation method
- d. Prepared sample labels containing all information necessary for effective sample tracking. Figure M.1 illustrates a typical sample label applicable to this purpose
- e. Standardized field tracking reporting forms to establish sample custody in the field prior to shipment. Figure M.2 presents a typical sample of a field tracking report form.

##### 2. Laboratory Operations:

- a. Identification of responsible party to act as sample custodian at the laboratory facility authorized to sign for incoming field samples, obtain documents of shipment (e.g., bill of lading number or mail receipt), and verify the data entered onto the sample custody records
- b. Provision for laboratory sample custody log consisting of serially numbered standard lab-tracking report forms. A typical sample of a standardized lab-tracking report form is shown in Figure M.3

(NAME OF SAMPLING ORGANIZATION)

Sample description \_\_\_\_\_  
 \_\_\_\_\_

Plant \_\_\_\_\_ Location \_\_\_\_\_  
 Date \_\_\_\_\_  
 Time \_\_\_\_\_  
 Media \_\_\_\_\_ Station \_\_\_\_\_  
 Sample type \_\_\_\_\_ Preservative \_\_\_\_\_

Sampled by \_\_\_\_\_

Sample ID number \_\_\_\_\_

Lab number \_\_\_\_\_

Remarks \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

Figure M.1. Example of General Sample Label.

W/O number _____		Page ____		
Field Tracking Report _____ (LOC-SN)				
Field sample code (FSC)	Brief description	Date	Time(s)	Sampler

Figure M.2. Sample of Field Tracking Report form.

W/O number _____		Page _____			
Lab-tracking report _____ (LOC-SN-FSC)					
Fraction code	X	Prep/anal required	Responsible individual	Date delivered	Date completed

Figure M.3. Sample of lab-tracking report form.

c. Specification of laboratory sample custody procedures for sample handling, storage and dispersment for analysis.

Additional guidelines useful in establishing a sample custody procedure are given in Section 2.0.6 of Reference 2, and Section 3.0.3 of Reference 3, and References 13 and 14.

M.5.8 Calibration Procedures and Frequency

Include calibration procedures and information:

1. For each major measurement parameter, including all pollutant measurement systems, reference the applicable standard operating procedure (SOP) or provide a written description of the calibration procedure(s) to be used
2. List the frequency planned for recalibration
3. List the calibration standards to be used and their source(s), including traceability procedures.

#### M.5.9 Analytical Procedures

For each measurement parameter, including all pollutant measurement systems, reference the applicable standard operating procedure (SOP) or provide a written description of the analytical procedure(s) to be used. Officially approved EPA procedures will be used when available. For convenience in preparing the QA Project Plan, Elements 6, 8, and 9 may be combined (e.g., Subsections M.5.6, M.5.8, and M.5.9).

#### M.5.10 Data Reduction, Validation, and Reporting

For each major measurement parameter, including all pollutant measurement systems, briefly describe the following:

1. The data reduction scheme planned on collected data, including all equations used to calculate the concentration or value of the measured parameter and reporting units

2. The principal criteria that will be used to validate data integrity during collection and reporting of data

3. The methods used to identify and treat outliers

4. The data flow or reporting scheme from collection of raw data through storage of validated concentrations. A flow-chart will usually be needed

5. Key individuals who will handle the data in this reporting scheme (if this has already been described under project organization and responsibilities, it need not be repeated here).

#### M.5.11 Internal Quality Control Checks

Describe and/or reference all specific internal quality control ("internal" refers to both laboratory and field activities) methods to be followed. Examples of items to be considered include:

1. Replicates

2. Spiked samples

3. Split samples
4. Control charts
5. Blanks
6. Internal standards
7. Zero and span gases
8. Quality control samples
9. Surrogate samples
10. Calibration standards and devices
11. Reagent checks.

Additional information and specific guidance can be found in References 17 and 18.

#### M.5.12 Performance and System Audits

Each project plan must describe the internal and external performance and system audits which will be required to monitor the capability and performance of the total measurement system(s).

The system audit consists of evaluation of all components of the measurement systems to determine their proper selection and use. This audit includes a careful evaluation of both field and laboratory quality control procedures. System audits are normally performed prior to or shortly after systems are operational; however, such audits should be performed on a regularly scheduled basis during the lifetime of the project or continuing operation. The on-site system audit may be a requirement for formal laboratory certification programs such as laboratories analyzing public drinking water systems. Specific references pertinent to system audits for formal laboratory certification programs can be found in References 19 and 20.

After systems are operational and generating data, performance audits are conducted periodically to determine the accuracy of the total measurement system(s) or component parts thereof. The plan should include a schedule for conducting performance audits for each measurement parameter, including a performance audit for all measurement systems. As part of the performance

audit process, laboratories may be required to participate in analysis of performance evaluation samples related to specific projects. Project plans should also indicate, where applicable, scheduled participation in all other interlaboratory performance evaluation studies.

In support of performance audits, the Environmental Monitoring Systems/Support Laboratories provide necessary audit materials and devices and technical assistance. Also, these laboratories conduct regularly scheduled interlaboratory performance tests and provide guidance and assistance in the conduct of system audits. To make arrangements for assistance in the above areas, these laboratories should be contacted directly:

Environmental Monitoring Systems Laboratory  
Research Triangle Park, NC 27711  
Attention: Director

Environmental Monitoring and Support Laboratory  
26 W. St. Clair Street  
Cincinnati, Ohio 45268  
Attention: Director

Environmental Monitoring Systems Laboratory  
P. O. Box 15027  
Las Vegas, NV 89114  
Attention: Director

#### M.5.13 Preventive Maintenance

The following types of preventive maintenance items should be considered and addressed in the QA Project Plan:

1. A schedule of important preventive maintenance tasks that must be carried out to minimize downtime of the measurement systems
2. A list of any critical spare parts that should be on hand to minimize downtime.

M.5.14 Specific Routine Procedures Used to Assess Data Precision, Accuracy and Completeness

It is Agency policy that precision and accuracy of data must be routinely assessed for all environmental monitoring and measurement data. Therefore, specific procedures to assess precision and accuracy on a routine basis on the project must be described in each QA Project Plan.

For each major measurement parameter, including all pollutant measurement systems, the QA Project Plan must describe the routine procedures used to assess the precision, accuracy and completeness of the measurement data. These procedures should include the equations to calculate precision, accuracy and completeness, and the methods used to gather data for the precision and accuracy calculations.

Statistical procedures applicable to environmental projects are found in Appendices A through L of this Volume and in References 2, 3, 12, 17, and 18. Examples of these procedures include:

1. Central tendency and dispersion (e.g., arithmetic mean, range, standard deviation, relative standard deviation, pooled standard deviation, and geometric mean)
2. Measures of variability (e.g., accuracy, bias, precision; within laboratory and between laboratories)
3. Significance test (e.g., u-test, t-test, F-test, and Chi-square test)
4. Confidence limits
5. Testing for outliers.

Recommended guidelines and procedures to assess data precision, accuracy and completeness are being developed.

M.5.15 Corrective Action

Corrective action procedures must be described for each project which include the following elements:

1. The predetermined limits for data acceptability beyond which corrective action is required
2. Procedures for corrective action
3. For each measurement system, identify the responsible individual for initiating the corrective action and also the individual responsible for approving the corrective action, if necessary.

Corrective actions may also be initiated as a result of other QA activities, including:

1. Performance audits
2. System audits
3. Laboratory/field comparison studies
4. QA Program audits conducted by QAMS.

A formal corrective action program is more difficult to define for these QA activities in advance and may be defined as the need arises.

#### M.5.16 Quality Assurance Reports to Management

QA Project Plans should provide a mechanism for periodic reporting to management on the performance of measurement systems and data quality. As a minimum, these reports should include:

1. Periodic assessment of measurement data accuracy, precision and completeness
2. Results of performance audits
3. Results of system audits
4. Significant QA problems and recommended solutions.

The individual(s) responsible for preparing the periodic reports should be identified. The final report for each project must include a separate QA section which summarizes data quality information contained in the periodic reports.

## M.6 QUALITY ASSURANCE PROJECT PLANS VERSUS PROJECT WORK PLANS

This document provides guidance for the preparation of QA Project Plans and describes 16 components which must be included. Historically, most project managers have routinely included the majority of these 16 elements in their project work plans. In practice, it is frequently difficult to separate important quality assurance and quality control functions and to isolate these functions from technical performance activities. For those projects where this is the case, it is not deemed necessary to replicate the narrative in the Quality Assurance Project Plan section.

In instances where specific QA/QC protocols are addressed as an integral part of the technical work plan, it is only necessary to cite the page number and location in the work plan in the specific subsection designated for this purpose.

It must be stressed, however, that whenever this approach is used a "QA Project Plan locator page" must be inserted into the project work plan immediately following the table of contents. This locator page must list each of the items required for the QA Project Plan and state the section and pages in the project plan where the item is described. If a QA Project Plan item is not applicable to the work plan in question, the words "not applicable" should be inserted next to the appropriate component on the locator page and the reason why this component is not applicable should be briefly stated in the appropriate subsection in the QA Project Plan.

## M.7 STANDARD OPERATING PROCEDURES

A large number of laboratory and field operations can be standardized and written as SOP. When such procedures are applicable and available, they may be incorporated into the QA Project Plan by reference.

QA Project Plans should provide for the review of all activities which could directly or indirectly influence data quality and the determination of those operations which must be covered by SOP's. Examples are:

1. General network design
2. Specific sampling site selection
3. Sampling and analytical methodology
4. Probes, collection devices, storage containers, and sample additives or preservatives
5. Special precautions, such as heat, light, reactivity, combustibility, and holding times
6. Federal reference, equivalent or alternative test procedures
7. Instrumentation selection and use
8. Calibration and standardization
9. Preventive and remedial maintenance
10. Replicate sampling
11. Blind and spiked samples
12. Collocated samplers
13. QC procedures such as intralaboratory and intrafield activities, and interlaboratory and interfield activities
14. Documentation
15. Sample custody
16. Transportation
17. Safety
18. Data handling procedures
19. Service contracts
20. Measurement of precision, accuracy, completeness, representativeness, and comparability
21. Document control.

#### M.8 SUMMARY

Each intramural and extramural project that involves environmental measurements must have a written and approved QA Project Plan. All 16 items described previously must be considered and addressed. Where an item is not relevant, a brief explanation of why it is not relevant must be included. It is Agency policy that precision and accuracy of data must be routinely assessed and reported on all environmental monitoring and

measurement data. Therefore, specific procedures to assess precision and accuracy on a routine basis during the project must be described in each QA Project Plan.

#### M.9 EXAMPLE OF PROJECT PLAN

For the convenience of the reader the following pages of this section contains an example of a QA project plan for ambient air monitoring. The format is retained as one would prepare a plan and hence not necessarily consistent with the Handbook format. The only exception is that the documentation is given on each page consistent with the Handbook.

M.9.1 Project Plan for Ambient Air Monitoring

A MODEL QA PROJECT PLAN

AMBIENT AIR MONITORING STUDY AROUND THE WEPCO POWER PLANT

QA PROJECT PLAN FOR IN-HOUSE PROJECT

APPROVAL:

EPA Project Officer: Thomas Swift Date 5/1/80  
EPA Supervisor: Thomas Smith Date 5/5/80  
EPA QA Officer: Harold Juff Date 5/10/80

TABLE OF CONTENTS

<u>Section</u>	<u>DESCRIPTION</u>	<u>Page</u>
1	Project Description	22
2	Project Organization and Responsibilities	22
3	QA Objectives in Terms of Precision, Accuracy, Completeness, Representativeness and Comparability	22
4	Sampling and Analysis Procedures	24
5	Sample Custody	24
6	Calibration Procedures	24
7	Data Analysis, Validation, and Reporting	25
8	Internal Quality Control Checks	25
9	Performance and System Audits	27
10	Preventive Maintenance	28
11	Specific Procedures to be Used to Routinely Assess Data Precision, Accuracy, and Completeness	28
12	Corrective Action	30
13	Quality Assurance Reports to Management	30

Distribution of Approved QA Project Plan:

1. Harold Smooth, QAD, EMSL/RTP
2. Thomas Swift, EMD, EMSL/RTP
3. Thomas Tuff, EMD, EMSL/RTP
4. Mary Pickford, EMD, EMSL/RTP
5. Mike Evans, EMD, EMSL/RTP
6. Gregory Thomas, QAD, EMSL/RTP
7. Ralph Niceguy, WEPCO

1. Project Description

The WEPCO power plant, located at Somewhere, Virginia, initiated a 12-mo ambient air monitoring project on April 1, 1980, to collect air quality data necessary for a construction permit for a new 200 meg-watt coal-fired boiler. WEPCO has established a monitoring network for total suspended particulates (TSP), SO<sub>2</sub> and NO<sub>2</sub> around the existing location where the new boiler will be constructed. EPA has received permission from WEPCO to monitor for TSP, SO<sub>2</sub> and NO<sub>2</sub> at WEPCO monitoring sites 2 and 5 for six mo starting July 1, 1980. Both WEPCO and EPA monitoring complies with monitoring and quality assurance requirements for Prevention of Significant Deterioration (PSD) monitoring. The purpose of the EPA study is to compare EPA and WEPCO results. In addition, EPA plans to compare the results from their continuous SO<sub>2</sub> monitors to results obtained by running the manual EPA Reference Method (pararosaniline method) every six days.

2. Project Organization and Responsibility

All EPA air monitoring and quality assurance will be performed by EPA personnel from the Environmental Monitoring Systems Laboratory, Research Triangle Park, North Carolina. The air monitoring will be performed by the Environmental Measurement Division (EMD) and the quality assurance by the Quality Assurance Division (QAD). The key personnel involved in the project, their project responsibility and line authority within EMSL are shown in Figure 1.

3. QA Objectives in Terms of Precision, Accuracy, Completeness, Representativeness and Comparability

All WEPCO sampling sites, including sites 2 and 5, were inspected by The State of Virginia Air Pollution Control Division and found to be valid and representative sampling sites. All 24-h integrated samples for TSP and SO<sub>2</sub> (by the Reference

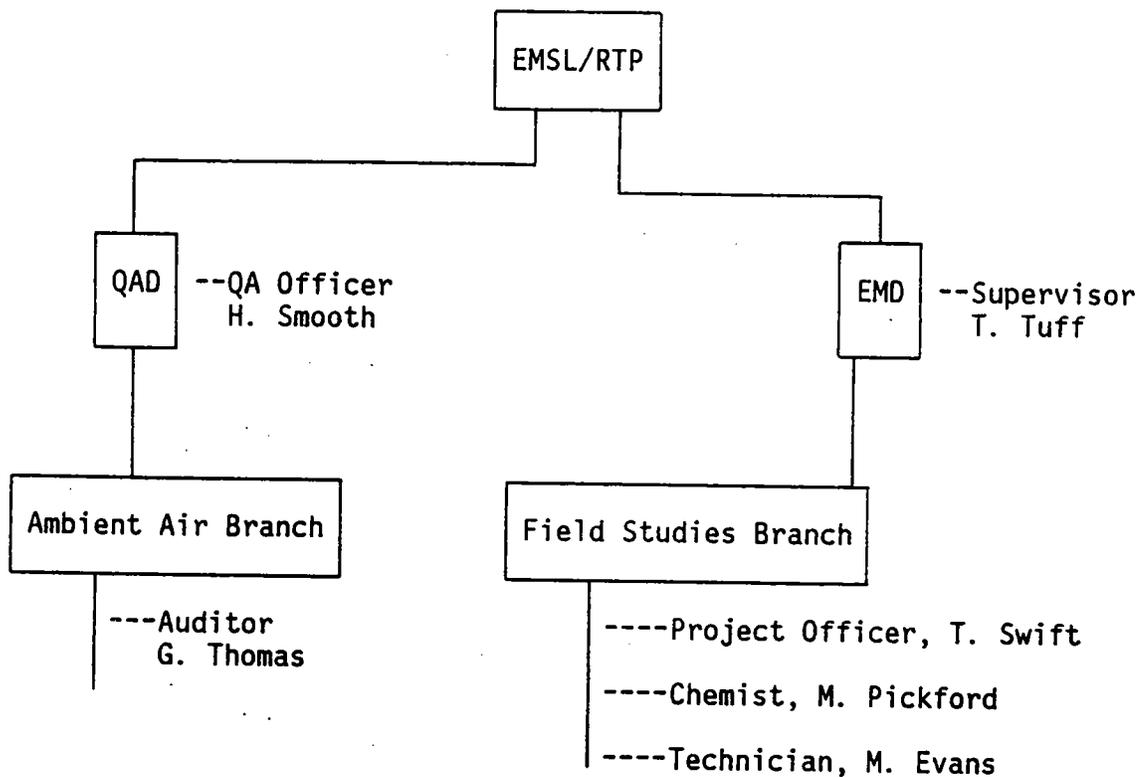


Figure 1. Project organization and responsibility.

Method) will be collected from midnight to midnight to correspond to calendar days. All results for TSP, SO<sub>2</sub> and NO<sub>2</sub> are calculated in µg/m<sup>3</sup> corrected to 25°C and 760 mm Hg so that results are comparable with WEPCO's data base.

The following QA objectives for precision, accuracy, and completeness have been used in the design of this study.

a. Completeness - Seventy-five (75) percent of all possible measurement data should be valid.

b. Accuracy - Each SO<sub>2</sub> and NO<sub>2</sub> continuous monitor results should agree within ±15 percent of audit concentration during each audit. Each SO<sub>2</sub> sample analysis audit for the SO<sub>2</sub> Reference Method should agree within the 90 percentile limits described in Section 2.1.8 of Volume II of this Handbook (EPA-600/

4-77-027a). Each TSP sampler flow audit should be within  $\pm 7$  percent of the audit flow value.

c. Precision - Current data are insufficient to give a good estimate for precision based on the quality assurance procedures required in Appendix B, 40 CFR 58 for PSD monitoring.

#### 4. Sampling and Analysis Procedures

All measurement methods used are EPA reference or equivalent methods. The following measurement methods will be used in this study.

a. Continuous SO<sub>2</sub> by Meloy SA185-2A flame photometric detector analyzers

b. Continuous NO<sub>2</sub> by Monitor Lab 8840 chemiluminescence analyzers

c. EPA Reference Method for SO<sub>2</sub> (pararosaniline method)

d. EPA Reference Method for TSP (Hi-Vol Method).

#### 5. Sample Custody

Since this is a research project, sample custody is not planned on this project.

#### 6. Calibration Procedures

All continuous monitors for SO<sub>2</sub> and NO<sub>2</sub> will be calibrated according to the manufacturer's recommended procedures and the recommendations in Section 2.0.9 of Volume II of this Handbook. Namely, each calibration shall include:

a. A zero concentration and three upscale concentrations equally spaced over the measurement range of 0 to 0.5 ppm

b. A daily Level 1 zero and span to be used to determine when recalibration is needed as per guidelines in Section 2.0.9 of Volume II of this Handbook.

Calibration and span gases for all continuous monitors for SO<sub>2</sub> and NO<sub>2</sub> shall be traceable to NBS, Standard Reference Materials using EPA Protocol No. 1 (Traceability Protocol for Establishing True Concentrations of Gases Used for Calibration

and Audits of Air Pollution Analyzers, Section 2.0.7 of Volume II). Specifically, cylinder gases of NO in N<sub>2</sub> at 50 ppm will be used for NO<sub>2</sub> monitors and SO<sub>2</sub> permeation tubes will be used for SO<sub>2</sub> monitors.

The calibration procedures described in the Reference Methods for TSP and SO<sub>2</sub> (pararosaniline method) will be followed. Recalibration shall be performed consistent with the guidance of Section 2.0.9 of Volume II of this Handbook.

#### 7. Data Analysis, Validation, and Reporting

The analysis and flow of data from the point of collection (raw data) through calculation and storage of validated concentrations (in µg/m<sup>3</sup>) is shown in Figure 2.

The SO<sub>2</sub> and NO<sub>2</sub> analyzers are calibrated in ppm. To convert ppm to µg/m<sup>3</sup> use the following equations:

$$\text{SO}_2 \text{ } \mu\text{g}/\text{m}^3 = \text{SO}_2 \text{ ppm} \times 2620$$

$$\text{NO}_2 \text{ } \mu\text{g}/\text{m}^3 = \text{NO}_2 \text{ ppm} \times 1880.$$

The equations for the calculation of SO<sub>2</sub> (pararosaniline bubbler method) and TSP concentrations are in the Reference Methods in Sections 2.1.6 and 2.2.6 of Volume II of this Handbook.

The principal criteria used to validate data are described in Subsection 9.1.4 of Section 2.0.9 for continuous methods (SO<sub>2</sub> and NO<sub>2</sub> analyzers) and Subsection 9.2.5 of Section 2.0.9 of Volume II of this Handbook for manual methods (TSP and SO<sub>2</sub> bubbler method).

#### 8. Internal Quality Control Checks and Frequency

The operational checks recommended in Section 2.0.9 of Volume II will be used in this project for internal quality control. A listing of the operational checks, the control limits for initiating corrective action, the planned corrective action, and the reference for more detailed description are shown in Figure 3.

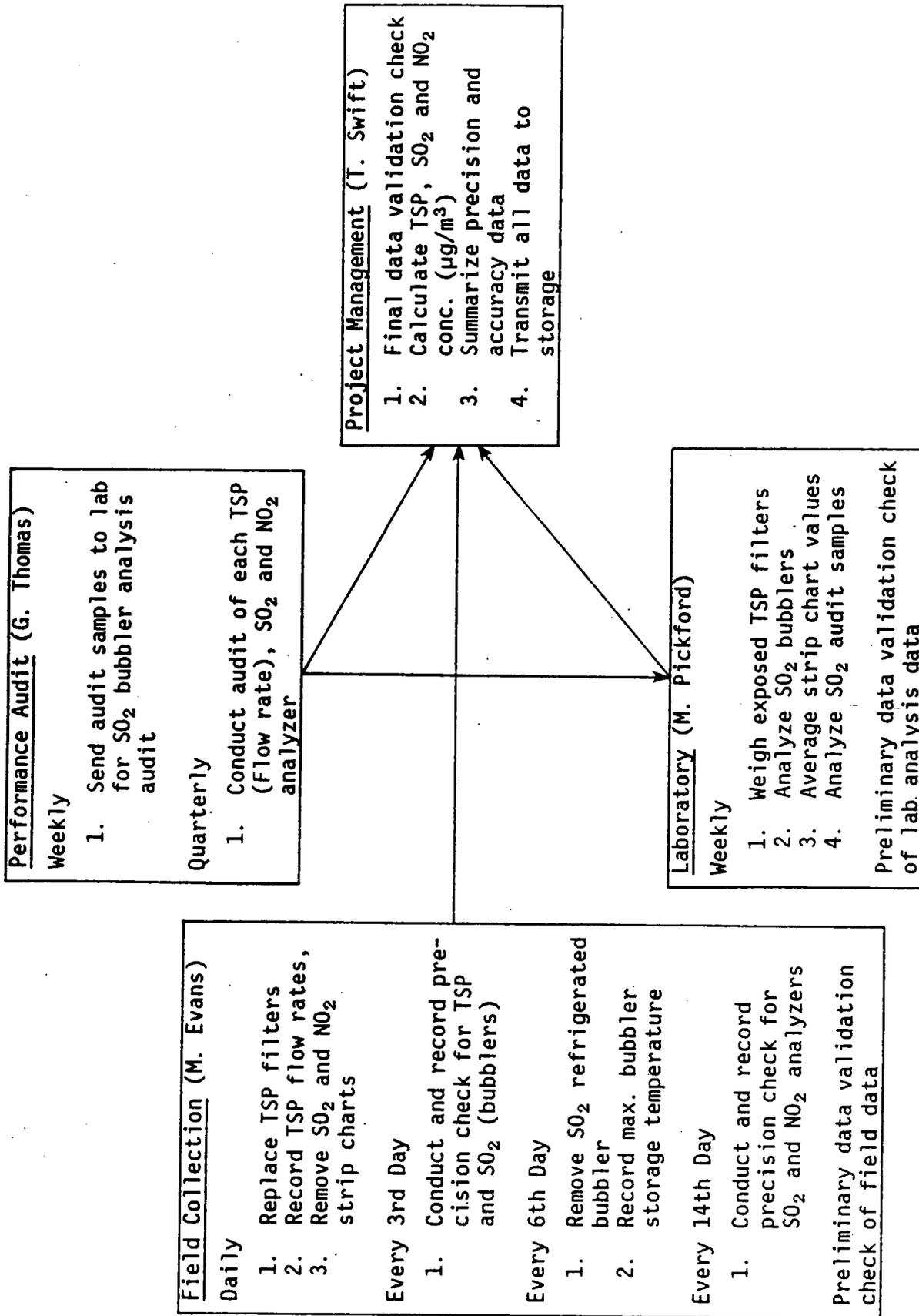


Figure 2. Data flow and analysis.

Measurement	Operation check	Control limit	Corrective action planned
Continuous SO <sub>2</sub> and NO <sub>2</sub>	daily level 1 span and zero drift check <sup>1</sup>	1. 3 std deviations 2. zero ±0.025 ppm 3. span ±15% 4. span ±25%	1. adjust analyzer 2. recalibrate 3. recalibrate 4. invalidate data
Manual SO <sub>2</sub> (Pararosaniline)	record bubbler temp during sampling and maintain low temp during shipment/storage <sup>1</sup>	temp must be between 5 and 25°C	invalidate sample
	sampling flow rate check each sample day <sup>1,2</sup>	±10%	invalidate sample
TSP	blank and standard solution each analysis day after every 10th sample <sup>1,2</sup>	1. blank absorbance ±0.03 units 2. std solution ±0.07 µg/ml	1. reanalyze previous 10 samples 2. reanalyze previous 10 samples
	sampling flow rate check each sample day <sup>1,3</sup>	±10%	recalibrate hi-vol sampler
	monthly reweigh a portion of exposed filters <sup>1,4</sup>	±5 mg	reweigh all exposed filters

<sup>1</sup>Section 2.0.9 of Volume II of QA Handbook.  
<sup>2</sup>Section 2.1.5 of Volume II of QA Handbook.  
<sup>3</sup>Section 2.2.4 of Volume II of QA Handbook.  
<sup>4</sup>Section 2.2.8 of Volume II of QA Handbook.

Figure 3. Internal quality control checks.

9. Performance and System Audits

Ambient air pollution measurements are scheduled to be initiated on July 1, 1980. A system audit is scheduled to be conducted during the week of June 23, 1980.

Performance audits to be conducted are the same type and on the same schedule as shown in Appendix B, 40 CFR 58 for PSD monitoring. Appendix B should be referred to for details. Briefly, the following performance audits and frequency will be conducted (based on Appendix B).

- a. Each continuous SO<sub>2</sub> and NO<sub>2</sub> analyzer will be audited quarterly with cylinder gases.
- b. For TSP, each hi-vol sampler will be audited quarterly at one flow rate between 40 and 60 cfm.
- c. For SO<sub>2</sub> bubbler samples, laboratory analyses will be audited each analysis day with one audit sample each in the range of 0.2 - 0.3, 0.5 - 0.6, and 0.8 - 0.9 µg SO<sub>2</sub>/ml. Note: This audit is described in Appendix A, not B, of 40 CFR 58.

#### 10. Preventive Maintenance

The preventive maintenance tasks and schedules recommended by the manufacturers of the SO<sub>2</sub> and NO<sub>2</sub> analyzers will be followed. The preventive maintenance recommended for TSP and the SO<sub>2</sub> Reference Method (bubblers) will be the same tasks and schedules described in Section 2.2.7 (for TSP) and Section 2.1.7 (for SO<sub>2</sub>) of Volume II of this Handbook.

The following spare materials will always be maintained on-hand during the project for daily checks and recalibrations:

- a. two extra SO<sub>2</sub> permeation tubes
- b. one extra zero cylinder gas
- c. one extra 50 ppm NO cylinder gas

#### 11. Specific Procedures to be Used to Routinely Assess Data Precision, Accuracy and Completeness

The results from performance audits described in Section 9 of this QA Project Plan are used to calculate accuracy for each measurement device. The audit frequency for each measurement device is also described in Section 9. The equations used to calculate accuracy are shown in: Appendix B of 40 CFR 58, for continuous SO<sub>2</sub> and NO<sub>2</sub>, and TSP; and Appendix A of 40 CFR 58 for

the SO<sub>2</sub> Reference Method. Example calculations for accuracy for each measurement device are shown in Section 2.0.8 of Volume II of this Handbook.

Precision check description and frequency for each measurement device is the same as shown in: Appendix B of 40 CFR 58 for continuous SO<sub>2</sub> and NO<sub>2</sub>, and TSP; and Appendix A of 40 CFR 58 for the SO<sub>2</sub> Reference Method. The results from these precision checks are used to calculate precision for each measurement device. The equations used to calculate precision are also shown in Appendices A and B. Example calculations for precision for each measurement device are shown in Section 2.0.8 of Volume II. A summary of the precision checks follows:

a. Each continuous SO<sub>2</sub> and NO<sub>2</sub> analyzer will be checked by the field operator every two weeks for span drift at a concentration between 0.08 and 0.09 ppm. Calculation of precision for each analyzer is based on a quarterly results.

b. The calculation of TSP precision is based on the operation of a second hi-vol sampler collocated at one of the two sites. This collocated sampler will be operated every third sampling day along with the regular hi-vol sampler. Calculation of TSP data precision is based on quarterly results and applies to both sampling sites.

c. The calculation of SO<sub>2</sub> precision for the Reference Method (bubbler technique) is based on the operation of a second bubbler system at one of the two sites. This collocated bubbler system will be operated every sixth day along with the regular bubbler system. Calculation of SO<sub>2</sub> data precision is based on quarterly results and applies to both sampling sites.

Data completeness will be calculated for each measurement device and is based on quarterly results. Completeness will be calculated as a percentage of valid data compared to the amount of data expected to be obtained under normal operations.

12. Corrective Action

Figure 3 describes internal quality control checks planned for each measurement. Control limits and planned corrective actions are also shown in Figure 3. The authority to conduct the planned corrective action when the control limits are exceeded is given to M. Evans for field operations and M. Pickford for laboratory operations.

13. Quality Assurance Reports to Management

Within 15 days following the end of the calendar quarter, precision, accuracy and completeness will be reported on each measurement system to: T. Tuff, Supervisor, EMD, EMSL/RTP; H. Smooth, EPA Project Officer; and R. Niceguy, WEPCO.

M.10 GLOSSARY OF TERMS

This glossary is specialized for the needs of developing QA project plans. The definitions do not agree precisely with those in Appendix A of this volume of the Handbook; however, they do agree in substance. One should refer to Appendix A for additional definitions or further information concerning the following definitions.

Audit - A systematic check to determine the quality of operation of some function or activity. Audits may be of two basic types: (1) performance audits in which quantitative data are independently obtained for comparison with routinely obtained data in a measurement system, or (2) system audits of a qualitative nature that consist of an on-site review of a laboratory's quality assurance system and physical facilities for sampling, calibration, and measurement.

Data Quality - The totality of features and characteristics of data that bears on their ability to satisfy a given purpose. The characteristics of major importance are accuracy, precision, completeness, representativeness, and comparability. These five characteristics are defined as follows:

1. Accuracy - the degree of agreement of a measurement X with an accepted reference or true value, T, usually expressed as the difference between the two values, X-T, or the difference as a percentage of the reference or true value,  $100 (X-T)/T$ , and sometimes expressed as a ratio, X/T.

2. Precision - a measure of mutual agreement among individual measurements of the same property, usually under prescribed similar conditions. Precision is best expressed in terms of the standard deviation. Various measures of precision exist depending upon the "prescribed similar conditions."

3. Completeness - a measure of the amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct normal conditions.

4. Representativeness - expresses the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition.

5. Comparability - expresses the confidence with which one data set can be compared to another.

Data Validation - A systematic process for reviewing a body of data against a set of criteria to provide assurance that the data are adequate for their intended use. Data validation consists of data editing, screening, checking, auditing, verification, certification, and review.

Environmentally Related Measurements - A term used to describe essentially all field and laboratory investigations that generate data involving (1) the measurement of chemical, physical, or biological parameters in the environment, (2) the determination of the presence or absence of criteria or priority pollutants in waste streams, (3) assessment of health and ecological effect studies, (4) conduct of clinical and epidemiological investigations, (5) performance of engineering and process evaluations, (6) study of laboratory simulation of environmental events, and

(7) study or measurement on pollutant transport and fate, including diffusion models.

Performance Audits - Procedures used to determine quantitatively the accuracy of the total measurement system or component parts thereof.

Quality Assurance - The total integrated program for assuring the reliability of monitoring and measurement data. A system for integrating the quality planning, quality assessment, and quality improvement efforts to meet user requirements.

Quality Assurance Program Plan - An orderly assemblage of management policies, objectives, principles, and general procedures by which an agency or laboratory outlines how it intends to produce data of known and accepted quality.

Quality Assurance Project Plan - An orderly assembly of detailed and specific procedures which delineates how data of known and accepted quality are produced for a specific project. (A given agency or laboratory would have only one quality assurance program plan, but would have a quality assurance project plan for each of its projects.)

Quality Control - The routine application of procedures for obtaining prescribed standards of performance in the monitoring and measurement process.

Standard Operating Procedure (SOP) - A written document which details an operation, analysis or action whose mechanisms are thoroughly prescribed and which is commonly accepted as the method for performing certain routine or repetitive tasks.

#### M.11 REFERENCES

1. Quality Assurance Handbook for Air Pollution Measurement Systems. Vol. I - Principles. EPA-600/9-76-005, March 1976.
2. Quality Assurance Handbook for Air Pollution Measurement Systems. Vol. II - Ambient Air Specific Methods. EPA-600/4-77-027a. May 1977.

3. Quality Assurance Handbook for Air Pollution Measurement Systems. Vol. III - Stationary Source Specific Methods. EPA-600/4-77-027b. August 1977.
4. Systems Audit Criteria and Procedures for Ambient Air Monitoring Programs. Section 2.0.11, Vol. II, QA Handbook. Currently under development and available from address shown in Reference 1 after July 1, 1980.
5. Techniques to Evaluate Laboratory Capability to Conduct Stack Testing.
6. Performance Audit Procedures for Ambient Air Monitoring Programs. Section 2.0.12, Vol. II.
7. Appendix A - Quality Assurance Requirements for State and Local Air Monitoring Stations (SLAMS). Federal Register, Vol. 44, No. 92, pp. 27574-81. May 10, 1979.
8. Appendix B - Quality Assurance Requirements for Prevention of Significant Deterioration (PSD) Air Monitoring. Federal Register, Vol. 44, No. 92, pp. 27582-84. May 10, 1979.
9. Appendix F - Procedure 1 - Quality Assurance Requirements for Gas Continuous Emission Monitoring Systems (CEMS) for Compliance. To be submitted as a proposed regulation to amend 40 CFR 60.
10. Test Methods for Evaluating Solid Waste - Physical/Chemical Methods. EPA SW-846. 1980.
11. Quality Assurance Guidelines for IERL-CI Project Officers. EPA-600/9-79-046. December 1979.
12. Handbook for Analytical Quality Control in Water and Wastewater Laboratories. EPA-600/4-79-019. March 1979.
13. NEIC Policies and Procedures Manual. Office of Enforcement. EPA-330-9-78-001, May 1978.
14. NPDES Compliance, Sampling and Inspection Manual. Office of Water Enforcement, Compliance Branch, June 1977.
15. Juran, J. M. (ed), Quality Control Handbook. Third Edition, McGraw Hill, New York. 1974.
16. Juran, J. M. and F. M. Gryna. Quality Planning and Analysis. McGraw Hill, New York. 1970.
17. Handbook for Analytical Quality Control and Radioactivity Analytical Laboratories. EPA-600/7-77-088. August 1977

18. Manual of Analytical Quality Control for Pesticides and Related Compounds in Human and Environmental Samples. EPA-600/1-79-008. January 1979.
19. Procedure for the Evaluation of Environmental Monitoring Laboratories. EPA 600/4-78-78-017. March 1978.
20. Manual for the Interim Certification of Laboratories Involved in Analyzing Public Drinking Water Supplies - Criteria and Procedures. EPA 600/8-78-008. August 1978.